Policing and Society

An International Journal of Research and Policy

ISSN: (Print) (Online) Journal homepage: www.tandfonline.com/journals/gpas20

'To whom am I speaking?'; Public responses to crime reporting via live chat with human versus AI police operators

Ben Bradford, Arabella Kyprianides, Will Andrews, Elizabeth Aston, Estelle Clayton, Megan O'Neill & Helen Wells

To cite this article: Ben Bradford, Arabella Kyprianides, Will Andrews, Elizabeth Aston, Estelle Clayton, Megan O'Neill & Helen Wells (19 Jan 2025): 'To whom am I speaking?'; Public responses to crime reporting via live chat with human versus AI police operators, Policing and Society, DOI: 10.1080/10439463.2025.2453437

To link to this article: https://doi.org/10.1080/10439463.2025.2453437

POLICING AND SOCIETY
https://doi.org/10.1080/10439463.2025.2453437

‘To whom am I speaking?’; Public responses to crime reporting via live chat with human versus AI police operators

Ben Bradford[a], Arabella Kyprianides[a], Will Andrews[b], Elizabeth Aston[c], Estelle Clayton[c], Megan O’Neill[d] and Helen Wells[b]

[a]Department of Security and Crime Science, University College London, London, UK; [b]School of Social Sciences, Keele University, Keele, UK; [c]Scottish Institute for Policing Research, Edinburgh Napier University, Edinburgh, Scotland; [d]School of Humanities, Social Sciences and Law, University of Dundee, Dundee, Scotland

ABSTRACT
Driven by social and technological change and the imperative to enhance efficiency, police have in recent years adopted various technologies to transform their interactions with the public. In the UK, these initiatives often fall under ‘transformation’ agendas, promoting ‘channel choice’ strategies to facilitate public interactions through various technologically mediated platforms, such as reporting crimes online using form-based or chat functions. Artificial Intelligence already plays a role in some of these interactions, which is likely only to increase in the future. In this study we examine preferences and perceptions in online crime reporting. Participants read a fictitious ‘chat’ between a victim of crime and a police operator identified as either a human or a chatbot. Although the chats were identical, we find a consistent preference for human operators over chatbots across all scenarios. Human operators were thought to provide clearer explanations, although there were no significant differences in judgements of interpersonal treatment or decision neutrality between human and chatbot operators. Participants also responded more positively to the process when (a) the crime involved was less serious and (b) when the outcome was active (police attendance) rather than passive (simple recording). Our findings underscore the importance of procedural justice and communication clarity in online crime reporting systems – and perhaps of human interaction when reporting crimes.

ARTICLE HISTORY
Received 25 June 2024
Accepted 9 January 2025

KEYWORDS
Online crime reporting; AI operator; human operator; procedural justice

Introduction

Police organisations have in recent years introduced various technologies that change how they interact with the public. In the United Kingdom, the site of the current study, the introduction of body-worn video, mobile data terminals and the Single Online Home (SOH; an online portal for the public to access information or contact police, for example via an online reporting form), and an apparently ever-increasing numbers of police social media accounts mean that police–public contact is increasingly likely to be technologically mediated in some capacity. Such developments are often heralded as ‘transformative’, and many coalesce around the notion of ‘channel choice’, increasing the number of ways people can interact with police. From a police perspective this

CONTACT Arabella Kyprianides a.kyprianides@ucl.ac.uk Department of Security and Crime Science, UCL, Shropshire House, 1 Capper street, London, WC1E 6JA

Supplemental data for this article can be accessed online at https://doi.org/10.1080/10439463.2025.2453437.

counts as a 'win-win' – the public are given greater choice, while many interactions shift online, thus becoming, at least in theory, easier to manage. Artificial Intelligence and other machine-based tools are increasingly part of the rhetoric, and in some cases the reality, of these developments, as they appear to offer the potential for police to handle an increasing range of interactions and, potentially, decisions without the need for input from expensive, slow, human operators.

This process of change has coincided with a period of intense pressure on police organisations, and on the institution itself. Policing in the UK has been buffeted by a series of high-profile cases of individual and organisational wrongdoing in recent years, including the murder of Sarah Everard by a serving police officer in 2021, continued fallout from the Hillsborough disaster[1] and the Stephen Lawrence enquiry,[2] increasing public dissatisfaction with the way police handle mundane reports of crime and disorder, and evidence of a long-standing and now potentially embedded decline in public trust and popular legitimacy.[3] In response, police have increasingly embraced the concept and language of procedural justice as a way to rebuild trust and legitimacy. The most visible example of this trend is His Majesty's Inspectorate of Constabulary and Fire & Rescue Services (HMICFRS) 'PEEL' inspections (the L stands for legitimacy). In its statutory inspections of forces across England and Wales, HMICRS considers the principle of procedural justice to make up a vital aspect of any assessment of the legitimacy (HMICFRS 2017). Improving the quality of police-public interaction stands at the heart of such efforts, not least because the contacts people have with police are something over which the latter have some control and which appear amenable to policy change and development.

The relationship between these two trends has received little attention, however, in policing or the academy (Wells *et al.* 2023). Technological developments in policing that insert technology 'into' interactions between police and policed have been initiated with little regard to how they will be received by the public. To date very little research has explicitly considered what impact increasing technological mediation will have on public trust and police legitimacy, or indeed how the mediating presence of technology affects how people judge the fairness of their interactions with police (for exceptions, see Hobson *et al.* 2021, Saulnier and Sivasubramaniam 2021, Aston *et al.* 2023, Hermstrüwer and Langenbach 2023, Saulnier 2023). The extant literature on procedural justice, police-public interaction and contact that has heavily influenced policing and policing scholarship has only just begun to consider what looks to be a fundamental shift in the nature of those interactions. Although technologically mediated contacts may still offer procedural justice, as studies on how social media companies interact with users have shown (Yuan and Lou 2020, Schoenebeck *et al.* 2021), exactly how this plays out in *policing* contexts remains under-explored.

This paper examines technologically mediated online crime reporting experiences by comparing 'live chat' interactions with police representatives and AI-operated chatbots.[4] Participants are presented with a scenario involving either graffiti on private property or burglary, and provided with an account of the online reporting of these crimes. The study investigates how the identity of the call handler – human or chatbot – and crime type/seriousness impact individuals' perceptions and judgments about the reporting process. Additionally, the research explores how the outcome of the call, either dispatching an officer promptly (an active response) or acknowledging the information without further action (a passive response), affects participants' perceptions and judgments.

Contacting police online

Online interaction has become a prominent feature of UK policing, with the Single Online Home (SOH) serving as a key manifestation of this trend. The SOH aims to offer a 'nationally consistent, locally branded service' through a digital police station that allows users across the UK to 'interact' with police through various functions (Wells *et al.* 2023). These include reporting crimes, traffic incidents, or missing persons; providing information on existing cases; applying for licenses or compensation; requesting reports; and giving feedback – all without directly encountering a human operator. Additionally, individual forces increasingly utilise live chats, including with chatbots (that is, machine or AI operators) for non-emergency matters (HMICFRS 2020).

This shift towards online communication represents a significant change from traditional police-public interactions which, when instigated by the public, relied on phone calls, in-person encounters, or occasional paper communication. The 1988 British Crime Survey, for example, found that 56% of people had contacted the police at least once in the 14 months before interview. Some 9% had called the 999 emergency number, 26% had telephoned a police station, 23% had approached a police officer on the street, and 5% had used some other mode such as, presumably, postal communication (Skogan 1990). By 2020, when 26% of people in England and Wales reported some form of contact with the police in the previous year, 6% had interacted with police over the phone; 2% at a police station; 2% via social media; 2% via email; 1% had approached officers on the street; and 1% used a police website (ONS 2022). Many of the other forms of contact reported were police instigated, including 'they knocked at my door' (6%), or related to public meetings and events (5% or more). There has, in other words, been a clear shift away from telephone and particularly face-to-face interaction, at least among public-initiated contacts.

To the extent that police have 'pushed' towards technologically mediated interaction two key motivations underpin these developments. Firstly, police organisations aim to manage calls for service more effectively and efficiently. The increasing volume and complexity of contacts have led to efforts to improve triage and implement automated systems (HMICFRS 2020, Wells *et al.* 2023). Secondly, the National Police Chief's Council (NPCC) seeks to match changing public expectations by offering online services that align with how people interact with other institutions, like banks and retailers (NPCC n.d.). This is believed to enhance service quality and improve public trust, though the link between online contact and trust remains largely untested. Moving police contact online (and thus, it is hoped, making it more timely and efficient) is assumed to enhance trust, that is, but precisely how and why this is the case is rarely explored, nor are any potential downsides explicitly considered.

Despite intentions to standardise and improve encounters with the public, there has been limited research on how technological mediation impacts procedural justice perceptions, satisfaction, and public trust. The current study seeks to address this gap by exploring how interactions with human operators and chatbots affect these perceptions, drawing on literature about human-machine interactions and procedural justice in technologically-mediated contexts.

Interacting with machines

Chatbots are algorithms or computer agents that provide outputs using rule-based, machine learning or artificial intelligence (AI) approaches to simulate conversation with a human end user (Jussupow *et al.* 2020). Research on how people perceive and interact with algorithms and AI – we use these terms interchangeably – suggests that they see algorithms as lacking agency and emotion but possessing greater accuracy and rationality than humans (Kleinberg *et al.* 2018, Lee 2018). This preference for accuracy can lead some to favour AI over human decision-makers, particularly in areas prone to human error and bias (McGuire 2021, Miller and Keiser 2021). In other words, some people who contact the police may *prefer* to speak to a machine, or at least not mind if they do, particularly if they think it is more likely to be 'right'. Moreover, they may – despite widespread concerns about algorithmic bias – believe the machine is less likely than a human operator to be biased or prejudicial, apparently because the automated processes were found to be more neutral, respectful and trustworthy. Saulnier and Sivasubramaniam (2021), for example, found that technologically mediated interactions at security checkpoints at airports enhanced perceptions of fairness and trustworthiness.

However, algorithmic decisions can also be perceived as less fair than those made by police officers (Hobson *et al.* 2021, Hermstrüwer and Langenbach 2023). Dietvorst *et al.* (2015) describe 'algorithmic aversion', wherein people may reject AI decisions due to perceived imperfections or distrust (Longoni *et al.* 2019). Algorithmic aversion describes a complex set of reactions to AI, which include: false expectations that affect responses to algorithmic decision-making (for example the idea that error is systematic, 'baked in' and irreparable); concerns about decision control and in general a sense that the decision-maker cannot be considered trustworthy; and an emphasis on the need for human decision-making in contexts marked by uncertainty (Burton *et al.* 2020). While Dietvorst *et al.* (2015, 2018) position aversion as emerging primarily from perceptions or experiences of imperfection (i.e. things going wrong, which seems to evoke a strong response when it is a machine making the error), other research has suggested that aversion can arise before such experiences (Longoni *et al.* 2019) – the aversion itself can thus be baked in to people's assessments of algorithms. This aversion differs from technological resistance, which stems from a preference for the status quo or rejection based on quality factors (Jussupow *et al.* 2020).

Relatedly, reliability, or making the correct decision, plays a crucial role in assessments of AI (Glikson and Woolley 2020, Kaplan *et al.* 2023). People may have a more instrumental relationship with machines than with people, accentuating their focus on accuracy and outcome favourability. Studies also show that AI decisions may be more accepted in routine tasks or low-cost scenarios, while use in more critical or complex situations may amplify negative reactions (Lee 2018, Mozafari *et al.* 2021, Nagtegaal 2021, Filiz *et al.* 2023). For example, Lee (2018) found that when the decision-maker (either algorithmic or human) was making a managerial decision about a mechanical task (for example scheduling employee's shift patterns), algorithm and human-made decisions were perceived as equally fair and trustworthy. However, when the decision-maker was considering a human task (e.g. hiring), algorithms were perceived as less fair and trustworthy and evoked more negative emotions.

Chatbots and procedural justice

Procedural justice is a fundamental lens through which individuals evaluate interactions with police agents (Tyler 2006, Tyler and Huo 2002). In technology-mediated environments, there is good evidence to suggest procedural justice remains important (e.g. Rabinovich-Einy and Katsh 2014, Saulnier and Sivasubramaniam 2021, Tyler *et al.* 2021), although its influence may manifest differently compared to human interactions, and in potentially complex and even contradictory ways. Research on trust in algorithmic decision-making highlights the importance of transparency, responsiveness, and procedural fairness in this context (Glikson and Woolley 2020). However, while algorithmic decision-makers may effectively demonstrate neutrality, other elements of procedural justice, such as voice and decision control, may be harder to convey digitally (Wells 2008, Saulnier and Sivasubramaniam 2021). This could lead to a shift in how different components of procedural justice are prioritised in human versus algorithmic encounters.

For example, the early iterations of procedural justice theory emphasised how having a voice, or input, in decision-making processes offers reassurance and contributes to judgments of both procedural justice and outcome-related concerns such as distributive justice (Thibaut and Walker 1975, Walker *et al.* 1979, Lind 2001). This implies that in online or algorithmic contexts, where processes are less transparent and where people may lack other cues to gauge outcomes, individuals may rely more heavily on cues of procedural justice to assess the overall fairness of decisions. By contrast, though, the literature on algorithmic aversion suggests that individuals might inherently distrust algorithmic decision-making, regardless of their direct experiences (Dietvorst *et al.* 2015, Longoni *et al.* 2019). As outlined above, this distrust can stem from perceptions of imperfection, concerns about decision control, and a preference for human involvement in uncertain situations. This may lead to different responses to online interactions with police chatbots, particularly in higher-cost or complex scenarios (Jussupow *et al.* 2020).

Another important aspect of procedural justice is explanation: the ability of the authority to clarify its decisions, demonstrating neutrality and impartiality (Tyler 2003). This has important implications in digital interactions, where users may trust algorithmic decision-makers less due to perceived limitations in their ability to provide comprehensive explanations – due, not least, to the 'black box' nature of AI decision-making (Mahmud *et al.* 2022). At the same time, some may view algorithms as inherently rational, neutral, and unbiased (key components of procedural justice) and therefore consider their decisions *more* trustworthy.

A further potential issue relates to the identity processes that lie at the heart of procedural justice theory (Sunshine and Tyler 2003, Tyler and Blader 2003). A central reason for the importance of procedural justice is that it communicates messages of status, belonging and inclusion within groups. On this account, police represent social groups many, even most, people find important – the nation, state, or community – and procedural justice signals to people interacting with officers that they, as representatives of those groups, recognise and affirm belonging within them (see Chan *et al.* 2023 for a recent review of these issues). This process may rely, though, on people's understanding that police officers are moral agents *capable* of such recognition. Moral agents are those governed by moral standards, who thus have both moral obligations (Himma 2009), such as to treat those subject to them with procedural justice, and the capacity to recognise the feelings, intentions and fundamental similarity – and thus moral standing – of others (Pesch 2020).

People may be less inclined to believe machines can be moral agents capable of recognising them in this way. Bellaiche *et al.* (2023), for instance, found that people preferred works of art they thought were produced by a human over those produced by an AI, *even though they were the same works*, seemingly because they thought the human work more morally authentic – that is, produced by an entity capable of moral thought and reasoning. Banks (2019) found that a variety of machine actors were perceived as less morally agentic than a human comparator; less likely to have a sense of right and wrong, less likely to refrain from actions with painful outcomes, and only able to behave on the basis of their programming. One implication of this perceived lack of moral agency may be that people are less attuned to signals of procedural justice in the behaviour of machines. They are inclined to believe those machines are incapable of recognising them as citizens or simply human beings worthy of respect, unable to see that treating others with respect is the right thing to do, and not able to judge the intrinsic moral worth of a person.

Extant work on automated systems in policing contexts seems to support these ideas. Both Wells (2008) and Merola *et al.* (2019) found that automated traffic managing processes can be seen as less fair than those involving people, even if the automated system provides objectively certain, or at least better, judgements, seemingly because such experiences may 'violate relational expectations by communicating undesirable (suspect or criminal) identities [without providing] the positive identity affirming information associated with procedurally just interpersonal treatment' (Saulnier and Sivasubramaniam 2021, p. 320).

The interaction between the nature of police authority and judgements of algorithmic decision-making might exert further influence. In digital interactions, users bring their social learning and preconceptions to the encounter (Nass and Moon 2000), including attitudes toward police authority. They bring, that is, a particular set of expectations to the encounter (one of which is that police will or at least should treat them with procedural justice), and they will look for evidence that these expectations have been met. If they do not believe that a machine can meet these expectations, for example because it lacks moral agency or meaningful authority, they may be primed to be less satisfied with what transpires simply because they went into the encounter assuming their interlocutor cannot provide some of what they require from it.

Thus, while it seems likely that procedural justice will remain important in technologically mediated police-public interactions, its dynamics might differ from traditional, human-to-human encounters, and perceptions of procedural justice might be influenced by both the nature of the interaction and individuals' existing attitudes toward both the police and 'machines' or AI. There is also much to suggest that people may be less attuned to signals of procedural justice in the behaviour or communication of machines.

Hypotheses

In this study we explore experimentally the impact of the reporting operator in online crime reporting experiences by comparing live chat with police representatives to chatbots with AI operators ('machines'). Participants in an on-line study read a 'chat' exchange between a police operative and someone reporting a crime. We manipulate the identity of that operator (human or chatbot), the seriousness of the crime involved (graffiti or burglary), and the immediate outcome offered to the protagonist ('active' police attendance vs 'passive' recording of the crime) but otherwise keep the 'chat' exchange the same. Drawing on the discussion above, six hypotheses guide our analysis.

H1: Given the same 'service', people will judge the human operator more favourably (fairer and more satisfactory) than the machine. The idea of algorithmic aversion suggests that all else equal people will prefer a human, while the notion that machines lack moral agency will mean that they will be less attuned to cues of procedural justice in interactions with the chatbot.

H2: Respondents in the burglary condition will favour the human operator by a greater margin than those in the graffiti condition. When the situation seems more serious there is a stronger preference for human interaction.

H3: Procedural justice will be a stronger predictor of perceptions of outcome fairness in the machine compared with the human conditions. Because AI decision-making processes are seen as more opaque, respondents will rely on fairness judgements to an (even) greater extent when they believe a machine is making the decision.

H4: Respondents will judge decisions made by humans as better explained. Algorithmic aversion indicates there will be a general unwillingness to 'listen' to what is 'said' by a machine, and people may believe it is impossible to understand how it makes its decisions.

H5: Respondents will judge decisions made by the machine as more neutral than decisions made by humans: there appears to be a relatively widespread view that this is a positive aspect of AI decision-making.

H6: Respondents will be more satisfied with the passive outcome in the human conditions: positive outcomes are more important for judging machines positively.

Methodology

Participant information

Participants were recruited from Prolific panels and completed the survey online in August 2023 (pilot) and September 2023 (full study). Participants were compensated £3.50 for their time. No exclusion criteria were applied beyond ensuring that all participants were located in the UK. The sample distribution was designed to reflect the UK population, as represented in the 2021 Census. A pilot study ($N = 80$) tested the effectiveness of the stimuli, specifically manipulating crime type (burglary vs. graffiti), operator type (human vs. AI), and outcome (active vs. passive). The results indicated that participants responded more positively to graffiti-related crimes, human operators, and active outcomes compared to burglary-related crimes, AI operators, and passive outcomes. Only results from the full study are reported here. [University name] granted us an ethics exemption certificate for this study.

The full study involved a sample of 640 participants. We intentionally oversampled to ensure that our study had sufficient power across all comparisons.[5] The sample represents a good mix of participants across potentially relevant variables such as gender, age and economic status. Gender distribution was nearly equal, with 48% male ($n = 348$) and 52% female ($n = 372$). Regarding age, 8% of participants fell into the 18–24 age group ($n = 60$), 22% were aged 25–34 ($n = 156$), 21% were in the 35–44 age category ($n = 150$), 16% were aged 45–54 ($n = 117$), 20% were between 55 and 64 ($n = 142$), 12% were aged 65–74 ($n = 83$), and 2% were 75 years or older ($n = 13$). The majority of participants were employed (63%, $n = 455$), with 15% being retired ($n = 110$), 9% self-employed ($n = 61$), 3% currently studying ($n = 20$), 3% unable to work ($n = 18$), 3% out of work or actively seeking employment ($n = 24$), 3% identified as homemakers ($n = 24$), and 1% either reported being out of work and not seeking employment ($n = 5$) or chose 'other' ($n = 4$). Education levels varied, with 55% holding at least a bachelor's degree or its equivalent ($n = 399$), 15% possessing a master's degree or its equivalent ($n = 109$), 14% having two or more A-levels or their equivalent ($n = 100$), 11% holding a diploma or its equivalent ($n = 82$), 10% having completed five or more GCSEs or their equivalent ($n = 75$), 5% with 1–4 GCSEs or their equivalent ($n = 38$), 1% reporting either skills for life ($n = 6$) or a doctoral degree ($n = 9$), and 2% selecting 'other' ($n = 12$). Regarding health, 21% of the participants reported having a long-standing illness, disability, or infirmity ($n = 152$), while 79% indicated that they did not have any such condition ($n = 569$).

The majority of participants (81%, $n = 584$) identified themselves as White, specifically as English, Welsh, Scottish, Northern Irish, or British. A small percentage of participants (4%, $n = 32$) indicated other White backgrounds, while 2% identified as Asian/Asian British (Indian: $n = 13$, Chinese: $n = 14$) or Black/African/Caribbean/Black British (African: $n = 13$). Additionally, 1% of participants reported their ethnicity as White (Irish: $n = 8$), Asian/Asian British (Pakistani: $n = 9$, Bangladeshi: $n = 5$), Mixed/Multiple ethnic groups ($n = 6$), Any other Asian background ($n = 9$), or Black/African/Caribbean/Black British (Caribbean: $n = 8$). A further 1% of participants chose not to disclose their ethnicity ($n = 7$). Ethnic groups not explicitly mentioned in this breakdown comprised less than 0.5% of the sample.

Experimental design

We exposed participants to a story involving either the presence of graffiti on private property or a house burglary. Following that, we guided them through the victim's online crime reporting experience (note that these were not crimes in progress). By comparing reporting via live chat with two different operators, one a police operator and the other an AI operator (chatbot), but holding the content of the chat constant, we investigated the impact of the operator and crime type on perceptions of and judgements about the process. Furthermore, we explored how the outcome of the reporting process interacted with perceptions of and judgements about the process. The active outcome entailed the decision to dispatch an officer to investigate the reported crime 'as soon as possible', while the passive outcome involved acknowledging the provided information without further action.

The study therefore employed a $2 \times 2 \times 2$ factorial design involving three independent variables: operator type (live chat with a police representative or an AI-operated chatbot), crime type (graffiti or burglary), and outcome (active or passive). This design allowed for a comprehensive exploration of the impact of operator type, crime type, and reporting outcome on participants' perceptions and judgments.

Procedure

All participants first received introductory materials providing information about the study and obtaining their consent to participate. Subsequently, participants were required to input their Prolific ID to ensure eligibility for compensation. They were then randomly assigned to one of two narratives: a story featuring unauthorised graffiti on private property or a house burglary. In the graffiti scenario, Jane, returning to her flat after work, sees the graffiti and opts to report it online, despite her mixed emotions. In the burglary scenario, Jane returns home to find her flat burgled, leading her to report the crime online, overwhelmed by a sense of violation. We emphasised that Jane was more affected by the burglary to indicate that this was a more serious crime, i.e. the distinction between the two crimes concerns the seriousness of the crime *to Jane*, but since there is no absolute scale of crime seriousness, we needed to make the effect of the crime on her clear to participants. The text of these scenarios is available in Appendix 2.

Subsequently, participants were again randomly assigned to experience the victim's online reporting process, either through reading a text based 'live chat' with a human ('Robin') or AI operator ('Bobbybot'). In this phase, Jane reports the crime (either graffiti or burglary, depending on the condition they were initially assigned to) to either Robin or Bobbybot. The reporting process involves providing details, experiencing a 5-minute delay, and concluding the text-based conversation with expressions of gratitude and assurance. Transcripts are again available in Appendix 2. The content of these chats, and how they unfolded, was based on observational and interview work conducted in police call handling centres, conducted as part of the larger project of which the current study is just one small part.

Participants were then presented with a series of questions regarding their perceptions of the fairness of the online conversation between Jane and Robin, or Jane and Bobbybot.

Subsequently, the chat between Jane and Robin/Bobbybot continued. Here, participants were randomly assigned to observe one of two possible outcomes of the reporting process. The active outcome involved dispatching an officer to investigate the reported crime 'as soon as possible', while the passive outcome merely acknowledged the information provided without further action. Transcripts of these chat conversations are available in Appendix 2.

Finally, participants responded to a series of questions regarding their perceptions of the fairness of the outcome presented in the conversation they had just read. They were also asked to rate their overall satisfaction with the process, and to answer a series of questions about their views on how the decision reached in the story was communicated to Jane. To close, participants were asked to provide their demographic information (gender, age, employment, education, disability, ethnicity), and were thanked for their participation and debriefed.

The study incorporated a manipulation check: participants were asked whether Jane was interacting with a human operator or an AI operator. All participants in the AI conditions correctly identified that Jane was talking to a machine. However, 65 out of 320 (20%) participants in the *human operator* conditions incorrectly answered that she was talking to a machine. We exclude these participants from our analysis on the basis that they apparently failed to notice our central experimental manipulation. However, the imbalance in these responses may reflect something significant about the way participants experienced the chat scenario, i.e. some seem to have assumed that a text-based live chat involved interaction with a machine.

Measures

The survey items either related directly to the content of the vignettes or were drawn from existing batteries. Question sets included: Perceptions of the fairness of the interaction process (procedural justice); Perceptions of the fairness of the outcome; Views on how the decision reached in the vignette was communicated to Jane; and Overall satisfaction with the process.[6]

Dependent variables

We used principal component analysis to assess the scale properties of the two constructs central to our analysis, perceptions of procedural justice and perceptions of outcome fairness. Having found reasonably good scale properties (i.e. all items loaded onto one main component), we calculated principal component scores from each separate model to create the two indices concerned. Perceptions of *procedural justice* was measured by a scale derived from nine items probing judgements about the procedural fairness of the police operative's behaviour (Jackson and Bradford 2019), including 'Thinking about the interaction between Jane and the police, would you agree or disagree that the police:' e.g. 'treated Jane with respect and dignity', 'Were impartial, i.e. acted in a fair and neutral manner', 'Listened to Jane before making decisions' ($a = .85$). All items loaded strongly onto one component, which was extracted and saved for analysis (mean = 0; SD = 1; min = −2.5; max = 2.7). Perceptions of *outcome fairness* was measured by a scale derived from three items ('Jane got the response from police that she deserved', 'Jane received an outcome similar to that others would get in the same situation', 'Some types of people receive a better service in these cases than others'; Tyler *et al.*, 2006) ($a = .82$). All items loaded strongly onto one component, which was extracted and saved for analysis (mean = 0; SD = 1; min = −3.4; max = 1.6).

Other outcome variables were measured by single items. We include two measures of procedural justice relating to the way the police response or outcome was decided (Jackson and Bradford 2019): 'What police were going to do was explained clearly to Jane' and the police 'Made decisions based on the facts'. Responses were on 5-point agree/disagree scales in each case. Finally, overall satisfaction with the process was measured by the item 'If you were Jane, would you have been satisfied with how the police dealt with this case?' (Bradford and Jackson 2010).

All study materials have been uploaded to a secure OSF site: [link].

Results

To address our hypotheses we use regression analysis, primarily for simplicity of presentation. We use linear regression models where the outcome variables are the measures of procedural justice and outcome fairness, and ordinal logistic regression models where the outcome variables are views on clarity of expression, decision-making based on facts, and overall satisfaction with the process.

Judgements about process fairness

Table 1 shows results from two linear regression models predicting perceptions of procedural justice. These provide partial tests of H1 (Given the same 'service', people will judge the human operator more positively than the machine) and H2 (Respondents in the burglary condition will favour the human operator by a greater margin than those in the graffiti condition). Model 1 shows that, holding the crime type constant, participants judged the human operator to be fairer than the chatbot. Holding the operator constant, police behaviour was judged to be fairer in the graffiti case than in the burglary case. Model 2 adds the interaction between operator and crime type, which is not significant in the model. In other words, the 'preference' for the human operator did not vary significantly by crime type.

Fitted values calculated from Model 1 illustrate the magnitude of these effects – notably, the difference between the most favourable condition (human, graffiti) and the least favourable (chatbot, burglary) was substantial, around .7 of a standard deviation.

Judgements about the outcome

Table 2 turns to judgements about outcome fairness, which provides further evidence in relation to H1 and H2. Model 1 shows that the outcome tended to be judged as fairer when there was a human compared with a chatbot operator, when the crime involved was graffiti rather than burglary, and when there was an active rather than a passive outcome. Considering potential interaction effects (Models 2–4), we found there was not a significant interaction between operator and outcome – the importance of the outcome did not vary by the nature of the operator. Neither did we find an interaction between operator and crime type. However, we did find a significant negative interaction between crime type and outcome – whether the outcome was active or passive mattered less when the crime was graffiti. Finally, Model 5 adds the three-way interaction between operator, crime type and outcome: this was also non-significant.

Table 1. Linear regression models predicting perceptions of process fairness.

	Model 1	Model 2
Human (ref: Bot)	0.32***	0.39***
Graffiti (ref: burglary)	0.42***	0.48***
Interaction		
Human × Graffiti		−0.14
Constant	−0.31***	−0.34***
R^2	0.07	0.07
N	575	575
Fitted values from Model 1 (95% C.I.)		
	Human	Bot
Graffiti	.42 (.28, .56)	.10 (−.03, .24)
Burglary	−.004 (−.14, .14)	−.31 (−.45, −.18)

Table 2. Linear regression models predicting perceptions of outcome fairness.

	Model 1	Model 2	Model 3	Model 4	Model 5
Human (ref: Bot)	0.27***	0.28*	0.28*	0.28***	0.27
Graffiti (ref: burglary)	0.42***	0.42***	0.43***	0.69***	0.67***
Proactive outcome (ref: passive)	0.43***	0.44***	0.43***	0.69***	0.68***
Interactions					
Human × Active		−0.01			0.04
Human × Graffiti			−0.02		0.03
Graffiti × Active				−0.52***	−0.48*
Human × Active × Graffiti					−0.1
Constant	−0.53***	−0.53***	−0.54***	−0.66***	−0.65***
R^2	0.12	0.12	0.12	0.14	0.14
N	575	575	575	575	575
Fitted values from Model 4 (95% C.I.)					
	Human		Bot		
Graffiti, Active	.48 (.32, .65)		.21 (.04, .37)		
Burglary, Active	.32 (.13, .50)		.03 (−.14, .20)		
Graffiti, Passive	.32 (.14, .50)		−.04 (−.14, .21)		
Burglary, Passive	−.37 (−.54, −.20)		−.65 (−.81, −.49)		

Fitted values from Model 4 again illustrate the magnitude of these effects. There is a full standard deviation difference in judgements between the most favourable condition (human, graffiti, active) to the least favourite (chatbot, burglary, passive). Note also that the human, graffiti, passive outcome is judged as fair as (if not fairer than) the chatbot, graffiti, active outcome.

We also estimated models 'within' each of the four groups formed by the initial operator by crime type manipulation, within which respondents were subsequently randomly assigned to active or passive responses (see Appendix 1). This allows us to hold constant operator and crime type – which affected judgements of procedural justice – while allowing the latter to vary, thus providing a test of H3. The response variable in all models was outcome fairness: the three explanatory variables were outcome, perception of procedural justice and the interaction between the two. Procedural justice was a strong predictor of outcome fairness in all models, but the interaction term was not significant in any.[7] The importance of procedural justice in shaping perceptions of outcome fairness did not seem to vary between the human and chatbot conditions.

Perceptions of decision-making

Next, we consider respondents' views on how the decision-making process (H4 and H5). Table 3 shows results from ordinal logistic regression models predicting views on how well the decision was communicated; Table 4 shows results predicting views on whether the decision was made based on the facts. Participants were significantly more likely to judge the human clearer than the chatbot (see Table 3), and were more likely to think that the decision was made based on the facts in the human conditions (see Table 4). In other words, they were not more likely to think the AI operator made more fact-based decisions, as predicted by H5. As above, we find that participants were more likely to judge the clarity of explanation positively, and that the decision was made based on the facts, in the graffiti (compared with the burglary) and the active (compared with the passive) outcome conditions.

Overall satisfaction with the process

Finally, we turn to overall satisfaction with the process – or, at least, how satisfied respondents thought Jane would have been. We use ordinal logistic regression to model responses to the five-category response variable to test H6 (Respondents will be more satisfied with the passive

Table 3. Ordinal logistic regression models predicting clarity of expression.

	Model 1	Model 2	Model 3	Model 4	Model 5
Human (ref: Bot)	0.57***	0.59**	0.63**	0.58***	0.72*
Graffiti (ref: burglary)	0.46**	0.46**	0.51*	0.60**	0.75*
Active outcome (ref: passive)	0.95***	0.96***	0.95***	1.08***	1.17***
Interactions					
Human × Active		−0.02			−0.19
Human × Graffiti			−0.12		−0.32
Graffiti × Active				−0.28	−0.46
Human × Active × Graffiti					0.4
N	575	575	575	575	575

* $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$.

outcome in the human conditions). Table 5 shows the results, which are strikingly similar to those presented above. Model 1 in Table 5 shows that overall satisfaction tended to be higher when there was a human compared with a chatbot operator, when the crime involved was graffiti rather than burglary, and when there was a active rather than a passive outcome. The interaction between operator type and outcome was non-significant (Model 2), implying that outcome was equally important for overall satisfaction in the human and chatbot conditions. The interaction between operator and crime type was also not significant (Model 3). Yet, we also find, again, that the outcome was more important in the burglary compared to the graffiti conditions (Model 4).

Discussion

Six hypotheses were tested in this study. Respondents generally favoured human operators over chatbots, indicating a preference for human interaction in the online crime reporting process. Participants tended to view the human operator as more procedurally just than the chatbot, even though the chat was the same in both cases, and they were more satisfied overall with the human operator (H1). While human operators were preferred overall, there was no significant difference in preference for human operators across different crime types, suggesting that this preference held regardless of the type of crime reported (H2). Procedural justice was identified as a strong predictor of outcome fairness in both human and chatbot-operated conditions, highlighting the importance of transparent and just reporting procedures (H3). Respondents perceived human operators as offering clearer explanations compared to chatbot operators (H4). We also found evidence to suggest humans, not chatbots, were more likely to be seen to be making decisions based on facts, contrary to H5, which predicted the opposite. Finally, we did not find evidence that positive outcomes are more important for machine operators. Participants were not more satisfied with the passive outcome in the human conditions: overall satisfaction was influenced by various factors including the type of operator and type of crime, but not by the interaction between operator type and outcome (H6).

Our core finding, that participants generally favoured human operators over chatbots in online crime reporting, aligns with the concept of algorithmic aversion (Dietvorst *et al.* 2015), where

Table 4. Ordinal logistic regression models predicting views on whether the decision was made based on the facts.

	Model 1	Model 2	Model 3	Model 4	Model 5
Human (ref: Bot)	0.40*	0.46*	0.47*	0.41**	0.53+
Graffiti (ref: burglary)	0.37*	0.38*	0.44*	0.78***	0.85**
Active outcome (ref: passive)	0.71***	0.76***	0.71***	1.12***	1.17***
Interactions					
Human × Active		−0.12			−0.1
Human × Graffiti			−0.15		−0.14
Graffiti × Active				−0.82**	−0.82+
Human × Active × Graffiti					0.01
N	575	575	575	575	575

+ $p < .1$, * $p < .05$, ** $p < .01$, *** $p < .001$.

Table 5. Ordinal logistic regression models predicting outcome satisfaction.

	Model 1	Model 2	Model 3	Model 4	Model 5
Human (ref: Bot)	0.54***	0.51*	0.26	0.55***	0.24
Graffiti (ref: burglary)	1.24***	1.24***	0.98***	1.56***	1.28***
Active outcome (ref: passive)	1.45***	1.43***	1.45***	1.77***	1.76***
Interactions					
Human × Active		0.06			0.04
Human × Graffiti			0.6		0.66
Graffiti × Active				−0.66*	−0.63
Human × Active × Graffiti					−0.09
N	575	575	575	575	575

* $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$.

people exhibit a preference for human decision-making due to concerns about trustworthiness and a perceived need for human involvement in uncertain situations. This resonates with procedural justice theory, which emphasizes the importance of transparency, voice and respect in establishing trust (Tyler 2003). It may be that people are, on average, disinclined to believe that machines can truly 'do' these things, perhaps because they believe that machines lack the intentionality and perhaps the moral expertise of a human, both of which may underly concerns about procedural justice: to feel one is being treated with respect (or disrespect) would seem to imply that one's interlocutor is capable of respect. To put it another way, procedural justice is inherently relational – it expresses and makes real social bonds (Sunshine and Tyler 2003, Tyler and Blader 2003) – and people may, again, be less likely to believe they can have a relationship with a machine. They subsequently read interactions with machines in this light, one implication being that even when confronted with the same cues of procedural justice as those in the human conditions, participants in the machine conditions were less likely to pick up and respond to them, thus judging the machine interactions as consistently less fair.

While some studies have suggested that algorithmic policing decisions may be viewed as more or less fair depending on context (Hobson *et al.* 2021, Hermstrüwer and Langenbach 2023), our results identify what seems to be a persistent preference for human interaction in this specific context. This in turn seems to align with the idea that algorithmic aversion can exist even before encountering errors in AI systems (Longoni *et al.* 2019), possibly due to the perceived limitations in machine-mediated interactions noted above.

The concept of algorithmic aversion also indicates that there will be a general unwillingness to 'listen' to what is 'said' by a machine (Glikson and Woolley 2020, Jussupow *et al.* 2020). This aligns with our finding that respondents perceived human operators as offering clearer explanations compared to chatbot operators, despite identical police responses being offered in both cases. It is plausible to suggest that the perceived (in)ability to offer an explanation, a key component of procedural justice (Tyler 2003), may be crucial in shaping trust in digital interactions. We also found that respondents tended to believe the human operator was making decisions based on facts, which would seem to contradict the idea that algorithms can be seen as more neutral. The most likely explanation here is perhaps that this merely reflects the underlying preference for human operators (i.e. participants 'read across' from their generally more positive view of the human operator), but further work could usefully explore this finding, not least because more neutral decision-making is one of the potential benefits of AI brings to policing.

By contrast, our finding that there was no significant difference in the strength of preferences for human operators across crimes of different seriousness contradicts the notion that people prefer human interaction when the situation involved is more serious or consequential (Lee 2018, Mozafari *et al.* 2021). Moreover, our results did not support the hypothesis that participants would be more satisfied with passive outcomes in the human conditions; or, to put it another way, that people are more outcome-oriented when they are dealing with machines. Similarly, the finding that procedural justice was a strong a predictor of outcome fairness in both human and chatbot-operated conditions challenges the hypothesis that AI decision-making processes, being more opaque, would lead respondents to rely more heavily on fairness judgments when they believe a machine is making the decision. That said, it is also the case that procedural justice was no *less* important in the AI conditions.

Given the apparent importance of outcome related factors in judgements of AI (Kaplan *et al.* 2023), in a crime reporting scenario one would perhaps expect that compared with a human, interaction with a chatbot would been seen in more instrumental terms, and that people would attend more to outcomes and less to process when they are dealing with a machine. This is not what we found. More research is needed to explore this issue, but one possible explanation is that in policing contexts people remain strongly attuned to the relational and value-bearing aspects of police behaviour, and this persists even when the actor is non-human. Those who report crime to the police do so not just because they want their stolen goods back, or a crime reference number, but because they seek reaffirmation of their status and worth in society (Pemberton and Mulder 2023). This may be the case even when they know that the police representative is a machine, not a human. The centrality of relational concerns in crime reporting scenarios may be one reason that, even in the vicarious machine encounter created by our vignettes, people remain attuned to procedural justice.

In sum, the overall conclusion from this study seems to be that people do care about, and identify, procedural justice (or injustice) in interactions with machines – it is just that they tend to judge the behaviour of entities they believe to be human as more procedurally fair than identical behaviour on the part of entities they believe to be machines. In effect, the latter start from a lower base.

It is also worth considering some of our other findings. As well as the consistent preference to human operators, participants also judged the graffiti reporting process more favourably than the burglary process, and the active outcome as better than the passive, although this effect was largely confined to the burglary conditions. The latter is perhaps not surprising – i.e. people value an active police response to crime events, particularly when these are more serious – but the former suggests that online reporting of the type considered here should be confined to less serious crimes. This is indeed the current position across the UK forces we are aware of. Greater moves towards automation might not, on this evidence, be well received by those who need to use such systems.

Our findings overall echo key themes covered in the introduction of the paper, particularly the shift towards digital policing and its impact on public perceptions of procedural justice. The movement to manage calls for service more efficiently and align with changing public expectations (Wells *et al.* 2023) reflects a balance between streamlining processes and maintaining trust (NPCC n.d.). However, the strong preference for human operators in our study suggests that this balance is delicate. The tendency for people to distrust algorithmic entities (Dietvorst *et al.* 2015), and to emphasise the relational aspects of interactions within when these involve a machine, indicates a need for online policing tools to address procedural justice components, such as transparency, voice, and neutrality to build and sustain public trust. This underscores the importance of managing technological mediation effectively, considering the psychological elements that underpin perceptions of procedural justice (Tyler 2003) and ensuring that digital systems are designed to foster positive interactions, even in the absence of human involvement.

Policy implications

Our findings offer several practical implications for online crime reporting systems. First, the preference for human operators over chatbot operators highlights the continued importance of incorporating human interaction or assistance in digital policing systems. This preference echoes the concept of algorithmic aversion (Dietvorst *et al.* 2015), emphasising the need for human involvement to build and sustain trust in online reporting mechanisms. Police organisations should aim to maintain access to human operators wherever possible, balancing digital innovation with interpersonal support.

Second, while preferences for human operators did not vary significantly across different crime types, perceptions of procedural justice differed notably between graffiti and burglary reporting conditions. This indicates that crime type may influence people's expectations of fairness and procedural justice, with more serious offenses possibly prompting higher expectations of procedural justice. Further research should explore these dynamics to inform how police can tailor their online systems and responses to varying public expectations (Tyler 2003).

Third, procedural justice was a strong predictor of perceptions of outcome fairness and overall satisfaction in both human and chatbot conditions, underscoring the need for transparent, fair, and procedurally just interactions. This aligns with existing literature, which emphasises the importance of elements like neutrality and voice in fostering trust (Rabinovich-Einy and Katsh 2014, Glikson and Woolley 2020). Police organisations considering automated systems should focus on enhancing clarity and comprehensibility in responses, ensuring clear and respectful communication, adherence to established procedures, and fair treatment.

Finally, outcome satisfaction was influenced by both the type of operator and the type of crime, indicating a holistic approach is needed to manage online reporting processes (Mozafari *et al.* 2021). Active responses may lead to higher satisfaction, particularly in serious offenses, but managing expectations and communicating effectively when outcomes are passive is essential. This holistic management can help balance procedural and distributive justice considerations, making online reporting systems more effective and trusted globally.

Limitations and future directions

While this study has provided valuable insights into the impact of reporting methods on the online crime reporting experience, several limitations should be acknowledged, along with suggestions for future directions. The study relied on hypothetical scenarios and vignettes to simulate reporting experiences. Although this approach allowed for controlled conditions, it will not have captured the complexities, emotions, and nuances of real-life crime reporting situations. Future research should aim to complement this approach with the analysis of real-world data or qualitative interviews with actual crime victims and witnesses to provide a more comprehensive understanding of the phenomena under investigation. The study also focused on only two crime types, namely graffiti and burglary. Expanding the range of crime types considered would be beneficial in order to ascertain whether the findings generalise across different offence categories. Variations in crime severity and characteristics may influence reporting behaviours and perceptions differently, and this might particularly be the case with crime types. In light of this, our conclusions in relation to H2 in particular should thus be interpreted cautiously.

The context in which the study was conducted may also be important. Cultural, regional, and jurisdictional factors may play a significant role in shaping individuals' perceptions and behaviours in reporting crimes. Within any one jurisdiction, different individuals and groups may have different needs and preferences in relation to dealing with Chatbots/AI, which might be based on factors as varied as age and prior individual or group experiences of police bias and prejudice. Future research should investigate whether the results described above hold true across diverse locations and populations and if they do not, why.

Conclusion

This study tested six hypotheses regarding online crime reporting preferences. It found that respondents generally preferred human operators over chatbots for reporting crimes, regardless of the type of crime, judging the human to human interaction to be fairer and more satisfactory (H1 and H2). That said, procedural justice was identified as a key factor affecting outcome fairness in both human and chatbot-operated conditions (H3), and respondents perceived human operators as providing clearer explanations compared to chatbot operators (H4). However, people did not judge decisions made by chatbots to be more neutral than decisions made by humans – in fact, the reverse (H5). Finally, outcome satisfaction varied depending on factors such as the type of operator and the type of crime reported, with active outcomes leading to higher satisfaction (H6).

Collectively, our findings emphasise the importance of a user-centred approach when developing and implementing online crime reporting systems. Ensuring procedural justice, clear communication, and human interaction when needed can lead to higher user satisfaction and trust in the online reporting process. If they wish to enhance public trust and popular legitimacy via the mechanisms heralded by 'channel choice', police organisations should continuously monitor and improve their online reporting systems based on user feedback to enhance public engagement and cooperation in reporting and preventing crime. Over-reliance on automated systems, were this to occur, might undermine this process.

Although this study has provided valuable insights, future research should consider incorporating real-world data and qualitative interviews, broadening the scope of crime types studied, exploring cross-cultural variations, and adopting mixed-methods approaches to better address the contextual and behavioural aspects of online crime reporting. Such efforts can ultimately lead to improved online reporting systems and enhanced public engagement with police.

Notes

1. In 1989, 94 Liverpool Football Club fans died in a crush at Hillsborough Stadium in Sheffield. A further three died subsequently from their injuries. At the time, police and press colluded to blame the fans for the deaths. Via a long process of enquiries, not completed until 2016, it subsequently emerged that errors by the police and stadium managers lead to the disaster.
2. Stephen Lawrence was a young black man murdered by racists in London in 1993. The police investigation into his death was subsequently found to have been riddled with errors, racism and corruption.
3. See for example the YouGov confidence tracker, available at: https://yougov.co.uk/topics/politics/trackers/how-much-confidence-brits-have-in-police-to-deal-with-crime?crossBreak=female
4. These chats are offered by way of a small icon often in the corner of a website, or a pop-up box, to visitors to a police website seeking information or to report issues.
5. For a $2 \times 2 \times 2$ factorial design, with a medium effect size (Cohen's d = 0.5), an alpha level of 0.05, and a power of 0.8, a sample size of approximately 34 participants per cell is necessary to detect both main effects and interactions with the desired power.
6. Participants were also presented with questions regarding their reporting preference as well as general questions about their levels of trust in the police, perceptions of police legitimacy, and their future intentions to engage or cooperate with the police; but we do not utilise these measures in the analysis presented.
7. For these interactions $p > .10$ in every case, except the graffiti, chatbot, model where the interaction between procedural justice and outcome was b = .26, $p = .052$). There is thus some indication that if the operator was a chatbot and the crime was less serious procedural justice was less important if the outcome was active, but given the *p*-value and other results noted caution is needed in interpreting this finding.

Data availability statement

All study materials will be uploaded to a secure OSF site: [link].

Disclosure statement

No potential conflict of interest was reported by the author(s).

Funding

This work was supported by the Economic and Social Research Institute [grant number ES/V00283X/1].

Notes on contributors

Ben Bradford (UCL) is a Professor of Global City Policing at the Department of Security and Crime Science, where he is the Director of the Centre for Global City Policing.

Arabella Kyprianides is a Research Fellow in Policing at UCL and an active member of both the Centre for Global City Policing (UCL) and the Keele Policing Academic Collaboration (KPAC).

Will Andrews is a Lecturer in Criminology & Social Science at Keele University.

Elizabeth Aston is the Director of the Scottish Institute for Policing Research (SIPR) and a Professor of Criminology at Edinburgh Napier University.

Estelle Clayton is a research fellow exploring technologically mediated policing at Edinburgh Napier University.

Megan O'Neill l is a Professor (Teaching and Research) in Human Geography at the University of Dundee.

Helen Wells is a Senior Lecturer in Criminology at Keele University.

References

Aston, E.V., *et al.*, 2023. Information sharing in community policing in Europe: building public confidence. *European journal of criminology*, 20 (4), 1349–1368.

Banks, J., 2019. A perceived moral agency scale: development and validation of a metric for humans and social machines. *Computers in human behavior*, 90, 363–371.

Bellaiche, L., *et al.*, 2023. Humans versus AI: whether and why we prefer human-created compared to AI-created artwork. *Cognitive research: principles and implications*, 8, 42.

Bradford, B. and Jackson, J., 2010. Trust and confidence in the police: A conceptual review. *SSRN*. doi:10.2139/ssrn.1684508.

Burton, J.W., Stein, M.-K., and Jensen, T.B., 2020. A systematic review of algorithm aversion in augmented decision making. *Journal of behavioral decision making*, 33, 220–239.

Chan, A., Bradford, B., and Stott, C., 2023. A systematic review and meta-analysis of procedural justice and legitimacy in policing: the effect of social identity and social contexts. *Journal of experimental criminology*, 17 (3), 1–58.

Dietvorst, B., Simmons, J., and Massey, C., 2015. Algorithm aversion: people erroneously avoid algorithms after seeing them err. *Journal of experimental psychology: general*, 144 (1), 114–126.

Dietvorst, B.J., Simmons, J.P., and Massey, C., 2018. Overcoming algorithm aversion: people will use imperfect algorithms if they can (even slightly) modify them. *Management science*, 64, 1155–1170.

Filiz, I., *et al.*, 2023. The extent of algorithm aversion in decision-making situations with varying gravity. *PLos one*, 18 (2), e0278751.

Glikson, E. and Woolley, A.W., 2020. Human trust in artificial intelligence: review of empirical research. *Academy of management annals*, 14 (2), 627–660.

Hermstrüwer, Y. and Langenbach, P.. 2023. *Fair governance with humans and machines.* MPI collective goods discussion paper, no. 2022/4, Psychology, Public Policy, and Law (forthcoming), Available at SSRN. doi:10.2139/ssrn.4118650.

Himma, K.E., 2009. Artificial agency, consciousness, and the criteria for moral agency: what properties must an artificial agent have to be a moral agent? *Ethics and information technology*, 11, 19–29.

HMICFRS, 2017. PEEL: Police legitimacy 2017: a national overview. Available from: https://assets-hmicfrs.justiceinspectorates.gov.uk/uploads/peel-police-legitimacy-2017-1.pdf [Accessed 4 Oct 2023].

HMICFRS, 2020. A call for help Police contact management through call handling and control rooms in 2018/19. Available from: https://www.nottinghamshire.pcc.police.uk/Document-Library/Public-Information/HMIC-Reports/2020/HMICFRS-Report-A-call-for-help.pdf [Accessed 4 Oct 2023].

Hobson, Z., *et al.*, 2021. Artificial fairness? Trust in algorithmic police decision-making. *Journal of experimental criminology*, 19, 165–189.

Jackson, J. and Bradford, B., 2019. Blurring the distinction between empirical and normative legitimacy? A methodological commentary on 'police legitimacy and citizen cooperation in China'. *Asian journal of criminology*, 14, 265–289.

Jussupow, E., Benbasat, I., and Heinzl, A., 2020. Why are we averse towards algorithms? A comprehensive literature review on algorithm aversion. *28th European Conference on Information Systems (ECIS)*.

Kaplan, A.D., *et al.*, 2023. Trust in artificial intelligence: meta-analytic findings. *Human factors: the journal of the human factors and ergonomics society*, 65 (2), 337–359.

Kleinberg, J., *et al.*, 2018. Advances in big data research in economics. *AEA papers and proceedings*, 108, 22–27.

Lee, M., 2018. Understanding perception of algorithmic decisions: fairness, trust, and emotion in response to algorithmic management. *Big data & society*, 5 (1), 1–16.

Lind, E.A., 2001. Fairness heuristic theory: justice judgments as pivotal cognitions in organizational relations. *In:* J. Greenberg and R. Cropanzano, eds. *Advances in organizational justice*. Stanford, CA: Stanford University Press, 56–88.

Longoni, C., Bonezzi, A., and Morewedge, C., 2019. Resistance to medical artificial intelligence. *Journal of consumer psychology*, 46 (4), 629–650.

Mahmud, H., *et al.*, 2022. What influences algorithmic decision-making? A systematic literature review on algorithm aversion. *Technological forecasting and social change*, 175, 121390.

McGuire, M.R., 2021. The laughing policebot: automation and the end of policing. *Policing and society*, 31 (1), 20–36.

Merola, L.M., Lum, C., and Murphy, R.P., 2019. The impact of license plate recognition technology (LPR) on trust in law enforcement: a survey-experiment. *Journal of experimental criminology*, 15, 55–66.

Miller, S.M. and Keiser, L.R., 2021. Representative bureaucracy and attitudes toward automated decision making. *Journal of public administration research and theory*, 31 (1), 150–165.

Mozafari, N., Hammerschmidt, M., and Weiger, W., 2021. *That's so embarrassing! When not to design for social presence in human-chatbot interactions*. Proceedings of the international conference on information systems. Available from SSRN: https://ssrn.com/abstract=4305288

Nagtegaal, R., 2021. The impact of using algorithms for managerial decisions on public employees' procedural justice. *Government information quarterly*, 38 (1), Article 101536.

Nass, C. and Moon, Y., 2000. Machines and mindlessness: social responses to computers. *Journal of social issues*, 56 (1), 81–103.

NPCC, n.d. Digital policing. Available from: https://www.npcc.police.uk/NPCCBusinessAreas/ReformandTransformation/Digitalpolicing.aspx [Accessed 4 Oct 2023].

ONS, 2022. Contact and trust with the police, year ending March 2016 to March 2020. Available from: https://www.ons.gov.uk/peoplepopulationandcommunity/crimeandjustice/adhocs/15340contactandtrustwiththepoliceyearendingmarch2016tomarch2020 [Accessed 4 Oct 2023].

Pemberton, A. and Mulder, E., 2023. Bringing injustice back in: Secondary victimization as epistemic injustice. *Criminology & Criminal justice*. doi:10.1177/17488958231181345.

Pesch, U., 2020. Making sense of the self: an integrative framework for moral agency. *Journal for the theory of social behaviour*, 50 (1), 119–130.

Rabinovich-Einy, O. and Katsh, E., 2014. Digital justice: reshaping boundaries in an online dispute resolution environment. *International journal of online dispute resolution*, 1 (1), 5–36.

Saulnier, A., 2023. Surveilled subjects and technologically mediated law enforcement: reflecting on relational concerns. *In:* M. Kwet, ed. *The Cambridge handbook of race and surveillance*. London: Cambridge University Press, 304–322.

Saulnier, A. and Sivasubramaniam, D., 2021. Procedural justice concerns and technologically mediated interactions with legal authorities. *Surveillance & society*, 19 (3), 317–337.

Schoenebeck, S., *et al.*, 2021. Youth trust in social media companies and expectations of justice: accountability and repair after online harassment. *Proceedings of the ACM on human-computer interaction*, 5 (CSCW1Article No. 2), 1–18.

Skogan, W.G., 1990. Police and public in England and Wales: a British crime survey report. Available from: https://www.ojp.gov/ncjrs/virtual-library/abstracts/police-and-public-england-and-wales-british-crime-survey-report [Accessed 4 Oct 2023].

Sunshine, J. and Tyler, T., 2003. Moral solidarity, identification with the community, and the importance of procedural justice: The police as prototypical representatives of a group's moral values. *Social psychology quarterly*, 66, 153–165.

Thibaut, J.W. and Walker, L., 1975. *Procedural justice: A psychological analysis*. London: John Wiley & Sons.

Tyler, T., 2003. Procedural justice, legitimacy, and the effective rule of law. *Crime and justice*, 30, 283–357.

Tyler, T., 2006. *Why people obey the Law*. 2nd ed. Princeton, NJ: Princeton University Press.

Tyler, T., *et al.*, 2021. Social media governance: can social media companies motivate voluntary rule following behavior among their users? *Journal of experimental criminology*, 17 (1), 109–127.

Tyler, T.R. and Blader, S.L., 2003. The group engagement model: procedural justice, social identity, and cooperative behavior. *Personality and social psychology review*, 7 (4), 349–361.

Tyler, T.R. and Huo, Y.J., 2002. *Trust in the law: Encouraging public cooperation with the police and courts*. Russell Sage Foundation.

Walker, L., Lind, E., and Thibaut, J.W., 1979. The relation between procedural and distributive justice. *Virginia law review*, 65, 1401–1420.

Wells, H., 2008. The techno-fix versus the fair cop: procedural (in)justice and automated speed limit enforcement. *British journal of criminology*, 48 (6), 798–817.

Wells, H.M., *et al.*, 2023. 'Channel shift': technologically mediated policing and procedural justice. *International journal of police science & management*, 25 (1), 42–52. doi:10.1177/14613557221132962.

Yuan, S. and Lou, C., 2020. How social media influencers foster relationships with followers: the roles of source credibility and fairness in parasocial relationship and product interest. *Journal of interactive advertising*, 20 (2), 133–147.