Philosophy & Technology (2021) 34:1581–1606
https://doi.org/10.1007/s13347-021-00476-1

RESEARCH ARTICLE

Good Proctor or "Big Brother"? Ethics of Online Exam Supervision Technologies

Simon Coghlan[1,2] · Tim Miller[1,2] · Jeannie Paterson[2,3]

Received: 10 November 2020 / Accepted: 24 August 2021 / Published online: 31 August 2021

Abstract

Online exam supervision technologies have recently generated significant controversy and concern. Their use is now booming due to growing demand for online courses and for off-campus assessment options amid COVID-19 lockdowns. Online proctoring technologies purport to effectively oversee students sitting online exams by using artificial intelligence (AI) systems supplemented by human invigilators. Such technologies have alarmed some students who see them as a "Big Brother-like" threat to liberty and privacy, and as potentially unfair and discriminatory. However, some universities and educators defend their judicious use. Critical ethical appraisal of online proctoring technologies is overdue. This essay provides one of the first sustained moral philosophical analyses of these technologies, focusing on ethical notions of academic integrity, fairness, non-maleficence, transparency, privacy, autonomy, liberty, and trust. Most of these concepts are prominent in the new field of AI ethics, and all are relevant to education. The essay discusses these ethical issues. It also offers suggestions for educational institutions and educators interested in the technologies about the kinds of inquiries they need to make and the governance and review processes they might need to adopt to justify and remain accountable for using online proctoring technologies. The rapid and contentious rise of proctoring software provides a fruitful ethical case study of how AI is infiltrating all areas of life. The social impacts and moral consequences of this digital technology warrant ongoing scrutiny and study.

Keywords Ethics · Artificial intelligence · Universities · Machine learning · Online assessment · Proctoring

✉ Simon Coghlan
simon.coghlan@unimelb.edu.au

[1] School of Computing and Information Systems, The University of Melbourne, Melbourne, Australia

[2] Centre for AI and Digital Ethics (CAIDE), The University of Melbourne, Melbourne, Australia

[3] Melbourne Law School, The University of Melbourne, Melbourne, Australia

1 Introduction

Recently, online exam supervision technologies have been thrust into the public spotlight due to growing demand for online courses in higher education (Ginder et al., 2019). On top of that, there have been huge global surges in online proctoring during lockdowns in the COVID-19 pandemic (Flaherty, 2020). Although educational institutions can supervise remote exam-takers simply by watching live online video such as via Zoom, online proctoring (OP) software programs offer more sophisticated monitoring functions which can be human-led and/or automated. Some now expect that OP technology will become the "new normal" in higher education around the world (Selwyn et al., 2021). Indeed, many institutions and students have already accepted them.

Yet such technologies have also generated controversy, confusion, media interest, and even legal action against a critic (Mullin, 2021). Some students have vigorously protested being compelled to submit to such monitoring (White, 2020). Concerns have even reached the higher levels of politics, with several US senators raising worries about the discriminatory potential of the software (Chin, 2021b). Due to its contentious and sensitive nature, even researching the subject has proved difficult. For example, when Selwyn et al. interviewed Australian students, activists, and university staff, they discovered some wariness about freely expressing their views (Selwyn et al., 2021). Some universities have defended their use of OP technologies, claiming they are sufficiently safe and sometimes necessary for students to complete their degrees. Others, such as the University of Illinois in Urbana Champaign, have retreated from their initial plans for using them as a result of opposition (Chin, 2021b; White, 2020). Still others, including Oxford and Cambridge, reportedly rejected them earlier in the pandemic (Clausen, 2020). As time goes on, institutions may of course alter their attitude towards the technology.

At the root of disagreement between concerned students and some universities are questions about the ethics of OP technologies. These ethical questions have not been comprehensively studied. This essay sets outs to explore them in philosophical detail and, moreover, to do so by drawing on the field of AI ethics. The paper should assist students, educators, and institutions in making informed judgments about the appropriateness of OP systems and the safeguards needed to protect the interests of students and educators alike. But the exploration has wider implications and meaning. Online proctoring illustrates how novel digital technologies including artificial intelligence (AI) can impact in new and interesting ways, and for better and for worse, on many different aspects of our lives. Our analysis also begins to illuminate some respects in which emerging digital technologies could affect educational practice and indeed society and cultural values more broadly (Selwyn et al., 2021).

OP software platforms first emerged in 2008 (ProctorU, 2020b) and are now booming. A 2020 poll found that 54% of higher educational institutions use them (Grajek, 2020). Reduced access to campuses due to COVID-19 has greatly elevated their attractiveness for universities and some students (Chin, 2021b).

Consequently, the financial value of OP software is predicted to jump significantly in coming years (Partners, 2021). Increasingly, OP software contains AI and machine learning components that analyze exam recordings to identify suspicious examinee behaviors or suspicious items in their immediate environment. OP companies, which can make good profits from their products (Chin, 2020), claim that automating proctoring increases the scalability, efficiency, and accuracy of exam supervision and the detection of cheating. These features have an obvious attraction for universities, some of which believe that the benefits of OP technologies outweigh any drawbacks. However, the complexity and opacity of OP technologies, especially their automated AI functions (Hagendorff, 2020), can be confusing. Furthermore, some (though not all) students complain of a "creepy" Big Brother sense of being invaded and surveilled (Hubler, 2020). Predictably, some bloggers are instructing students how to bluff proctoring platforms (Binstein, 2015).

Scholars have only just begun exploring remote and automated proctoring from a range of viewpoints, including pedagogical, behavioral, psychological, and technical perspectives (Asep & Bandung, 2019; Cramp et al., 2019; González-González et al., 2020). Nonetheless, and despite vigorous ethical discussion in regular media (Zhou, 2020), blog posts (Torino, 2020), and social media, the ethics of emerging OP technologies has so far received limited scholarly analysis (but see Dawson, 2020; Selwyn et al., 2021; Swauger, 2020b) and no sustained moral philosophical treatment. We aim to fill that gap. Although moral assessments can be informed by empirical data about online and in-person proctoring—such as data about test-taker behavior (Rios & Liu, 2017) and grade comparisons (Goedl & Malla, 2020)—moral assessments depend crucially on philosophical analysis. In the following ethical analysis, we identify and critically explore the key notions of academic integrity, fairness, non-maleficence, transparency, privacy, autonomy, liberty, and trust as they apply to OP technologies.

Some of these ethical notions are prominent in the new field of AI ethics (Jobin et al., 2019), which is burgeoning as AI penetrates increasingly into various facets of our lives, including education. In this paper, we suggest that OP platforms are neither a silver bullet for remote invigilation nor, as some would have it, a completely "evil" technology (Grajek, 2020). It is difficult to judge if and when OP technologies are justified, not least when students are effectively compelled to submit to them (Selwyn et al., 2021). Although students may sometimes be offered alternative exam arrangements, there is often limited capacity for students to give genuine, informed consent.

We shall suggest that OP technologies may be justified in some cases—if the ethical issues we discuss are adequately addressed and those using the platforms are properly accountable. However, institutions will also need to confront some wider possible implications of choosing to use OP technologies. This ethical analysis will help to inform concerned parties (students, educators, universities, etc.) while setting out important ethical considerations for educational institutions who are considering OP platforms, including how they might devise appropriate governance frameworks and processes to ensure that they remain accountable for their decisions. It will also provide a context for various future empirical and theoretical investigations of OP technologies.

The essay is structured as follows. The Background section provides context concerning exam invigilation and the central technological capabilities of popular OP programs. The Philosophical Approach section identifies and explains central moral principles and values relevant to the OP debate. The Discussion section applies in detail the selected principles and values to the technology. Finally, the Conclusion summarizes the ethical lessons for educational institutions and others and suggests questions for them and for further research.

2 Background

Because we are providing one of the first detailed ethical analyses of OP software, we need to give some relevant background concerning education, assessment, technology, and OP platforms. Digital technologies are used in education in various ways. Plagiarism detection tools like Turnitin are widely available for uncovering academic dishonesty and teaching good academic practice. Emerging AI teaching systems can adapt to the learning needs of individual pupils (Bartneck et al., 2021). AI-based predictions of student performance (Sweeney et al., 2015) have been used to create summative grades (Hern, 2020)**.** This has sometimes caused controversy. For example, a public scandal erupted in the UK when distressed students objected to the way their grades were predicted by an algorithm after exams were cancelled during COVID-19 (Simonite, 2020). Nonetheless, many technologies are being increasingly adopted in higher education.

Examinations have a long history in both the West and the East. Written public examinations first took place in Imperial China. Centuries later, exams in academia became established in British universities (Kellaghan & Greaney, 2019); their advent in the 1800s gave rise to the first institutional invigilators. Today's proctors, who possess varying standards of professionalism and expertise (Rios & Liu, 2017), may also be employed by specialist agencies. Proctors also support stressed students (Sloboda, 1990) and provide equitable exam environments. They are thus required to meet some of the ethical obligations of educational institutions to provide fair and equitable academic assessment. Although not all instructors use exams for assessment, exams still enjoy wide support (Butler and Roediger, 2007). Because exams can be readily invigilated, instructors can have greater confidence that the work is the student's own.

Today's OP software can be easily integrated into existing university learning management systems. Reports of the exam session generated by the technology can then be uploaded to a dashboard for convenient review. Although different OP platforms perform broadly similar functions, they sometimes differ, such as in their level of monitoring. Platforms can variously allow pure automated proctoring and/or the addition of human invigilators, either from universities or OP companies themselves. Given this variety and flexibility between and within various OP platforms, we shall describe the more important and/or ubiquitous features. Obviously, OP capabilities may increase in time, potentially raising new ethical issues.

2.1 Monitoring and Control of Devices

Typically, students must download OP programs or install a web browser extension (which may be deleted post-exam) and permit the commandeering of their computer's microphone and camera. Different programs allow different degrees of monitoring. They can variously capture screen images, access web page content, block browser tabs, analyze keyboard strokes, and change privacy settings (Norman, 2020).

2.2 Candidate Authentication

OP software can record IP addresses, names, and email addresses and can request a password or ask other questions to verify candidates' identity. Programs typically require candidates to display an officially recognized ID card and photo to be matched against their faces by a live proctor (or, conceivably, an AI algorithm). Some programs can analyze the keystroke cadence of typed names to yield biometric substitutes for handwritten signatures; one program can even request biometrics like fingerprints (Examity, 2020). Programs offering more ID data point checks may improve reliability of authentication, while those offering fewer checks may be championed by purveyors as less privacy intrusive.

2.3 AI-Based and Human Online Proctoring

Online exam invigilation by algorithms with or without supplementation by a person raise some of the strongest concerns. Examinees may be prompted to activate their webcam and turn their device around 360° to "scan" the room for unauthorized materials and family, friends, or housemates (Examity, 2020; Proctorio, 2020). Some programs can detect other devices like mobile phones. The face and body of the candidate can also be monitored, either by means of automated or live human proctoring.

Some AI algorithms can conduct voice and facial recognition but more commonly perform facial detection and analysis. Many automated proctoring systems are trained using machine learning (ML), which is a particular AI technique that uses data to learn a model. ML algorithms can be trained on thousands of video examples to recognize movements of eyes and head that appear to correlate with suspicious behavior, like repeatedly glancing away from the screen. This would include a set of "negative" examples of people purported to not be engaged in misconduct and a set of "positive" examples of people engaged in behaviors that appear to be misconduct, such as talking and unusual eye movements. Of course, these types of behaviors can also be quite normal (talking to oneself, glancing around while thinking), but the OP systems claim only to detect the behaviors, not misconduct itself.

Given these negative and positive examples, an ML algorithm would be trained to automatically identify the behaviors. Typically, the final model is not 100% accurate: There will be false negatives (failing to flag the suspicious behavior) and false positives (flagging non-suspicious behavior), and the job of the data scientist is to improve the training or the data to reach a threshold of accuracy that they are happy with—e.g., 95% correct identification.

Once a trained OP model is included into an OP system, it then raises "red flags" during the exam (allowing the human invigilator to immediately intervene) or afterwards. Given the fact that the technology is not 100% accurate, such human intervention is crucial. How and why the false positives are managed would be up to the particular educator or their institution. Some OP companies claim that the combination of AI and trained human proctors provides greatest accuracy and reliability:

> *The exciting thing about innovating with machine learning technology is that our system is continuously learning, adapting and getting smarter with every exam. ProctorU's goal in introducing AI into proctoring is not to replace humans but, rather, to strengthen the accuracy of proctoring by assisting humans in identifying details such as shadows, whispers or low sound levels, reflections, etc., that may otherwise go unnoticed.* (ProctorU, 2020a)

Companies claim that well-designed AI can also mitigate human bias and error (Proctorio, 2020) and surpass the human ability to accurately detect cheating. Video and audio recordings and analyses are typically stored for a period of weeks or months on company-owned or other servers before being deleted.

Before beginning our ethical examination, it is worth noting the four key stakeholders and their perspectives in relation to OP technologies. The stakeholders are students, educators, institutions, and companies. Some students have embraced or adapted to OP platforms, while others still strenuously object to them. Educators too have various and differing views about OP technology. Typically, students are not given the option to refuse OP platforms if they object on principle, although students with special needs may be offered alternatives. Institutions often grant discretion to educators, but this does not mean that there could not sometimes be pressure on them to use technology that the institution has paid for. OP software provides convenience and flexibility for educational institutions, but they also carry reputational risks, including adverse media attention and public relations headaches from, for example, personal data leakage and hacking. In Australia, it has tended to be the smaller universities that have most adopted them (Selwyn et al., 2021). There is a risk for educators and institutions of prosecuting misconduct against students who are innocent. Clearly, companies have strong profit-driven motivations for promoting OP options. This provides further risks for universities, not least when company personnel assist with the technical setup and the invigilation itself. Nonetheless, companies often argue that their products meet legal and ethical requirements.

3 Philosophical Approach

This essay employs an analytical philosophical approach which includes a range of moral principles and values. The principles are broadly drawn from the burgeoning field of AI ethics (Lin et al., 2017) which often refers to the principles of fairness, non-maleficence, transparency, privacy, accountability, and respect for autonomy.

These principles—plus what we are calling the values of academic integrity, liberty, and trust—also have relevance to educational practice and educational philosophy (Curren et al., 2003). Values such as liberty and trust in technology (Jacovi et al., 2021) and in educational institutions are important for democratic society more generally. These moral notions can arise at the intersection of digital technologies and education. An example is the ethics of using data analytics to measure student performance (Kitto & Knight, 2019).

Compared to the philosophy of education, AI ethics (and more broadly digital ethics) is young and still under development. AI ethics principles have occasionally been criticized for their lack of practical specificity and theoretical rigor and for sidelining wider issues of economic and racial injustice (Kind, 2020). The principles may also be misused in the prosecution of personal or corporate interests (Floridi, 2019). Additionally, principles such as fairness may be used in confusingly different ways (Mulligan et al., 2019). However, these ethical ideas provide a starting point for scrutinizing AI as a socio-technical system. Furthermore, our use of such principles goes some way toward fleshing them out and specifying their application to a novel, concrete socio-technological case. We also indicate links between these moral ideas and wider social issues and trajectories in the context of rapid technological change.

In a global survey of AI guidelines, Jobin et al. identified the ethical principles of transparency, justice and fairness, non-maleficence, responsibility, privacy, beneficence, freedom and autonomy, trust, sustainability, dignity, and solidarity (Jobin et al., 2019). In another recent study, Floridi et al. highlight the ideas of beneficence, non-maleficence, autonomy, justice, explicability, and accountability (Floridi et al., 2018). Ethical notions such as these feature in many other discussions in the AI ethics literature.

Some of the moral notions we are employing, such as autonomy and non-maleficence, feature, albeit with some differences, in the more mature tradition of medical ethics, where they have been developed in detail (Mittelstadt, 2019). Scholars are now exploring how various principles and values apply in medical ethics as AI begins to transform healthcare (Laacke et al., 2021; Quinn et al., 2021). Medical ethics developed significantly in response to abuses of human research subjects, especially disadvantaged and oppressed groups. In those cases, powerful medical, scientific, and educational institutions were often not held accountable for overriding the autonomy and the right to genuine informed consent belonging to less powerful human subjects. As in medical ethics, the ethical principles and values we deploy here may be regarded as non-committal amongst normative theories such as utilitarianism, virtue ethics, and deontology. Such theories may accommodate these so-called mid-level ethical notions (Beauchamp & Childress, 2001).

Furthermore, normative theories may be used to justify and deepen the notions and their operation. Shannon Vallor, for example, has provided an extensive basis for such ethical ideas in the form of a virtue ethics applied to contemporary socio-technical developments. Her "technomoral virtues," which include justice, honesty, humility, care, civility, wisdom, and courage, help to determine what it is to live well amidst the challenging socio-technical trajectories of the twenty-first century (Vallor, 2016).

Here, we briefly introduce the selected principles and values. When we return to them in more detail in the Discussion, it should become clear why they are especially important for understanding the ethics of OP technologies. Fairness is commonly referred to in AI ethics as well as in moral and legal philosophy. Concerns about fairness may encompass an absence of illegitimate bias, equity considerations of accessibility and opportunity, treating people as ends in themselves and not merely as means, and procedural justice. The concept of fairness is sometimes connected to the values of transparency and accountability in AI ethics (Jobin et al., 2019).

Transparency can refer to the degree to which the determinations or predictions of AI systems are revealed to relevant parties in ways that those parties prefer and can understand. Although transparency is not necessarily or always an ethical good, it is associated with more basic ethical ideas such as justice and respect for autonomy sufficiently frequently that it is often treated as a key ethical principle in AI Ethics. Accountability relates to the ethical responsibility of those designing and using technology to implement appropriate responses and mechanisms for ensuring that the other principles and responsibilities are upheld. Respect for autonomy is a widely prized modern notion which unsurprisingly features both in AI ethics and the philosophy of education (Siegel et al., 2018). It represents a broad commitment to allowing each individual to determine their own personal values and make their own choices, within the general framework of acceptable conduct determined by the society in which they live. Non-maleficence cautions against doing harm to others and requires that any harm done must be morally justified. Privacy is highly relevant to AI ethics because new digital technologies often collect, process, retain, and interpret vast amounts of personal and sensitive data, and indeed in many cases are enabled by such data.

The value of academic integrity is widely regarded as critically important in education (Bretag, 2018; Dawson, 2020). Digital technologies, however, pose some threats to academic integrity, while also providing anti-cheating responses (e.g., plagiarism checking software). Academic integrity involves "commitment from students, faculty, and staff to demonstrate honest, moral behavior in their academic lives" (International Center for Academic Integrity, 2021). It requires the nourishment of conditions in which honest and genuine teaching and learning can take place. Although educational institutions are partly motivated by reputational concerns, most nonetheless regard academic integrity as a vital intrinsic value. There can be severe disciplinary sanctions for academically dishonest students and faculty alike. Some scholars call for universities to not just penalize perpetrators, but to actively promote "positive values of honesty, trust, fairness, respect, responsibility, and courage" amongst students, staff, and the institutional culture (Bretag, 2018).

Finally, the values of liberty and trust are increasingly important to public concerns about fast-moving technologies, data gathering, surveillance, and the like (Zuboff, 2015). These and other ethical notions are also clearly relevant to the treatment of students by educational institutions. Furthermore, some of these values and principles are also implicated in the civic responsibilities and cultural roles of universities. As we shall now see, these moral concepts help to illuminate the ethics of OP technologies.

Table 1 Ethical principles and values and their implications for OP exam technology

Ethical principle	Implications for OP exam technology
Fairness	Equitable access to technology and remote exam settingsEqual, not biased nor discriminatory, determination of cheating
Transparency	Transparent use and explanation of the nature of the technology and its selected functionsTransparent use of AI-based "red flags"
Non-maleficence	Effective and safe application of the technology which does not cause harm to the subject
Privacy	Privacy in collection and security of personal data and exposure of body, behavior, and home spaces
Respect for autonomy	Examinee autonomous choice regarding personal data use, use of AI, video recordings, strangers as proctors
Accountability	Governance, auditing, and other mechanisms to ensure that the entity using the technology is vigilant and responsive in respect to the risks of harm or misuse Processes for individuals to appropriately contest outcomes
Ethical value	Implications for OP exam technology
Academic integrity	Ensuring academic honesty, rigor, excellence, and institutional reputation
Liberty and trust	Potential wider effects on freedoms, use of digital technologies, and society's trust in AI, universities, etc

4 Discussion: Applying the Principles and Values

We can now apply in detail the moral principles and values outlined above to OP technologies. Note that the ethical issues discussed below may pertain to some OP functions but not to others. The technology may give institutions some discretion over which capabilities are used. After first discussing academic integrity, we examine fairness, non-maleficence, transparency, privacy, autonomy, liberty, and trust as they apply to OP technology. We touch on accountability in the closing section. Violations of one principle can overlap with violations of others. For example, privacy violations may cause certain harms and so also be violations of non-maleficence. Table 1 summarizes these values and principles and some of their possible implications for our case.

4.1 Academic Integrity

Academic integrity, a vital value in academia, can be threatened by student lack of awareness, dishonesty, and misconduct. Studies show that student cheating is critically affected by institutional culture (Bretag, 2018). Forms of academic dishonesty and misconduct include impersonation and contract cheating, unauthorized use of cheat notes, and the copying of exam answers from fellow students or online sites. OP programs target these illicit activities. There are several ethical reasons why it is vital to prevent academic dishonesty (Kaufman, 2008), which we can briefly enumerate.

First, the value and viability of courses and universities depend on their academic integrity and educational rigor. Second, permitting cheating is unfair on students who are academically honest. Third, knowledge that others are cheating can create for honest students an invidious moral choice between self-interest (e.g., where class rankings matter) and personal integrity, as well as causing a hurtful sense of both being taken advantage of by fellow students and let down by the university. Fourth, universities arguably bind themselves to providing students with (in some sense) a moral education alongside an intellectual education, minimally by nourishing a favorable academic culture in which academic integrity and honesty are salient (Dyer et al., 2020).

Although universities rightly encourage in students an autonomous and sincere commitment to honest behavior—and may institute other mechanisms like ethics policy statements and student honor codes (McCabe & Trevino, 2002)—the failure to invigilate where necessary to prevent cheating above a certain level can, amongst other things, convey the impression that academic honesty is unimportant, thereby negatively affecting the institutional culture (Bretag, 2018). What that precise damaging level is in any particular case requires a difficult judgement. Yet although its effects may be hard to judge, it would be too quick to simply dismiss the idea that student cheating can sometimes have corrosive effects on academic integrity and all that entails (cf. Swauger, 2020b). In fact, there is evidence that, for example, the behavior of peers directly influences other students' academic honesty (Brimble, 2016).

Some studies suggest that students more often cheat in online testing environments than traditional exam rooms (Srikanth & Asmatulu, 2014), although there are conflicting views (Stuber-McEwen et al., 2009). Cheating may help students to achieve higher grades in assessments, but it may also degrade their learning and longer-term interests. For all these reasons, both universities and students have significant interests in the maintenance of academic integrity and its flourishing as an institutional value. Consequently, the need to preserve academic integrity and a corresponding culture of honesty and awareness constitutes a non-trivial reason for considering the use of OP technologies.

4.2 Fairness—Equity and Accessibility in AI Use

Remote invigilation by means of OP technology promises accessibility benefits not only for students who study remotely but also for students who officially study on campuses but who have reasons for sitting exams from home. Institutions may save on costs of hiring exam centers and professional invigilators, which is important in the face of severe budgetary constraints exacerbated by COVID-19 lockdowns.[1] OP can facilitate a greater range of online course offerings and benefit students from distant places or with limited mobility. The technology allows exams to be scheduled day or night, which could particularly benefit, for example, parents. For such

[1] Note however that OP technology can cost many thousands of dollars for institutions (Grajek, 2020) and this may serve as another disincentive to their uptake.

reasons, OP may help promote more equitable outcomes, including for traditionally excluded social groups.

However, some students may be disadvantaged by OP in certain ways. This includes students who lack reliable internet connections—an online exam may be voided if the internet even momentarily disconnects. Some students lack appropriate devices, such as web cameras, and home or other environments in which to sit exams. To be fair, OP providers largely appear eager to ensure that their programs are executable on numerous devices and do not require super-fast internet connections. Also, universities may loan devices and arrange for select students to sit exams on-campus or at other locations. A potential drawback in some such cases is that doing an exam later than the rest of the cohort may delay course progression or graduation. Hence, there are logistical issues with fairness implications for institutions to consider.

4.3 Fairness—Bias, Discrimination, and AI-facilitated Determination of Cheating

OP platforms using AI and/or live proctors may increase fairness for honest students by identifying relatively more cases of cheating. Some proponents claim that digital proctoring does better on this score than traditional invigilation where the ratio of proctors to test-takers is very low (Dimeo, 2017). This claim, of course, would need to be backed by empirical studies. In addition, the potential for OP software to create unfairness also needs to be considered. Unfairness can relate to inequitable outcomes and unjust processes. This may play out in different ways. For example, false-negative identification of cheating may constitute unfairness for non-cheating students, while false positives may result in unfairness for those examinees. Unfairness may flow from the use of AI, from remote human invigilators, or from both together. This needs to be spelt out.

The general fairness problems created by the use of AI and ML are the subject of vigorous contemporary public and academic discussion in AI ethics. Deployment of ML has starkly exposed its potential for inaccuracy and bias. Notorious cases of inaccuracy and bias in ML include automated reviews of curriculum vitae for job applications that favor male candidates, determinations of parole conditions for offenders that apparently discriminate against people of color, and the disproportionate allocation of policing to disadvantaged communities (O’Neil, 2016). Thus, machine bias can create substantial unfairness (Jobin et al., 2019).

Facial recognition technology too has been criticized as inaccurate and has even resulted in legal action, despite the fact that the ML algorithms may have been trained on thousands or millions of images (Peters, 2020). Energetic debate has similarly centered on the so-called biases that can afflict ML. Again, facial recognition software has been associated with bias in the (mis)recognition of certain racial groups and gender (Buolamwini & Gebru, 2018). Recently, reports have emerged that facial recognition functions in OP systems sometimes have more difficulty recognizing darker skin tones (Chin, 2021a). Affected students may therefore have additional trouble undertaking online exams or may be denied access altogether. Misrecognition not only represents an equity of access issue with associated harms such as distress; it is also a matter of racial fairness, since certain groups of people who have darker skin have often suffered historical discrimination and injustice, and this includes discrimination and injustice in education. Thus, overcoming the technical problems, finding acceptable workaround solutions, or simply not using facial recognition functions are ethical imperatives.

Another form of bias or discrimination may arise through the model used by the OP for "normal" or "acceptable" exam behavior. OP companies refer to flagging suspicious gestures and even tracking eye movements. Yet people with disabilities or who are neuro-diverse may not always behave in a way that is recognized by these processes (Swauger, 2020a), and this may lead to false positives as such people are red flagged for cheating through the manifestation of their behaviors.

Bias can creep into ML through input of skewed and poorly representative training data or through the mechanisms of pattern-searching (Mehrabi et al., 2019). Presumably, this could occur in the training and operation of ML in online cheating analysis. As OP platforms accumulate increasingly larger data sets on which to train, their reliability should increase. But bias and inaccuracy may never be fully eliminated, and some forms of unfairness may not be solvable by purely technical means (Selbst et al., 2019), leaving the potential for students to be unfairly charged with cheating. Nevertheless, unfairness in socio-technical systems need not always be the outcome of ML bias. OP companies stress that it is, after all, not the AI algorithm that ultimately makes a judgment about academic dishonesty, but a knowledgeable human being, such as the course instructor. Furthermore, instructors may choose which settings they will and won't use—for instance, they might choose to disable or ignore AI algorithms that track eye movements.

But this flexibility does not totally eliminate ethical concerns. For example, instructors may have unwarranted faith in the red flags, such as the automated flagging of "suspicious" head movements. The problem is magnified when we consider the conscious and unconscious inclination for some people to over-trust AI (Dreyfus et al., 2000). Even where psychological bias is absent, instructors may be unsure how to interpret some red flags and may draw incorrect inferences from them. Certain flagged events, such as when the test-taker is plainly replaced with another person, are relatively easy to assess. But more subtle flags may be much harder to appraise. These could include flags for "low audible voices, slight lighting variations, and other behavioral cues" (ProctorU, 2020b). Further, if the ML element is intended to enhance detection of cheating over and above a human observer carefully attending to the same images (etc.), then it follows that an independent level of trust is intended to be invested in the AI assessment. As ML technology advances, greater epistemic weight will likely be placed on its judgments, thereby potentially elevating the risks of unfairness.

4.3.1 Non-maleficence

Reliance on OP technology raises risks of harm for both students and universities. As mentioned, there is a risk to students of false claims of cheating that did not occur. Wrongful allegations of academic misconduct, especially where there is no process for contestability, may affect job prospects, self-confidence, and personal trust in the university. For universities, false negatives and positives could more broadly undermine social trust in the integrity of the institution.[2] Therefore, it is important that such systems are effective, that they work with a sufficient degree of accuracy, and that there is clarity about their reliability. But, as we have noted, the operation of such systems is often opaque, and although claims are made about accuracy, the OP company websites rarely if ever cite rigorous studies to justify their claims and to eliminate concerns about false positives (e.g. Examity, 2020; Proctorio, 2020).

One could imagine hypothetical situations that clearly involve unfairness and harm related to assumed belief, group membership, or behavior. Suppose, for example, that some future AI proctoring system red flags the presence in the examinee's room of white supremacist propaganda or pornographic material. Or suppose that the AI system is biased towards red-flagging suspicious eye movements in people with disabilities or assumes that black students need closer monitoring than white students. Again, imagine the program casts doubt on students' honesty purely from a brief unintelligible exchange of words with someone who happens to enter the room. The emergence of such systems would obviously be cause for alarm.

AI-led or human-led, post-exam determinations of cheating differ from in-person or live remote invigilation, where the primary anti-cheating mechanism is typically to warn students at the precise time of the potential infraction (e.g., when students are seen conversing). Unlike subsequent review of captured OP data, that latter mechanism does not depend on an official charge of academic dishonesty, but on its immediate prevention. Some test-takers may simply have idiosyncratic exam-taking styles, or disabilities and impairments, that trigger specious AI red flags. Even falsely suggesting that these individuals are academically dishonest, let alone accusing and penalizing them, would potentially be unfair and harmful. Even though such a false suggestion or imputation is less morally serious than a false official condemnation, they are still morally serious. Further, recipients of spurious insinuations are likely to receive them as an injustice and to feel corresponding hurt. In addition, such individuals, in an effort to avoid this potentially wrongful treatment, may be forced to disclose personal idiosyncrasies or impairments, compromising their privacy and potentially doing them harm in the process.

4.3.2 Transparency

Uncertainty may persist about how precisely the AI identifies "cheating behavior." Some OP company websites are more transparent than others about how their AI systems work. But even with some explanation, it can be confusing and difficult to gain an adequate understanding of how they compute red flags and how reliable those determinations are.[3] One representative explains that their company uses an:

[2] We discuss trust further below.

[3] For one example of an attempt to explain how AI-based judgments are made using a "credibility index," see Mettl (2018). This program allows users to select which indicators or patterns (e.g., eye movements) to incorporate and which to exclude from the AI analysis, as well as the weight they carry.

> *Incredibly futuristic AI Algorithm that auto-flags a variety of suspicious cases with 95%+ accuracy…With AI-driven proctoring, the algorithms will soon become trained enough to prevent cheating 100%, a guarantee that a physical invigilator cannot always promise.* (Kanchan, 2019)

Compared to, perhaps, the plagiarism detection tool Turnitin, proctoring AI may strike users as highly opaque (Castelvecchi, 2016). The problem of AI "black boxes" is one reason why ethicists stress the moral need for transparency in AI (Reddy et al., 2020).Transparency in this context may work at different levels and different times. At the outset, students need to understand enough about the OP process to know what is expected of them in exams so as not to trigger a red flag. Students will also need information on how to contest any adverse finding and their rights of appeal, and those who are accused of cheating will need to know the basis on which that allegation is made.

Admittedly, not all of this information need be presented to students upfront, especially given concerns about information overload and about students gaming the system. Nonetheless, academic fairness requires that the evidence and procedures on which accusations of cheating are made are generally defensible and transparent. To reduce the risks of unfairness and emotional harm, OP companies and universities should be transparent about how the technology works, how it will be used in particular circumstances, and how it will impact on students, including those with disabilities.

4.4 Privacy

OP technologies raise moral concerns about privacy which privacy laws, and university policies and governance, may not adequately address (Pardo & Siemens, 2014)—especially given that many jurisdictions have privacy laws that have not been amended to adjust to the data collecting capacities of new digital technologies (Australian Government, 2020). OP technologies collect several kinds of data, including the capturing and storage of information from devices, the gathering of biometric and other ID details, and the audio and video recording of students and their environment. It should not be surprising that some students have a sense of "Big Brother invading their computers" (Dimeo, 2017).

Privacy is a large philosophical topic. A rough distinction can be made between private and public domains (Warren & Brandeis, 1890). Privacy can relate to the (non)exposure to other individuals of one's personal information and one's body, activities, belongings, and conversation (Gavison, 1980; Moore, 2003). However, what is private for one person may, in a recognizable sense, not be private for another (Moore, 2015). For example, I may strongly prefer that no strangers gaze inside my bedroom and watch me studying, whereas you, who do not draw your curtains, may not care. Exposure of my bedroom and activities to passers-by represents in my case a loss of privacy; in your case, it does not. So, there is an intelligible sense in which the determination of privacy and its breaching can turn partly on individual perspectives about the personal. At the very least, we can say that the moral seriousness of exposure is plausibly related, to some extent, to these individual preferences.

While the moral "right" to privacy may sometimes justifiably be infringed (e.g., in law enforcement), it is still a vital right or interest. For some philosophers, privacy's value essentially reduces to the value of liberty and autonomy (Thomson, 1975), i.e., to a person's ability to act and make choices about their own lives. The right to privacy is also sometimes seen as the enabler of other important rights, such as, for example, freedom of expression and association. For other thinkers, privacy's importance relates to possible harms resulting from public exposure of the personal (Rachels, 1975), such as social embarrassment and deleterious financial or employment repercussions. We might regard privacy's importance not as confined to a single philosophical conception, but to a range of conceptions that cover respect for autonomy, the causation of various kinds of harm, and so on. Information or data privacy raises unique challenges for all these perspectives because of the scope and the longevity of the inferences that may be drawn from data about an individual.

OP technologies may threaten personal privacy in several ways. Reports exist of inadvertent capture of personal or sensitive information, such as in one case a student's credit card details that were accidentally displayed on their computer screen (Chin, 2020). While technology designers might address some such risks, there are additional risks concerning data security. Captured information can be stored in encrypted form on host servers such as Microsoft and Amazon servers. For their part, many OP companies claim that they have no access to this encrypted information and therefore cannot view video recordings or obtain sensitive personal data. Furthermore, purveyors claim to be compliant with legal protections like the EU's General Data Protection Regulation (GDPR), which carries heavy penalties for breach.

Companies may also have internal rules against sharing data with third parties and for commercial gain (Dennien, 2020). Universities too are required to have stringent cybersecurity and privacy policies. However, there can be no absolute guarantee against leakage of data or successful cyberattacks on servers used by companies or universities. The maintenance of such privacy is never completely certain: These kinds of cyber risks are always present with any data collected by any institution (ANU, 2019). It is nonetheless possible that students may feel particularly anxious about the possible loss of the kinds of sensitive personal information (e.g., video recordings, certain data from personal computers) collected by OP technologies.

As we saw, some people worry that OP platforms are especially intrusive because they readily facilitate video (and audio) capture of examinees and its live or subsequent review by a person. This concern may be countered by proponents of OP technologies as follows: Students necessarily relinquish aspects of their privacy in education. In-person invigilation, which is morally uncontroversial, is already privacy-invasive: Strangers or instructors watch students like hawks, scrutinizing their activities and personal belongings. On this moral view, online proctoring is essentially the same in an ethical sense as in-person invigilation.

However, this argument is highly contestable. We may start by noting that

> *Students who have used Examity say it feels much weirder than proctoring with a professor…They're being watched closer up, by a stranger, and in a*

> *place more private than a classroom…students described their experiences as everything from "uncomfortable" to "intrusive" to "sketchy."* (Chin, 2020)

One element of the privacy intrusion relates to human invigilators seeing into the home environment of the student, such as bedrooms or loungerooms. Another element is that of other human beings watching video of the faces and upper bodies of students themselves. To make the analogy with traditional invigilation more truly comparable, then, we must imagine an in-person supervisor sitting near the examinee and staring at them throughout the exam. Such observation would include scrutinizing the student's expressions or micro-expressions, perhaps with the help of an AI facial detection/recognition device.

Furthermore, OP may allow the human invigilator, who may reside locally or on the other side of the world, to re-watch the video and to use its pause function, potentially in private. In traditional exam rooms, the presence of other students, instructors, and invigilators provides a degree of security and protection. In contrast, the student who is invigilated by OP technology cannot know, even when they are given assurances by universities and OP companies, how the online human proctor uses the video. For example, students cannot be sure that they are not being leered at or that online proctors have not shared their images shared with third parties.

This online proctoring scenario should strike us as potentially more invasive of privacy than in-person invigilation, irrespective of whether students—including students whose histories and psychologies render them particularly averse to being closely watched by strangers—have the additional concern that viewers may take a prurient interest in them. Besides, some students evidently have that view. Furthermore, because (as we suggested) what constitutes a loss of privacy turns in one sense partly on the individual's own perspective, OP in those cases is more intrusive of privacy for the students who feel invaded. It follows that the hurdle for justifying its use is that much higher.

4.5 Respect for Autonomy

Autonomy might be restricted by online proctoring in several ways. For example, it may require students to avoid doing things they can often do in traditional exams, such as muttering to themselves, looking to the side, and going to the bathroom—lest they raise automated red flags about suspicious behavior. Some students may simply prefer not to be invigilated by AI or by online human proctors or to have their images and personal data collected and viewed.

Philosophers often regard respect for autonomy as a fundamental ethical value (Christman, 2018). Autonomy in this sense implies self-governance, or the ability of a rational and mature agent to form and act on decisions and personal values free of compulsion and manipulation. As we saw, some philosophers ground the value of privacy in the value of autonomy. We should be clear, however, that respect for autonomy is not reducible to respecting privacy: Respect for autonomy can apply to the use of personal information even where loss of privacy is not at stake (e.g., because data is anonymized). Further, respecting autonomy may require providing agents with genuine opportunities to decide whether to consent to the action in question.

Of course, consent is sometimes purely formal, such as in ticking a box or clicking a button (Paterson & Bant, 2020). For a more robust or truer kind of consent, the choice must be made voluntarily and exercised with liberty and without coercion, and the chooser must have adequate knowledge of the nature, risks, and benefits of committing to or refraining from the relevant action (cf. Clifford & Paterson, 2020). This requires transparency about the nature and potential effects of OP programs. A robust standard of genuine consent would also allow students to be able not to consent to OP without penalty and to freely choose instead a human invigilator. If this option is unavailable, then the consent cannot be considered genuine consent. From the perspective of a university, such discretion granted to students to decide whether to participate in an examination-related process may be unmanageable. However, if the institutional decision is to compel participation, then this demonstrates the need for other processes to be put in place to protect students’ interests.

Responding to legal argument opposing compulsory Proctorio invigilation at the University of Amsterdam, the Amsterdam District Court found that students have no right to choose not to use the university processes (Persbureau, 2020). An ethical case might also be made that autonomy and the prima facie requirement for informed consent are already justifiably restricted in education. Educational limitations on liberty extend, quite obviously, to the prevention of cheating and more (as when personal student information is collected for enrolment). One early student criticism of Turnitin likened its use to the coercive drug testing of students (Glod, 2006). Such moral objections are now often (though not universally) considered exaggerated. Indeed, our attitudes towards novel technologies can change with familiarity and understanding. However, deciding when it is justified to limit autonomy for the sake of academic integrity requires moral (and not just legal) judgment.

Most ethicists acknowledge that coercion and compulsion are sometimes justified, most obviously when the freedom is likely to result in significant harm to others (Gaus et al., 2020; Mill, 1966). But even then, respect for autonomy may imply that limitations upon autonomy be minimized wherever possible and that relevant information be provided transparently to students who are under compulsion from their universities. This would include information related to the above concerns, along with others. For example, OP companies may use data derived from student exams to train ML algorithms (ProctorU, 2020b) without the students being (adequately) informed that their data will be so used. Such use may arguably produce good outcomes (e.g., improving accuracy and reducing AI bias), but institutions that fail to investigate data use arrangements and/or inform students accordingly disrespect their autonomy.

4.6 Wider Social Implications: Liberty and Trust

To conclude our investigation of the ethics of OP technologies, we shall briefly discuss some of its potential wider implications. The current context is one of rapid technological change. Shannon Vallor claims that “the unpredictable, complex, and destabilizing effects of emerging technologies on a global scale make the shape of the human future increasingly opaque" (Vallor, 2016, p. 1). Whatever one makes of that, it is undeniable that the present rate of technological change presents both significant potential benefits and risks for society. We have explained some potential benefits and harms of OP technologies; here, we indicate some of their possible broader socio-ethical dangers (Selwyn et al., 2021). Although the risks we mention here are admittedly much less certain, they are real enough to consider when forming a comprehensive ethical judgment about this emerging socio-technical example.

The concerns raised earlier about the intrusive and invasive nature of OP technologies have a possible connection to broader technological and social trajectories. To give some stark examples, these socio-technical trajectories may include increased biometric monitoring and surveillance in education (Andrejevic & Selwyn, 2020) and their normalization in society (Adams & Purtova, 2017); the step-by-step evolution of a security state (Reiman, 1995); commercialization of technology in education; personal data being used in unexpected ways or "anonymized" data being publicly reidentified (Culnane et al., 2017); constrictions of the private domain (Nissenbaum, 2009); "nudging" that prompts us to think and behave in certain desired ways; and the spread of "nonhuman" judgments that threaten human rights. Consider, for example, contemporary public concerns arising from facial recognition, dubious employment of personal information scraped from social media, heightened tracking and tracing during the COVID-19 pandemic, AI decision-making in jurisprudence, and more (Feldstein, 2019).[4]

In the field of education, digital technologies, not least the internet, have enabled both exciting new options for teaching and learning and concerning new modes of academic dishonesty and fraud. The evolution of so-called e-cheating, discussed in detail by education scholars such as Phillip Dawson, requires novel and creative responses (Dawson, 2020). But although there can be beneficial technological responses to technologically mediated problems, it is at least arguable that OP technology could also (if only modestly) contribute to the above worrying social trajectories. Such risks are, as we stress, very difficult to assess; but that does not mean they may be ignored. As we noted, some universities have chosen and will choose not to use OP technologies. We can also perhaps expect that many universities will make diligent efforts to protect and foster respect for privacy, liberty, autonomy, and fairness (Kristjánsson, 2017) when they do use them.

Such efforts are entirely proper. Indeed, it may be suggested (though we cannot argue it here) that universities should recognize and reaffirm their standing as bulwarks against the natural proclivity of governments and powerful corporations to intrude into people's private lives and to chip away, deliberately or unthinkingly, at their freedoms. Many universities explicitly claim to be major contributors to the democratic social good in part through the education of ethically responsible leaders and global citizens. Yet some students and university staff evidently feel that OP

[4] In an example that comes from OP technologies, a website for "Proctortrack" said that its ""RemoteDesk" solution goes beyond exam proctoring to provide automated monitoring of people who work from home" (Kanchan, 2019).

platforms could damage a university's visible commitment to liberty and to earning trust. That such an effect could have wider reverberations is a reason for taking the ethical aspects of OP technologies seriously.

The weight of the above concerns will depend not only on cultural factors and differences but also partly on factors such as the extent of opposition to OP technology amongst students and staff, and the relative intrusiveness of various proctoring functions that have in the eyes of many vaguely Big Brother overtones. Disquiet would mount if OP platforms allowed, say, the undisclosed on-selling of test-taker data and use of AI to generate Uber-style ratings to indicate an examinee's honesty while closing off avenues for appeal and contestation.

In today's digital and cultural climate, none of these further possibilities may be blithely dismissed. At some point, universities may want to take a stand against not only the ethical risks to students and to their own reputations, but against the societal risks of endorsing particularly invasive and intrusive technologies. Whatever institutions decide, our objective in this section is to underline the point that OP technologies need to be considered not just from the perspective of their potential immediate and local effects, but also from the perspective of their more distant, wider, and longer-term potential effects on culture and society—even if those effects are much harder to measure and predict with any certainty.

5 Conclusion: Justification, Accountability, and Future Research

Debate and disagreement about the appropriateness of remote OP technologies in general, in remote education, and in difficult circumstances like pandemics that require more off-campus assessment are bound to continue. As we saw, there are considerations that speak in favor of OP technologies despite their drawbacks. Indeed, it is fair to acknowledge that in-person proctoring is not ethically perfect either: It can both miss cheating and similarly result in unfair accusations of academic dishonesty. Furthermore, we have accepted that it is vital to maintain academic integrity to protect both students and institutions. It is true that the pedagogical value of high-stake examinations is sometimes questioned[5]; and faced with the ethical problems of OP technology, some academics will adopt alternative assessments. But, on the assumption that high-stake exams have value and will persist in education, there are reasons for regarding at least some OP technologies and capabilities as representing acceptable "proctors" or proctors' assistants.

Nonetheless, the above analysis reveals that OP platforms raise ethical concerns over-and-above those affecting live and in-person exam invigilation. These concerns include an uncertain risk of academic unfairness associated with AI-informed judgment, further diminution of student privacy and autonomy, and

[5] Grajeck reports: "Three in ten institutions are considering broad changes to assessment. Exams are common, but they are only one way to assess learning. The [COVID-19] pandemic is providing 31% of institutions the opportunity to consider more authentic demonstrations of knowledge and skills" (Grajek, 2020).

(perhaps) increased distrust towards institutions that are bastions of social values. Another fear, partially dependent on these former fears, is that OP platforms could contribute to the social trajectories of growing surveillance, liberty and privacy loss, mining of massed personal data, and dubious instances of AI decision-making.

It is difficult to form a general recommendation about whether the benefits of OP technologies outweigh their risks. Educational institutions have several options that include (1) adopting a permissive approach; (2) rejecting OP technologies altogether; or (3) using them in some situations only. Our argument suggests that (1) is unjustified. As we noted, some institutions have opted for (2), others for (3). The choice of conditional and restricted adoption could also involve being selective about which OP companies and which OP functions (e.g., more or less privacy invasive ones) are engaged with.

At the very least, institutions choosing OP technologies should accommodate the ethical considerations we have presented in their assessment policies and governance plans. Drawing on our ethical analysis, we would also like to offer the following set of questions to help guide educational institutions and educators in their decision-making:

1. Are there alternative assessment types that are acceptable, or are closed-book exams essential (e.g., due to regulatory or professional requirements of certain courses)?
2. Can other arrangements be made for all or some students to have in-person invigilation (e.g., by delaying exams or offering alternative sitting options)?
3. Would academic integrity really be degraded to an unacceptable level if OP technologies were avoided?
4. Are the relevant OP technologies and their functions likely to be acceptable to informed students?
5. Is their use consistent with the institutions' (e.g., a university's) social role?
6. Could those technologies be phased out as soon as conditions permit (e.g., social distancing during pandemics)?

In addition to carefully weighing the risks and benefits of OP technologies in deciding whether or not to adopt them, educational institutions also need to have the right systems in place to remain accountable for such choices. Accountability operates at several levels and points in time. Ex ante, or before the technology is utilized, staff and students should be consulted and adequately informed about the impacts and the capabilities of selected OP technologies. For example, relevant information might include how cheating is determined and privacy affected. Information addressing key questions about OP technologies should be readily available.

In operationalizing the use of the technology in exam contexts, accountability may require suitable alternative options for students who cannot reasonably access, or for serious reasons object to, this form of monitoring. Ex post, or after the use of OP to particular exams, there should be clear and accessible avenues for students and staff to contest adverse outcomes. There should also be forums allowing feedback and concerns from students and staff to be aired and transparency provided in institutional responses. Importantly, there should be clear systems for oversight at a systemic level of the operation of any OP being used by the institution—which might include audits and other methods of scrutinizing results, as well as reviewing data provided by the OP provider.

Our analysis also has implications for OP designers and purveyors. Although OP companies do not have the same ethical responsibilities as educators and educational institutions, they do have some. Insofar, as those involved in designing, producing, and selling OP technologies aim to not only make money but respect the interests of educational institutions and their students, they should reflect on the potential harms as well as benefits of their products. Some functions and activities, such as using facial recognition and data without the fully informed consent of examinees, are clearly unethical. Furthermore, these companies should aim to make products that minimize as far as possible any negative impacts and risks, such as impacts on privacy, autonomy, and fairness. Simply offering a suite of functions for institutions to consider, and then claiming that it is the responsibility of consumers to choose if and how they will use them, is unjustifiable.

In conclusion, we hope that future investigations, both theoretical and empirical, take up the question of the implications of OP technologies. At the level of theory, this could mean further studying their various consequences for students, educators, and institutions, including their implications for academic integrity, fairness, privacy, harm, transparency, autonomy, and accountability. It might also encompass investigations into the social, cultural, and ethical implications of these and similar products, such as their possible wider effects on liberty and community trust in emerging technologies and in those who design and deploy them.

Studies that evaluate the empirical benefits and harms of these technologies are also needed. Do automated OP technologies discourage misconduct, or do they just change the way people engage in misconduct? Do they truly detect academic dishonesty, or do they just detect unusual behaviors that resemble it? Are educators successfully able to argue misconduct cases using evidence from these technologies? These questions, and more, are of critical importance for understanding and weighing any practical benefits that may or may not result from these new forms of online proctoring.

Acknowledgements We would like to thank the reviewers for their very helpful comments.

Data availability Not applicable.

Code Availability Not applicable.

Declarations

Conflict of Interest The authors declare no competing interests.

References

Adams, S., & Purtova, N. (2017). Introducing the special issue "rethinking surveillance: Theories, discourses, structures, and practices." *Philosophy & Technology, 30*(1), 5–7. https://doi.org/10.1007/s13347-016-0237-z

Andrejevic, M., & Selwyn, N. (2020). Facial recognition technology in schools: Critical questions and concerns. *Learning, Media and Technology, 45*(2), 115–128. https://doi.org/10.1080/17439884.2020.1686014

ANU. (2019, October 1). *ANU releases detailed account of data breach*. Australian National University News. https://www.anu.edu.au/news/all-news/anu-releases-detailed-account-of-data-breach. Accessed 24 Jul 2021

Asep, H. S., & Bandung, Y. (2019). A design of continuous user verification for online exam proctoring on M-learning. *International Conference on Electrical Engineering and Informatics (ICEEI), 2019*, 284–289.

Australian Government. (2020). *OAIC welcomes privacy act review*. Office of the Australian Information Commissioner. https://www.oaic.gov.au/updates/news-and-media/oaic-welcomes-privacy-act-review/. Accessed 24 Jul 2021

Bartneck, C., Lütge, C., Wagner, A., & Welsh, S. (2021). *An introduction to ethics in robotics and AI*. Springer Nature. https://library.oapen.org/handle/20.500.12657/41303. Accessed 24 Jul 2021

Beauchamp, T. L., & Childress, J. F. (2001). Principles of biomedical ethics. *Oxford University Press, USA*. https://doi.org/10.1136/jme.28.5.332-a

Binstein, J. (2015). How to cheat with proctortrack, examity, and the rest. *Jake Binstein*. https://jakebinstein.com/blog/on-knuckle-scanners-and-cheating-how-to-bypass-proctortrack/. Accessed 24 Jul 2021

Bretag, T. (2018). Academic integrity. In M. A. Hitt (Ed.), *Oxford Research Encyclopedia of Business and Management*. Oxford University Press. https://doi.org/10.1093/acrefore/9780190224851.013.147. Accessed 24 Jul 2021

Brimble, M. (2016). Why students cheat: An exploration of the motivators of student academic dishonesty in higher education. In T. Bretag (Ed.), *Handbook of Academic Integrity* (pp. 365–382). Springer Singapore. https://doi.org/10.1007/978-981-287-098-8

Buolamwini, J., & Gebru, T. (2018). Gender shades: Intersectional accuracy disparities in commercial gender classification. *Proceedings of Machine Learning Research, 81*, 15.

Butler, A. C., & Roediger, H. L., III. (2007). Testing improves long-term retention in a simulated classroom setting. *European Journal of Cognitive Psychology, 19*(4–5), 514–527. https://doi.org/10.1080/09541440701326097

Castelvecchi, D. (2016). Can we open the black box of AI? *Nature News, 538*(7623), 20.

Chin, M. (2020, April 29). *Exam anxiety: How remote test-proctoring is creeping students out*. The Verge. https://www.theverge.com/2020/4/29/21232777/examity-remote-test-proctoring-online-class-education. Accessed 24 Jul 2021

Chin, M. (2021a, January 5). *ExamSoft's proctoring software has a face-detection problem*. The Verge. https://www.theverge.com/2021/1/5/22215727/examsoft-online-exams-testing-facial-recognition-report. Accessed 24 Jul 2021

Chin, M. (2021b, January 28). *University will stop using controversial remote-testing software following student outcry*. The Verge. https://www.theverge.com/2021/1/28/22254631/university-of-illinois-urbana-champaign-proctorio-online-test-proctoring-privacy. Accessed 24 Jul 2021

Christman, J. (2018). Autonomy in moral and political philosophy. In E. N. Zalta (Ed.), *The Stanford Encyclopedia of Philosophy* (Spring 2018). Metaphysics Research Lab, Stanford University. https://plato.stanford.edu/archives/spr2018/entries/autonomy-moral/. Accessed 24 Jul 2021

Clausen, M. (2020, July 29). *Student proctoring software gets first test under EU privacy law*. https://news.bloomberglaw.com/tech-and-telecom-law/student-proctoring-software-gets-first-test-under-eu-privacy-law. Accessed 24 Jul 2021

Clifford, D., & Paterson, J. (2020). Consumer privacy and consent: Reform in the light of contract and consumer protection law. *Australian Law Journal*, *94*(10). https://www-westlaw-com-au.eu1.proxy.openathens.net/maf/wlau/app/document?&src=search&docguid=I5872e29203ff11eb99dafba9e329f6c2&epos=1&snippets=true&fcwh=true&startChunk=1&endChunk=1&nstid=std-anz-highlight&nsds=AUNZ_JOURNALS&isTocNav=true&tocDs=AUNZ_AU_JOURNALS_TOC&context=38&extLink=false&searchFromLinkHome=true. Accessed 24 Jul 2021

Cramp, J., Medlin, J. F., Lake, P., & Sharp, C. (2019). Lessons learned from implementing remotely invigilated online exams. *Journal of University Teaching & Learning Practice, 16*(1), 10.

Culnane, C., Rubinstein, B. I., & Teague, V. (2017). Health data in an open world. *ArXiv Preprint.* https://arxiv.org/pdf/1712.05627. *Accessed 24 Jul 2021*

Curren, R., Markie, P., & Mathews, G. (2003). *A companion to the philosophy of education*. John Wiley & Sons, Incorporated. http://ebookcentral.proquest.com/lib/unimelb/detail.action?docID=214150. Accessed 24 Jul 2021

Dawson, P. (2020). *Defending assessment security in a digital world: Preventing e-cheating and supporting academic integrity in higher education*. Routledge.

Dennien, M. (2020). UQ students raise privacy concerns over third-party exam platform. *Brisbane Times*. https://www.brisbanetimes.com.au/national/queensland/uq-students-raise-privacy-concerns-over-third-party-exam-platform-20200419-p54l77.html. Accessed 24 Jul 2021

Dimeo, J. (2017). *Online exam proctoring catches cheaters, raises concerns*. Inside Higher Ed. https://www.insidehighered.com/digital-learning/article/2017/05/10/online-exam-proctoring-catches-cheaters-raises-concerns

Dreyfus, H., Dreyfus, S. E., & Athanasiou, T. (2000). *Mind over machine*. Simon and Schuster.

Dyer, J. M., Pettyjohn, H., & Saladin, S. (2020). Academic dishonesty and testing: How student beliefs and test settings impact decisions to cheat. *Journal of the National College Testing Association, 4*(1), 30.

Examity. (2020). Auto proctoring. *Examity*. https://examity.com/auto-proctoring/. Accessed 24 Jul 2021

Feldstein, S. (2019). *The global expansion of AI surveillance* (Vol. 17). Carnegie Endowment for International Peace.

Flaherty, C. (2020). *Online proctoring is surging during COVID-19*. https://www.insidehighered.com/news/2020/05/11/online-proctoring-surging-during-covid-19. Accessed 24 Jul 2021

Floridi, L. (2019). Translating principles into practices of digital ethics: Five risks of being unethical. *Philosophy & Technology, 32*(2), 185–193. https://doi.org/10.1007/s13347-019-00354-x

Floridi, L., Cowls, J., Beltrametti, M., Chatila, R., Chazerand, P., Dignum, V., Luetge, C., Madelin, R., Pagallo, U., Rossi, F., Schafer, B., Valcke, P., & Vayena, E. (2018). AI4People—An ethical framework for a good AI society: Opportunities, risks, principles, and recommendations. *Minds and Machines, 28*(4), 689–707. https://doi.org/10.1007/s11023-018-9482-5

Gaus, G., Courtland, S. D., & Schmidtz, D. (2020). Liberalism. In E. N. Zalta (Ed.), *The Stanford Encyclopedia of Philosophy* (Fall 2020). Metaphysics Research Lab, Stanford University. https://plato.stanford.edu/archives/fall2020/entries/liberalism/. Accessed 24 Jul 2021

Gavison, R. (1980). Privacy and the limits of law. *The Yale Law Journal, 89*(3), 421–471.

Ginder, S. A., Kelly-Reid, J. E., & Mann, F. B. (2019). Enrollment and employees in postsecondary institutions, Fall 2017; and Financial Statistics and Academic Libraries, Fiscal Year 2017: First Look (Provisional Data). NCES 2019–021Rev. *National Center for Education Statistics*.

Glod, M. (2006, September 22). Students rebel against database designed to thwart plagiarists. *Washington Post*. http://www.washingtonpost.com/wp-dyn/content/article/2006/09/21/AR2006092101800.html. Accessed 24 Jul 2021

Goedl, P. A., & Malla, G. B. (2020). A study of grade equivalency between proctored and unproctored exams in distance education. *American Journal of Distance Education, 34*, 1–10. https://doi.org/10.1080/08923647.2020.1796376

González-González, C. S., Infante-Moro, A., & Infante-Moro, J. C. (2020). Implementation of E-proctoring in online teaching: A study about motivational factors. *Sustainability, 12*(8), 3488. https://doi.org/10.3390/su12083488

Grajek, S. (2020). *EDUCAUSE COVID-19 QuickPoll results: Grading and proctoring*. https://er.educause.edu/blogs/2020/4/educause-covid-19-quickpoll-results-grading-and-proctoring. Accessed 24 Jul 2021

Hagendorff, T. (2020). The ethics of AI ethics: An evaluation of guidelines. *Minds and Machines, 30*(1), 99–120. https://doi.org/10.1007/s11023-020-09517-8

Hern, A. (2020, August 21). Ofqual's A-level algorithm: Why did it fail to make the grade? *The Guardian*. https://www.theguardian.com/education/2020/aug/21/ofqual-exams-algorithm-why-did-it-fail-make-grade-a-levels. Accessed 24 Jul 2021

Hubler, S. (2020, May 10). Keeping online testing honest? Or an Orwellian overreach? *The New York Times*. https://www.nytimes.com/2020/05/10/us/online-testing-cheating-universities-coronavirus.html. Accessed 24 Jul 2021

International Center for Academic Integrity. (2021). *Academic integrity*. https://academicintegrity.org/. Accessed 24 Jul 2021

Jacovi, A., Marasović, A., Miller, T., & Goldberg, Y. (2021). Formalizing trust in artificial intelligence: Prerequisites, causes and goals of human trust in AI. *Proceedings of the 2021 ACM Conference on Fairness, Accountability, and Transparency*, 624–635. https://doi.org/10.1145/3442188.3445923

Jobin, A., Ienca, M., & Vayena, E. (2019). The global landscape of AI ethics guidelines. *Nature Machine Intelligence, 1*(9), 389–399. https://doi.org/10.1038/s42256-019-0088-2

Kanchan, R. (2019, February 20). Top 5 proctoring solution and service providers. *Mettl*. https://blog.mettl.com/top-5-proctoring-solution-providers/. Accessed 24 Jul 2021

Kaufman, H. E. (2008). Moral and ethical issues related to academic dishonesty on college campuses. *Journal of College and Character*, *9*(5).

Kellaghan, T., & Greaney, V. (2019). A brief history of written examinations. *In Public Examinations Examined, 1–0*, 43–74. https://doi.org/10.1596/978-1-4648-1418-1_ch3 The world bank.

Kind, C. (2020, August 23). The term 'ethical AI' is finally starting to mean something. *VentureBeat*. https://venturebeat.com/2020/08/23/the-term-ethical-ai-is-finally-starting-to-mean-something/. Accessed 24 Jul 2021

Kitto, K., & Knight, S. (2019). Practical ethics for building learning analytics. *British Journal of Educational Technology, 50*(6), 2855–2870.

Kristjánsson, K. (2017). Emotions targeting moral exemplarity: Making sense of the logical geography of admiration, emulation and elevation. *Theory and Research in Education, 15*(1), 20–37.

Laacke, S., Mueller, R., Schomerus, G., & Salloch, S. (2021). Artificial intelligence, social media and depression. A new concept of health-related digital autonomy. *The American Journal of Bioethics, 0*(0), 1–33. https://doi.org/10.1080/15265161.2020.1863515

Lin, P., Abney, K., & Jenkins, R. (2017). *Robot ethics 2.0: From autonomous cars to artificial intelligence*. Oxford University Press.

McCabe, D., & Trevino, L. K. (2002). Honesty and honor codes. *Academe, 88*(1), 37.

Mehrabi, N., Morstatter, F., Saxena, N., Lerman, K., & Galstyan, A. (2019). A survey on bias and fairness in machine learning. ArXiv:1908.09635 *[Cs]*. http://arxiv.org/abs/1908.09635. Accessed 24 Jul 2021

Mettl. (2018). *Mettl ProctorPlus—Experience the Real Power of AI*. https://www.youtube.com/watch?v=_YPu9X3TCXY&feature=youtu.be. Accessed 24 Jul 2021

Mill, J. S. (1966). On Liberty. In J. S. Mill & J. M. Robson (Eds.), *A selection of his works* (pp. 1–147). Macmillan Education UK. https://doi.org/10.1007/978-1-349-81780-1_1

Mittelstadt, B. (2019). Principles alone cannot guarantee ethical AI. *Nature Machine Intelligence, 1*(11), 501–507. https://doi.org/10.1038/s42256-019-0114-4

Moore, A. D. (2003). Privacy: Its meaning and value. *American Philosophical Quarterly, 40*(3), 215–227.

Moore, A. D. (2015). *Privacy, security and accountability: Ethics, law and policy*. Rowman & Littlefield.

Mulligan, D. K., Kroll, J. A., Kohli, N., & Wong, R. Y. (2019). This thing called fairness: Disciplinary confusion realizing a value in technology. *Proceedings of the ACM on Human-Computer Interaction, 3*(CSCW), 1–36. https://doi.org/10.1145/3359221

Mullin, J. (2021, February 23). *Student surveillance vendor Proctorio files SLAPP lawsuit to silence A Critic*. Electronic Frontier Foundation. https://www.eff.org/deeplinks/2021/02/student-surveillance-vendor-proctorio-files-slapp-lawsuit-silence-critic. Accessed 24 Jul 2021

Nissenbaum, H. (2009). *Privacy in context: Technology, policy, and the integrity of social life*. Stanford University Press.

Norman, D. (2020). Online exam proctoring. *D'Arcy Norman Dot Net*. https://darcynorman.net/2020/03/31/online-exam-proctoring/. Accessed 24 Jul 2021

O'Neil, C. (2016). *Weapons of math destruction: How big data increases inequality and threatens democracy*. Broadway Books.

Pardo, A., & Siemens, G. (2014). Ethical and privacy principles for learning analytics. *British Journal of Educational Technology, 45*(3), 438–450.

Partners, T. I. (2021, February 23). *Online exam proctoring market to reach US$ 1,187.57 million by 2027*. GlobeNewswire News Room. https://www.globenewswire.com/news-release/2021/02/23/2180512/0/en/Online-Exam-Proctoring-Market-to-Reach-US-1-187-57-million-by-2027-COVID-19-Impact-and-Global-Analysis-by-The-Insight-Partners.html. Accessed 24 Jul 2021

Paterson, J., & Bant, E. (2020). Contract and the challenge of consumer protection legislation. In T. Arvind & J. Steele (Eds.), *Contract Law and the Legislature: Autonomy, Expectations, and the Making of Legal Doctrine* (1st ed., pp. 79–106). Hart Publishing.

Persbureau, H. (2020, June 12). *Court sides with University of Amsterdam on use of surveillance software*. Cursor. https://www.cursor.tue.nl/en/news/2020/juni/week-2/court-sides-with-university-of-amsterdam-on-use-of-surveillance-software/. Accessed 24 Jul 2021

Peters, J. (2020, June 8). *IBM will no longer offer, develop, or research facial recognition technology*. The Verge. https://www.theverge.com/2020/6/8/21284683/ibm-no-longer-general-purpose-facial-recognition-analysis-software

Proctorio. (2020). *A comprehensive learning integrity platform*. https://proctorio.com/. Accessed 24 Jul 2021

ProctorU. (2020a). *Harnessing the power of artificial intelligence to improve online proctoring*. ProctorU. https://www.proctoru.com/harnessing-the-power-of-artificial-intelligence. Accessed 24 Jul 2021

ProctorU. (2020b). *ProctorU - The leading proctoring solution for online exams*. ProctorU. https://www.proctoru.com/. Accessed 24 Jul 2021

Quinn, T. P., Senadeera, M., Jacobs, S., Coghlan, S., & Le, V. (2021). Trust and medical AI: The challenges we face and the expertise needed to overcome them. *Journal of the American Medical Informatics Association, 28*(4), 890–894.

Rachels, J. (1975). Why privacy is important. *Philosophy & Public Affairs, 4*(4), 323–333.

Reddy, S., Allan, S., Coghlan, S., & Cooper, P. (2020). A governance model for the application of AI in health care. *Journal of the American Medical Informatics Association, 27*(3), 491–497.

Reiman, J. H. (1995). Driving to the panopticon: A philosophical exploration of the risks to privacy posed by the highway technology of the future. *Santa Clara Computer & High Tech. LJ, 11*(1), 27.

Rios, J. A., & Liu, O. L. (2017). Online proctored versus unproctored low-stakes internet test administration: Is there differential test-taking behavior and performance? *American Journal of Distance Education, 31*(4), 226–241.

Selbst, A. D., Boyd, D., Friedler, S. A., Venkatasubramanian, S., & Vertesi, J. (2019). Fairness and abstraction in sociotechnical systems. *Proceedings of the Conference on Fairness, Accountability, and Transparency, 19*, 59–68. https://doi.org/10.1145/3287560.3287598

Selwyn, N., O'Neill, C., Smith, G., Andrejevic, M., & Gu, X. (2021). A necessary evil? The rise of online exam proctoring in Australian universities. *Media International Australia, 107*, 2411–2502. https://doi.org/10.1177/1329878X211005862

Siegel, D. C., Harvey, Phillips, & Callan, E. (2018). Philosophy of education. In E. N. Zalta (Ed.), *The Stanford Encyclopedia of Philosophy* (Winter 2018). Metaphysics research lab, Stanford University. https://plato.stanford.edu/archives/win2018/entries/education-philosophy/. Accessed 24 Jul 2021

Simonite, T. (2020). *Meet the secret algorithm that's keeping students out of college*. Wired. https://www.wired.com/story/algorithm-set-students-grades-altered-futures/. Accessed 24 Jul 2021

Sloboda, J. A. (1990). Combating examination stress among university students: Action research in an institutional context. *British Journal of Guidance & Counselling, 18*(2), 124–136.

Srikanth, M., & Asmatulu, R. (2014). Modern cheating techniques, their adverse effects on engineering education and preventions. *International Journal of Mechanical Engineering Education, 42*(2), 129–140. https://doi.org/10.7227/IJMEE.0005

Stuber-McEwen, D., Wiseley, P., & Hoggatt, S. (2009). Point, click, and cheat: Frequency and type of academic dishonesty in the virtual classroom. *Online Journal of Distance Learning Administration, 12*(3). https://www.westga.edu/~distance/ojdla/fall123/stuber123.html. Accessed 20 Aug 2021

Swauger, S. (2020a). *Software that monitors students during tests perpetuates inequality and violates their privacy*. MIT Technology Review. https://www.technologyreview.com/2020/08/07/1006132/software-algorithms-proctoring-online-tests-ai-ethics/. Accessed 24 Jul 2021

Swauger, S. (2020b, April 2). *Our bodies encoded: Algorithmic test proctoring in higher education*. Hybrid Pedagogy. https://hybridpedagogy.org/our-bodies-encoded-algorithmic-test-proctoring-in-higher-education/. Accessed 24 Jul 2021

Sweeney, M., Lester, J., & Rangwala, H. (2015). Next-term student grade prediction. *IEEE International Conference on Big Data (big Data), 2015*, 970–975. https://doi.org/10.1109/BigData.2015.7363847

Thomson, J. J. (1975). The right to privacy. *Philosophy & Public Affairs, 4*(4), 295–314.

Torino, B. (2020, June 2). *Data privacy vs. European universities: Online proctoring*. Medium. https://medium.com/@brunnavillar/data-privacy-vs-european-universities-online-proctoring-be38e64fe080. Accessed 24 Jul 2021

Vallor, S. (2016). *Technology and the virtues: A philosophical guide to a future worth wanting*. Oxford University Press.

Warren, S. D., & Brandeis, L. D. (1890). The right to privacy. *Harvard Law Review, 4*(5), 193–220.

White, N. (2020, April 22). "Creepy" software to stop university students cheating in online exams. *Daily Mail Australia*. https://www.dailymail.co.uk/news/article-8243637/Creepy-software-used-stop-university-students-cheating-online-exams-amid-coronavirus.html. Accessed 24 Jul 2021

Zhou, N. (2020). Students alarmed at Australian universities' plan to use exam-monitoring software. *The Guardian*. https://www.theguardian.com/australia-news/2020/apr/20/concerns-raised-australian-universities-plan-use-proctorio-proctoru-exam-monitoring-software. Accessed 24 Jul 2021

Zuboff, S. (2015). Big other: Surveillance capitalism and the prospects of an information civilization. *Journal of Information Technology, 30*(1), 75–89.