Can Artificial Intelligence Cut the Gordian Knot? A Factorial Survey Analysis of AI Usage and the Legitimacy of Bureaucratic Decisions

Star, Daniel van de

Can Artificial Intelligence Cut the Gordian Knot?

A Factorial Survey Analysis of AI Usage and the Legitimacy of Bureaucratic Decisions

Author: Daniel van de Star

Programme: MSc Public Administration

Specialization: Public Management and Leadership

Primary Reader: Dr. Hsini Huang

Secondary Reader: Dr. Johan Christensen

Submission Date: January 5, 2024

Abstract

What is the effect of AI technology usage by bureaucrats under problematic conditions on the perceived legitimacy of bureaucratic decisions? Scholars argue that AI usage potentially exacerbates the negative consequences of misused bureaucratic discretion. Others suggest that AI usage might curtail bureaucratic discretion and increase outcomes of equity and efficiency. Existing empirical research demonstrates no significant difference between the perceived legitimacy of AI-assisted and human decision-making. This study aims to determine the effect of AI usage on the perceived legitimacy of bureaucratic decisions made under problematic principal-agent dilemma conditions. This effect is assessed across 96 survey respondents from the University of Leiden and the University of Amsterdam using experimental factorial survey analysis. The results of this research indicate that AI usage in decision-making has a significant positive effect on perceived legitimacy ($p < 0.001$). The main implication of this research is that AI usage can plausibly alleviate the impact of consequential bureaucratic decisions on perceived legitimacy by obscuring bureaucratic discretion. A second implication is that AI usage in bureaucratic decision-making exerts a notable effect on the perceived efficiency of bureaucratic decisions.

Keywords: Perceived legitimacy, bureaucratic discretion, street-level bureaucrat, principal-agent problem, ai technology

Contents

1.1. Problem Statement

The purpose of this study is to analyze the effect of AI usage on the perceived legitimacy of bureaucratic decisions made under problematic conditions. The research question of this study is: What is the effect of AI technology usage by bureaucrats under problematic conditions on the perceived legitimacy of bureaucratic decisions? Perceptions of legitimacy form an essential component of government decision-making. Problems of legitimacy also present a salient contemporary challenge to public administration. The OECD 2021 Survey on Drivers of Trust in Public Institutions finds that the public perception of government integrity is an issue in many countries (2021). Illegitimacy is a social problem “incompatible with the values of a significant number of people who agree that action is needed to alter the situation” (Rubington & Weinberg, 2011, p. 3). Scholars have extensively highlighted the consequences of government illegitimacy (Nye et al., 1997; Hunter & Bowman, 1996; Clark & Lee, 2001). If citizens do not believe the government to be legitimate, voting outcomes decrease, and participatory and successful governance declines (Schoon, 2022). Therefore, this topic is exceedingly relevant to public administrators

Bureaucrats are necessary for a functioning government, contributing to public administration in many beneficial ways. However, bureaucrats also have the potential to exacerbate perceived illegitimacy. Pathological behaviors of bureaucrats can directly impact how citizens perceive the institutions to which these bureaucrats belong. One example is the behavior of police officers in the death of George Floyd, which had a significant impact on perceptions of the police (Brantingham et al., 2022). Another notable example of such a phenomenon is the publicized undermining of the Trump administration by White House officials, which subsequently reinforced conspiracy theories about the deep state and discredited the legitimacy of government bureaucracy (Taylor, 2018; Rothkopf, 2022; Roig-Franzia, 2020). These examples demonstrate the assumption that bureaucrats are empowered to make problematic decisions. The inherent function of bureaucrats in the government as executors and enforcers of policy grants these agents much discretion and the potential to misuse this discretion.

Many scholars conceptualize government institutions as actors, but bureaucrats ultimately execute policy, making bureaucratic discretion integral to government decision-making (March & Olsen, 1989; Linders & Peters, 1990). Political philosophers have historically emphasized the importance of bureaucrats being just and procedural in enforcing policy (Shafritz & Hyde, 1992). Accordingly, "the civil servant… has a responsibility to be able to exercise discretion" so long as they follow institutional aims and are "guided by the public interest" (Plant, 2011, p. 470; Shafritz & Hyde, 1992, p. 77). Under certain conditions, such as those implicit in the principal-agent dilemma, bureaucrats pursue conflicting aims, shirk responsibility, or work inefficiently. Bureaucratic discretion under these conditions can inadvertently undermine institutional aims and contravene the public interest. More importantly, these situations can cause bureaucrats to make decisions that can impair the perceived legitimacy of the government (Schoon, 2022). Addressing the problem of bureaucratic discretion under problematic conditions is a priority for public administrators to understand how to prevent perceived illegitimacy.

The use of AI technology presents a potential solution to mitigating the problematic behaviors of bureaucrats. AI technology is developing at an unprecedented rate and with remarkable consequences (Moore & Intel.Com, 1965). It possesses the computing power necessary to solve problems more efficiently and fairly than a human could, overcoming the pathologies of misused bureaucratic discretion (Bullock et al., 2022; Agarwal, 2018). For these reasons, government officials in the public sector predict AI technology usage to increase government effectiveness and citizen satisfaction (Mergel et al., 2019). Bureaucratic decision-making under certain conditions threatens government legitimacy, and AI technology provides a potential solution to this problem. The question is whether there are unforeseen consequences of utilizing AI technology in government decision-making.

The development of AI technology will notably impact various elements of public administration and governance (Zerilli et al., 2019; Edwards & Veale, 2017; Cobbe, 2019; Concerned Researchers, 2021; Bannister & Connolly, 2014). AI usage in governance can have notable negative consequences. (Bullock et al., 2022; Wenger & Wilkins, 2009; Compton et al., 2022). For instance, researchers predict that the development of AI technology will lead to drastic unemployment (McGaughey, 2021). It is not unreasonable to assume that this technology will similarly affect the private and public sectors. The "rapid and unexpected change" associated with AI development will also burden the operational capacity of the government (Sharitz & Hyde 1998, p. 284; Ansoff, 1979). Government organizations notoriously struggle to adapt to disruptive changes, making transformative technology like AI likely to do the same (Janssen & Van Der Voort, 2016). AI technology usage will particularly impact organizational decision-making (Koteen, 1997). Since "decision-making is at the heart of administration," it is prudent to understand how AI usage will affect government decision-making. (Simon, 2013, p. x1). Thus, studying the effect of AI usage on bureaucratic decision-making presents a forward-looking research subject.

Medaglia et al. (2021) emphasize the need for research on AI and governance in public administration, as present research on the topic is limited. De Fine Licht and De Fine Licht (2020) study the role of transparency on the perceived legitimacy of AI-assisted government decision-making. Martin and Waldman (2022) analyze perceptions of the legitimacy of algorithmic decisions made by firms. Other researchers demonstrate the differences in perceived red tape and trustworthiness between AI and human-led government decision-making (Ingrams et al., 2021). Some research emphasizes potential worries about AI governance and political legitimacy (Erman and Furendal, 2022; Beckman et al., 2022). In general, many scholars express concerns about bias in automation (Agarwal, 2018; Alon‐Barkat & Busuioc, 2022). However, few empirical or evidence-based studies explore the effect of AI technology usage on perceived legitimacy, especially under problematic conditions (Charles et al., 2022).

1.2. Academic Relevance

This study explains how AI usage impacts the perceived legitimacy of bureaucratic decisions under problematic situations, such as those inherent in the principal-agent dilemma, as opposed to ordinary conditions. The principal-agent dilemma is a conventional problem in public administration whose consequent pathologies have been studied extensively by scholars. Considering the additional use of emergent AI technology in bureaucratic decision-making, a layer of ambiguity and complexity is involved. Will the inclusion of AI technology in bureaucratic decision-making worsen the pathologies of bureaucratic discretion under problematic conditions? The competing theories of AI technology suggest that the technology is both biased *and* neutralizing. Thus, the impact of using AI technology is not entirely known. This new development presents a compelling topic of research.

Existing research that comes closest to this topic is that by Starke and Lünich (2020). They measure the impact of AI usage in government decision-making on the perceived legitimacy of policy outcomes related to European Union (EU) budget distributions made at the organizational level. Using a factorial survey analysis, they test fully autonomous AI decision-making, AI-assisted decision-making, and human decision-making to determine AI technology's effect on various measures of legitimacy. They find no significant difference in perceptions of policy outcomes resulting from AI-assisted decision-making compared to human decision-making. Although Starke and Lünich (2020) present relevant insights into the subject at hand, they fail to realistically assess the phenomenon of AI usage in bureaucratic decision-making in several ways.

First and foremost, Starke and Lünich (2020) study decisions made at the "EU level" by "EU policy makers" (pp. 2, 14). Citizens' perceptions are at the forefront of the research topic, but Starke and Lünich's (2020) research features a supranational governmental organization with demonstrably less public participation than national governments have (European Parliament, 2019; International IDEA, 2024). Furthermore, citizens often do not directly interact with government organizations but interact with street-level bureaucrats who represent those organizations (Lipsky, 1980). If citizens' perceptions are at the forefront of the research, then vignettes should display relatable street-level bureaucrats and not foreign and distanced EU figures. Moreover, the decisions represented in the vignettes ought to be salient to citizens to be relevant. An organizational decision to marginally decrease the EU budget is feasibly less impactful to citizens than, for instance, the decision by a tax auditor to issue them a fine for tax fraud.

Finally, the research fails to consider the ubiquity of problematic conditions that can impact bureaucratic discretion. The potential benefit of AI usage in decision-making lies in its ability to reduce inefficiency and bias. By presenting a non-problematic procedural and agreed-upon organizational decision, Starke and Lünich (2020) miss the opportunity to truly measure the potential contributions of AI technology to governmental decision-making. Medical research, for example, studies curative medicine on sick people, not healthy people. If public administrators wish to understand the potential impact of AI technology as a solution to governance-related problems, research should test the technology in problematic situations.

In contrast, this study provides value to the field of public administration by building on Starke and Lünich's (2020) research and exploring the effect of AI usage on the perceived legitimacy of decisions made by street-level bureaucrats in the context of realistic governance-related conditions where bias and inefficiency *are* present. Furthermore, this study focuses on the salient decisions of familiar street-level bureaucrats in relatable scenarios, as opposed to studying the decisions of high-level EU officials at the organizational level. This study aims to recreate Starke and Lünich's (2020) research more realistically by introducing assumptions about individual bureaucratic discretion and the ubiquity of conventional principal-agent dilemmas in government decision-making. The data collection method employed by this study is a digital Qualtrics survey administered to participants in group chats comprising university students at UL and UVA. A factorial vignette research design is used to isolate the effect of AI usage by bureaucrats on perceived legitimacy. Participants are randomly assigned to a control vignette or treatment vignette. The data analysis is performed using Stata 18. The primary statistical analyses this study uses to estimate the effect of AI usage in bureaucratic decision-making on perceived legitimacy are ordered logistic models and Spearman's rank correlation.

1.3. Roadmap

In the introduction, this study presents the problem statement and addresses this study's contribution to the field of public administration. Next, the theoretical framework outlines the theory underpinning this study and states the hypotheses this study aims to answer. In the methodology section, this study's research design is explained, sampling and participants are discussed, the primary dependent and independent variables are operationalized, methods of data analysis are highlighted, the validity and reliability of the study are explored, and ethics are considered. In the analysis section, the descriptive statistics of the study are provided and statistical analyses are carried out to test the hypotheses. In the discussion section, the interpretation of the data is provided and potential implications are reviewed. Finally, a conclusion is drawn, potential future research is suggested, and policy recommendations are made.

2. Theoretical Framework

2.1. Legitimacy

2.1.1. Defining Legitimacy

According to Weber's definition in 1919, legitimacy can be defined as the right to rule. Weber identified three types of legitimacy: traditional, legal, and charismatic legitimacy (Weber, 1985). Traditional legitimacy is also known as normative legitimacy, which is based on normative beliefs that justify whether a party or leader should be allowed to rule (Levitov, 2016; Risse‐Kappen & Stollenwerk, 2018). Legal legitimacy, on the other hand, describes the right to rule under current legal standards (Roth, 2000). Charismatic legitimacy refers to the justification of the rule based on a leader's achievements (Weber, 1985).

In this study, legitimacy is defined according to Suchman's (1995) definition as "a generalized perception or assumption that the actions of an entity are desirable, proper, or appropriate within some socially constructed system of norms, values, beliefs, and definitions" (p. 574). Therefore, this study defines legitimacy in terms most closely related to Weber's concept of traditional or normative legitimacy (Gilley, 2009).

Legitimacy is a complex and controversial concept that involves various understandings that can contradict and overlap (Stillman, 1974). It is important to keep in mind that these understandings are not interchangeable. Additionally, the legitimacy of a state or institution can come from many sources beyond citizens' judgment of its institutions (Risse‐Kappen & Stollenwerk, 2018). For example, democratic stability, which is closely connected to perceived legitimacy, can be influenced by international oil prices (Brückner et al., 2012). This means that there are several potential factors that can affect the measurement of legitimacy and make understanding it more complicated.

Different institutions within government can also be perceived differently in terms of legitimacy. Thus, just because one institution is considered illegitimate does not mean that the entirety of the government is. Perceptions of legitimacy can also change over time as people obtain new information (Risse‐Kappen & Stollenwerk, 2018). Cultural and socioeconomic factors can influence perceptions of legitimacy (Brandt et al., 2020). Thus, legitimacy is not perceived the same by every citizen. A prime example of this is the police, which minorities in the United States disproportionately believe to be racially biased (Morin & Stepler, 2013; Gau & Brunson, 2009; Brunson, 2007; Wortley et al., 1997). Legitimacy is a complex notion and perceptions of legitimacy are not easily ascertained. However, perceptions of legitimacy are essential to effective governance.

2.1.2. The Importance of Legitimacy

For democratic institutions to be successful and governance to be effective, public opinion must be favorable and consider institutions legitimate (Dye, 1998; Risse‐Kappen & Stollenwerk, 2018; Samuelson, 1995). Scholars have emphasized the role of legitimacy in understanding political institutions (DiMaggio & Powell, 1983). Legitimacy is used to justify state violence (Jackson, 2018), uphold norms and practices (Johnson et al., 2006), increase organizational effectiveness (Lipsky, 2011), and reduce the need for collective mobilization (Suchman, 1995). In addition, perceptions of legitimacy play a crucial role in the functioning of specific government institutions such as the police, tax agencies, and health departments. Perceived illegitimacy can decrease cooperation with the police, the reporting of crimes, and adherence to the law (Tyler, 2009, 2011). If citizens perceive tax authorities as more legitimate they are more compliant to tax laws (Hartl et al., 2015). During the COVID-19 pandemic, those who had lower trust in the government, which is closely related to perceived legitimacy, were less likely to follow measures designed to mitigate the spread of the virus. (Georgieva et al., 2021). Therefore, perceived legitimacy is essential to maintaining effective government operations across a wide range of institutions.

2.1.3. Mediators of Legitimacy

Since legitimacy is a concept downstream from many other normative judgments, assessing legitimacy relies on many mediating variables. For instance, if a government is corrupt and appropriates funds or discriminates against its citizens, this can be reasonably assumed to be detrimental to its perceived legitimacy. Similarly, if citizens distrust the government, this might impact the government's perceived legitimacy. Trust, for instance, is especially relevant to concerns about data integrity, misinformation, and technological complexity (Gerlich, 2023). There is an evident link between trust and legitimacy (Jackson & Gau, 2016; Kappmeier & Fahey, 2022). Other mediators of perceived legitimacy include efficiency and fairness. Efficiency contributes to upholding institutional legitimacy (Jeong & Kim, 2019). Fairness can also shape perceptions of legitimacy and vice-versa, especially in legal settings (Melamed, 2012; Franck, 1998; Farrar, 2022).

Furthermore, considering efficiency and fairness as factors that influence the perceived legitimacy of an institution or government is conducive to the definition of legitimacy used in this study. Stone (2011) highlights efficiency and equity as core public policy values and goals that policymakers strive to achieve.[1] Since the goals of efficiency and fairness meet Suchman's (1995) definition, these measures feasibly contribute to perceived legitimacy.[2] Moreover, equity and efficiency are particularly relevant to this study as mediators of legitimacy because they correspond to theories on fairness, bias, equity, and efficiency that are prominent in the literature that discusses the potential consequences of using AI technology.

2.2. The Pathologies of Bureaucratic Discretion

2.2.1. Street-Level Bureaucrats

Dahl (1947) remarks that “major problems revolve around human beings and that human beings consequently cannot be ignored in the study of public organizations" (Denhardt, 1999, p.76). This assumption is especially true as it relates to implementing policy and citizen-state interactions. Street-level bureaucrats are front-line government employees who interact face-to-face with citizens and possess distinct and subjective authority in policy enforcement. Examples include police officers, tax auditors, and health inspectors. In contrast to assumptions about procedural or rule-based work in the bureaucracy, street-level bureaucrats possess “substantial discretion in the execution of their work” (Lipsky, 1980, p.3). Street-level bureaucrats selectively enforce policy depending on what they deem appropriate or correct (Nordesjö et al., 2020). In other words, they possess the authority to discriminate in their decisions. Discrimination is not always a negative phenomenon. In some cases, bureaucrats use their discretion to bend the rules in favor of individuals who require it (Zacka, 2017). However,

[1] Fairness is a value closely-related to equity and is a synonym of the term (Thesaurus, 2023).
[2] According to Suchman’s (1995) definition of legitimacy, something is legitimate if it is “desirable, proper, or appropriate” (p. 574).

discrimination is subjective, meaning that bureaucrats are prone to potentially misusing their discretion.

The prevalence of discriminatory behavior is often determined by context (Christensen et al., 2012). For instance, in stressful environments, bureaucrats are more likely to discriminate (Vedung, 2015; Tummers et al., 2015). In doing so, bureaucrats often use stereotypes to ease their decision-making process (Lipsky, 1980; Maynard-Moody & Musheno, 2003; Prottas, 1979; Harrits & Møller, 2014; Epp et al., 2014; Dubois, 2010). According to Lipsky (1980), bureaucrats do not necessarily discriminate due to prejudiced beliefs. Instead, he argues that discrimination is a coping mechanism to deal with stress. In doing so, bureaucrats service clients differently to “maximize personal or agency resources” (p. 107). Mennerick (1974) expands on this notion and argues that bureaucrats make decisions to eliminate or reduce conflict in the workplace. Frontline bureaucrats even implement policy differently in response to pressure from political actors and the media to avoid blame (Hinterleitner & Wittwer, 2022). Bureaucrats possess notable discretion, use their power to discriminate, and make their jobs easier by selectively enforcing policy, and thus are susceptible to making mistakes.

2.2.2. The Principal-Agent Dilemma

Many scholars have challenged the notion of an inefficient bureaucracy, but in certain conditions, bureaucrats who possess considerable discretion make poor decisions (Rainey & Steinbauer, 1999; Pierre & Peters, 2017; Svara, 2001). As aforementioned, context determines bureaucratic behavior. The principal-agent dilemma explores the problematic consequences of malfunctioning bureaucracy. Principal-agent theory models the relationship between actors as consisting of one or more superior principals and one or more subordinate agents (Jensen & Meckling, 1976). A classic configuration of the model comprises an elected official as a principal and a bureaucrat as an agent. Another configuration is that of a citizen as the principal and a bureaucrat as the agent. The principal-agent dilemma builds on agency theory, which operates under the assumption that unless work is monitored or rewarded, “employees will put in as little effort as they can get away with” (Organization For Economic Co-Operation And Development, 2005, p.32; Pratt & Zeckhauser, 1985; Arrow, 1984). Under the assumptions of agency theory, unless an individual expects a reward proportional to their effort, they will not be motivated to complete a task unless motivated otherwise (Lawler, 1971; Vroom, 1964). Principal-agent theory similarly assumes that actors are utility-maximizing and that there is a low level of trust between principals and agents (Jensen & Meckling, 1976). Principal-agent theory suggests that street-level bureaucrats are more likely to malfunction under certain conditions. Three such conditions are information asymmetry, goal ambiguity, and moral hazard.

2.2.2.1. Information Asymmetry.

Information asymmetry describes an environment with a remarkable difference in the knowledge and expertise between the principal and agent. The term is closely related to the economic theory of adverse selection, which describes a situation in which a seller of a good knows more than a buyer and uses this knowledge to their advantage (Akerlof, 1970). One reason that information asymmetry develops is because principals and agents occupy roles with different responsibilities and tasks. The fact that “bureaucracy…almost completely avoids public discussion of its techniques, [despite] public discussion of its policies," reflects the information asymmetry present in the citizen-government configuration of the principal-agent dilemma (Shafritz & Hyde, 1992, p. 102). This situation is problematic because bureaucrats can “monopolize expertise,” allowing them to define the terms of their relationship with their principal (Aberbach et al. 1981, p. 8; Crawford & Guasch, 1983). As Shafritz and Hyde (1992) describe it, “every trained technician…has a profound sense of omniscience and a great desire for complete independence in service of society. When employed by the government [they know] exactly what the people need better than they do themselves” (p. 86). Information asymmetry can enable biased and inefficient decision-making, which has the power to undermine perceptions of legitimacy.

2.2.2.2. Goal Ambiguity.

Goal ambiguity describes a situation in which an agent must fulfill multiple – sometimes conflicting – goals, leading to uncertainty. If goals are not adequately defined, an agent must interpret them, which leads to incorrect behaviors. Goal ambiguity is better understood when considering the need for a “unity of command” in organizational administration, the need for a single superior. Per Shafritz and Hyde (1992), “a workman subject to orders from several superiors will be confused, inefficient, and irresponsible” (p. 85). There is a demonstrable link between the quality of goal setting and performance (Latham & Locke, 2002). Weber emphasizes the importance of hierarchies in bureaucratic organizations, stressing the need for predictability of results (Denhardt, 1999). Therefore, if bureaucrats are not attuned to organizational goals, there can be discord in performance. Alternatively, if goals are displaced, a similar result can occur. Displacement describes the process when “an instrumental value becomes a terminal value (Shafritz & Hyde, 1992, p. 103). An example might be a police officer pursuing minor traffic stops to fulfill quotas instead of prioritizing violent crimes, which in turn conflicts with the general goal of the police to promote safety.

2.2.2.3. Moral Hazard.

Arrow (1984) defines moral hazard as a “hidden action” by an agent (p. 38). A classic example in economic literature is the example of insurance. If a person is medically insured, they might be more likely to engage in dangerous activities, knowing that if something goes amiss, the insurance company will pay for potential medical bills (Finkelstein, 2014). In this study, moral hazard describes a situation in which an agent makes a decision because they do not bear the risk of that decision. A bureaucrat might make a decision without considering all risks because they have limited accountability. After all, bureaucrats make decisions on behalf of a government institution and do not personally bear direct responsibility.

2.2.2.4. Principal-Agent Dilemma and Perceived Legitimacy.

Principal-agent problems notably result in consequences that undermine the legitimacy of the institutions to which the agent belongs. For instance, a bureaucrat could engage in discriminatory behaviors due to goal ambiguity or displacement. A practical example of this type of circumstance is stop-and-frisk policies. Stop-and-frisk allows police officers to conduct body searches on citizens without a warrant. These policies disproportionately affect minority citizens (Baker, 2010). When principals subsequently incentivize police officers to meet arrest quotas, they contradict the police's aim to contribute to the public good because they are inadvertently discriminating against citizens of color (Sparrow, 2015). Thus, the condition of goal ambiguity empowers poor decision-making by a bureaucrat, which leads to policy enforcement that perpetuates perceptions of the police as unfair and illegitimate. Subsequently, minority citizens are more wary of police officers, compliance decreases, and police officers pursue more arrest quotas by stopping minorities that they deem suspicious. It is not difficult to imagine how such a condition can get out of hand.

Such inadvertent discrimination due to a principal-agent problem can have detrimental consequences for the legitimacy of governmental institutions (National Institute of Justice, 2013). These consequences have been demonstrated in correctional officers in prisons (Sparks &

Bottoms, 1995), tax agents in the Internal Revenue Service (Organization For Economic Co-Operation And Development, 2004), officers in the United States Department of Safety (Mazerolle et al., 2013), among other bureaucratic institutions. When the agent's principal is a public manager, the principal-agent dilemma might manifest in operational inefficiencies. However, if the agent's principal is the citizen they serve, inefficient behavior can lead to significant lapses in the perceived legitimacy of government institutions.

2.3. Contemporary Uses of AI in Government

AI is a powerful tool with many applications in the public and private sectors. AI is notably adaptive and not constricted by pre-programmed code. For this reason, scholars have emphasized its potential for use in dynamic organizational contexts (Dwivedi et al., 2021; Sun & Medaglia, 2019). There has already been extensive investment in AI technology in the private sector (McKinsey, 2020). According to KMPG, 77% of government leaders want a "more aggressive approach" to AI adoption (Alva, 2021). Some consider AI usage in governance "inappropriate or far off" (Edwards & Veale, 2017, p. 45). In contrast, many advocate the potential benefits of implementing AI technology in public administration (Zerilli, 2021; Charles et al., 2022; Alva, 2021). Contemporary employment of AI by the bureaucracy highlights the potential advantages of using AI technology – efficiency, objectivity, and predictability (Bullock et al., 2022).

Contemporary AI technology is far from the sentient entity envisioned in science fiction, although its applications are numerous and varied. AI algorithms today are used principally as "decisional aides," assisting human decision-making (Alon-Barkat & Busuioc, 2022, p. 154). Other functions of AI include machine learning (Smola & Vishwanathan, 2008), neural networks (Priddy & Keller, 2005), case-based reasoning (Kolodner, 1992), language processing such as in

ChatGPT (Chowdhury, 2003), multi-agent systems (Wooldridge, 2009), machine reasoning (Bottou, 2014), computer vision (Klette, 2014), internet-of-things (Shafiq et al., 2022), autonomous vehicles such as Tesla (Thrun, 2010), the simplification of public policy (Rice & Zorn, 2019), facial recognition (Viola & Jones, 2004), and prediction (Helsby et al., 2018; Lee, 2018). The applications of AI are just about as expansive as the concerns about the rapid development and use of AI. Possible problems associated with the increased use of AI are the potential for misinformation, the lack of privacy, the presence of algorithmic bias, the lack of transparency, the weaponization of AI, and existential risks related to the development of sentient AI (Russell & Norvig, 2019; Cellan-Jones, 2014; Nicas, 2018; Savage, 2022).

2.4. How AI Technology Affects Bureaucratic Discretion

An implication of using AI technology in government is its potential effect on bureaucratic discretion (Bovens & Zouridis, 2002). Busch and Henriksen (2018) assert that AI can "influence or replace human judgment" (p. 4). Existing research emphasizes that when AI technology assumes discretionary authority in decision-making, the potential benefits include scalability, quality, and cost-effectiveness (Young et al., 2019).[3] The demonstrated effect of AI on bureaucratic decision-making is varied. On one hand, AI technology can empower bureaucrats to make biased or inefficient decisions. On the other hand, AI technology can replace human judgment in decision-making and reduce the chance of bias or inefficiency.

2.4.1. The Curtailment Thesis

If bureaucrats use AI technology in government decision-making, it might reduce the influence of human discretion. The curtailment thesis suggests that technology will co-opt bureaucratic discretion as digitalization increasingly controls bureaucratic decision-making

[3] Many terms describe this phenomenon, including "artificial discretion," "automated discretion," and "digital discretion" (Young et al., 2019, p. 301; Zouridis et al., 2019, p. 314; Busch & Henriksen, 2018, p. 4).

(Bovens & Zouridis, 2002; Buffat, 2013). If bureaucrats exercise less discretion, there is less chance for poor decision-making. McGregor (2006) argue that the ordinary bureaucrat "works as little as possible, lacks ambition…is self-centered, indifferent to organizational needs, [and] resistant to change" (Shafritz & Hyde, 1998, p. 217). Under these assumptions, AI technology can provide an opportunity to alleviate the behavioral sources of inefficiency in government decision-making (Alon‐Barkat & Busuioc, 2022).

Furthermore, if less of the bureaucratic decision-making process relies on subjective judgments by human bureaucrats and more on objective judgments grounded in big data, bureaucratic decisions could be even more efficient and fair. The counter-argument to this assumption is that, as aforementioned, AI technology is prone to biased programming, relies on flawed training databases, and is susceptible to the bias of its users (Agarwal, 2018). Despite drawing on a supposedly objective source, human interpretations of data and subsequent policy applications are still subjective in nature (Boyd & Crawford, 2012).

2.4.2. The Enablement Thesis

AI technology could also empower bureaucrats to exercise their discretion even more than usual. The enablement thesis argues that the additional action resources provided by automation can enable bureaucratic discretion (Buffat, 2013). At face value, this presents a threatening consequence of AI usage by bureaucrats – especially in conditions where bureaucrats are empowered to make poor decisions. For instance, in the case of the Child Benefits Scandal, if AI technology equips a bureaucrat with more resources, the bureaucrat could exercise more discriminatory policy decisions. Another potential negative consequence of using AI technology under the assumptions of the enablement thesis is that bureaucrats can justify biased police outcomes by appealing to arguments of objectivity.

There are also potential positive consequences of the enablement thesis. AI technology reduces uncertainty, as it can co-opt much of the stressful decision-making process of bureaucrats (Gajduschek, 2003). According to the assumptions made by Lipsky (1980), the bureaucrat is principally self-interested and motivated to maximize agency resources and minimize stress and conflict. A different way of understanding the enablement thesis is by conceptualizing the potential of AI technology to reduce the stress and workload of bureaucrats by providing more resources to the bureaucrats. If a government agent has more time and access to tools and resources to do their job, it is plausible that they are more equipped for their work and less stressed and overworked. According to Lipsky's (1980) theory of coping, which argues that much of discrimination and biased decision-making by bureaucrats is due to stress, it is feasible that AI usage by bureaucrats could minimize discrimination and bias by reducing the bureaucrats' experience of stress and uncertainty.

2.4.3. Obscuring Discretion

AI usage in government can potentially affect the perception of bureaucratic discretion by citizens. Research by De Boer and Raaphorst (2021) finds that automation, for instance, does two things: street-level bureaucrats become more accommodative and people's experience of bureaucratic discretion is reduced. Surprisingly, despite the use of discretion by bureaucrats, the experience of bureaucratic discretion by citizens is reduced. The results of De Boer and Raaphorst's (2021) research can be partly explained by the theory of negativity bias, which is the proclivity for people to attribute greater salience to negative experiences than positive experiences (Vaish et al., 2008). Thus, perhaps citizens are less likely to notice and experience bureaucratic discretion when bureaucratic discretion positively impacts them. However, another possible explanation for the findings of De Boer and Raaphorst (2021) is that using automation obscures bureaucratic discretion. If bureaucratic discretion is obscured, a bureaucrat could feel more comfortable executing their discretion. Moreover, bureaucrats might even feel comfortable contradicting the rules they are intended to enforce – e.g., being more accommodative to the needs of citizens.

This conclusion is consistent with Jorna and Wagenaar's (2007) report, which finds that technologies do not reduce bureaucratic discretion but merely obscure it. When a bureaucrat utilizes AI technology in decision-making, the bureaucrat shares their responsibility for the decision with the technology. This is consequential in several ways. First, if a bureaucrat makes a poor decision under problematic conditions, the mere fact that AI technology is involved in the decision-making process could potentially reduce the negative assessment of the decision. It is ultimately the bureaucrat, not the technology that makes the decision. However, since the bureaucrat's discretion is obscured, the perceived responsibility for the decision lies in the technology. Second, such a setup empowers the bureaucrat to shirk responsibility for their decision. In the example of the Dutch Child Benefits Scandal, for instance, it would be plausible for a tax agent to blame the discriminatory algorithm rather than take responsibility for the problematic decisions themselves. Third, if citizens have preconceived prejudices against bureaucratic decision-making, then obscuring bureaucratic discretion could ameliorate potential problems associated with those prejudices. In the case of minority citizens in the United States, for instance, removing the perceived source of bias – police officers – from decision-making in law enforcement could benefit perceptions of legitimacy and fairness. Thus, employing AI technology in government decision-making provides a potential buffer against negative judgment by citizens, as perceived responsibility is seemingly diverted away from government agents and towards the technology they employ. In practice, this might result in shirking behaviors by bureaucrats. However, AI usage in government decision-making also provides a safeguard against negative perceptions.

2.5. Is AI Technology Neutral or Biased?

There is a general assumption that AI technology is neutral. This assumption is not necessarily correct. The implicit bias of programmers, the complexity of AI, and the training databases used by the technology make it prone to bias. Agarwal (2018) argues that the biases of those who program AI technology manifest in the technology. Furthermore, users and programmers cannot always identify bias because the machine-learning algorithms and code that govern AI are often highly complex (Agarwal, 2018).

Furthermore, AI functioning depends on training data, which informs its decisions and logic. Databases that rely on the Internet to obtain data reflect the biases of Internet users. The internet is not by any means a neutral space. Access to the Internet consolidates in wealthier countries with the infrastructure required for such technology. Therefore, users who consume and produce content on the Internet tend to share similar demographic characteristics (Ritchie et al., 2023). The homogeneity of internet users can influence the objectivity of training databases. One example of such a biased training database is the “Common Crawl,” which archives internet data monthly for open-source use (*Common Crawl,* 2023). Google Images relies on this database and frequently reflects biased results. In some of its worst results, when a user searches, gorilla, monkey, or ape, Google images will show images of black people (Alexander, 2018; Hern, 2018). If the AI technology used in government relies on such databases, the potential consequences could be dire.

Moreover, the rules that govern AI technology aim to emulate a logic of neutrality, emphasizing objectivity and statistical reasoning. Boyd and Crawford (2012) argue that, in reality, the line between objectivity and subjectivity is often ambiguous. When people use a tool or interpret data, subjective human judgment biases the results. A hammer is morally ambiguous, but a criminal can use it to commit a crime. Alternatively, a person can use a hammer to build a home. Even though analytical applications such as AI algorithms are touted by experts as objective, employing such algorithms can often perpetuate discrimination and unequal outcomes (Ferguson, 2017). Relying solely on statistical means can lead to statistical discrimination (Phelps, 1972). This type of discrimination comes from pattern recognition inherent in AI models, which often does not capture the more quantitatively intangible elements that play a role in a particular phenomenon. For example, analyzing crime statistics without context could lend itself to the conclusion that black individuals are predisposed to crime (Federal Bureau of Investigation, 2019). Whilst this is not the case, one can imagine the potential policy measures that a public administrator could take as a result of assuming that the AI technology they employ is wholly objective.

The concept of automation bias is related to the biased programming of AI technology. Automation bias describes the “overreliance on algorithmic advice even in the face of warning signals from other sources” (Alon‐Barkat & Busuioc, 2022, p.153). Research has found a tendency for automation bias in published research papers across many fields of study (Goddard et al., 2012). Contrary to these findings, other research claims that automation bias is not present in the decision-making of civil servants in the Netherlands (Alon‐Barkat & Busuioc, 2022). However, as recently as 2020, the Dutch Data Protection Authority was investigating the Dutch Tax Authorities for their involvement in the Child Benefits Scandal. In this scandal, the Dutch Tax Authorities wrongfully denied childcare benefits to thousands of people. The scandal involved the overreliance on an algorithm that determined whether applicants for child benefits were considered high-risk and likely to commit fraud. Investigators deemed the algorithm discriminatory because it took applicants' country of origin and dual-citizenship status into account in assigning applications to the high-risk category and denying them certain benefits. The Dutch Data Protection Authority stated that the algorithm was "designed and used" in a discriminatory manner (Autoriteit Persoonsgegevens, 2020). Despite a series of complaints about the denial of benefits, the Tax Agency continued to employ this algorithm anyway. This scandal provides a prime example of automation bias and how AI technology can reflect the biases of its programmers. One significant consequence of automation bias includes discrimination, as demonstrated in the Child Benefits Scandal (Noble, 2018).

Alternatively, evidence suggests that AI usage improves outcomes of fairness in practice. Algorithmic automation by AI can have a notable effect on reducing errors in government decision-making. Compton et al. (2022) found that automated state-client interactions significantly reduce discrimination arising from administrative errors in social welfare policies. Moreover, research by Wenger and Wilkins (2009) found that women receiving unemployment insurance increased without detriment to male applicants. This research suggests that if government policies are already biased against certain groups, reliance on automated AI algorithms can remove bias in decision-making. Ultimately, AI technology has the potential to both advance equity or discrimination – to undermine or uphold legitimacy – depending on its use and the context in which it is used (Brundage, 2018). Although these studies provide valuable insights into the potentially positive consequences of using AI technology, this study aims to answer how AI usage in bureaucratic decision-making will affect perceptions. Will citizens perceive AI technology usage in governance as a tool that empowers illegitimate decision-making or as an assurance of neutral competence that can alleviate the problems associated with bureaucratic discretion?

2.6. The Effect of AI Usage on Perceived Legitimacy

Existing research by Starke and Lünich (2020) finds that unsupervised use of AI technology is detrimental to perceived legitimacy. The necessity for a human operator is also found in other studies (Fritsch et al., 2022). For example, autonomous AI vehicles are skilled at avoiding collisions, whereas humans demonstrate the opposite effect (Scanlon et al., 2021). Despite this, research by Hidalgo et al. (2021) finds that AI failure is judged worse than human failure (Hidalgo et al., 2021). Thus, it would make sense that fully autonomous AI decision-making decreases perceptions of legitimacy. However, Starke and Lünich (2020) also find people do not judge AI-assisted decision-making differently from human decision-making. Several confounding factors in their study would contradict applying their findings to this study. For one, the research is performed at the EU level and addresses only EU budgetary policy. In contrast, this study focuses on street-level bureaucrats who make decisions about more relatable policies.

This study distinguishes how AI usage in bureaucratic decision-making affects actual legitimacy compared to perceived legitimacy. From a functional standpoint, AI technology can increase legitimacy by "identifying pressing societal issues, forecasting potential policy outcomes, and evaluating policy effectiveness" (Starke & Lünich, 2020, p. 1). AI usage also increases fairness and equity in policy outcomes (Compton et al., 2022; Wenger & Wilkins, 2009). AI usage also demonstrates the potential for increased efficiency in government decision-making (Bullock et al., 2022; Alon‑Barkat & Busuioc, 2022). But what effect does the use of AI usage exert on perceived legitimacy and measures of legitimacy such as efficiency and fairness?

People perceive AI technology as an efficient neutral entity and an equalizer (Hang & Chen, 2022; Agarwal, 2018, p. 920). In general, AI technology and AI decision-making are perceived positively (Sartori & Bocca, 2022; Araujo et al., 2020). Furthermore, research demonstrates that people evaluate AI decision-making as being “on par or even better” than decision-making by human experts (Araujo et al., 2020, p. 611). Automated decision-making in government generally elicits positive feedback from citizens compared to human decision-making (Gerlich, 2023). Consistent with Suchman's (1995) definition of legitimacy, if people assess AI decision-making as being better than human decision-making, this would equate to a greater perceived legitimacy. Subsequently, this study will assume that perceptions of AI decision-making will positively influence the perceived legitimacy, efficiency, and fairness of bureaucratic decisions that implicate AI technology.

The potential pathologies of bureaucratic discretion are at the forefront of this study on AI usage. This study measures whether bureaucratic decisions in problematic situations will be perceived differently if bureaucrats use AI technology in the decision-making process. Conditions of principal-agent dilemma, for instance, empower bureaucrats to exercise their discretion in a way that leads to inefficient or biased outcomes. Evidence suggests that bureaucrats use various means to minimize their workload and stress, even to the extent that they discriminate based on social class, ethnicity, and sex. As Denhardt (199) writes: “all actors bring to their interactions with other preferences and concerns that affect their behavior” (p. 28). It is plausible to assume that in situations in which bureaucratic discretion is high, principal oversight is minimal, and principal-agent problem conditions exist, bureaucrats will use AI technology to make their jobs easier or exert their biases. These assumptions are consistent with the enablement thesis (Buffat, 2013). However, if AI usage obscures bureaucratic discretion, as the research by De Boer and Raaphorst (2021) demonstrates, then AI usage would decrease the impact of such pathological bureaucratic behaviors. Thus, the obscurement theory suggests that AI usage by bureaucrats will positively impact perceived legitimacy, efficiency, and fairness.

2.7. Hypotheses

*H*1.a: The perceived legitimacy of bureaucratic decisions made under problematic conditions increases when bureaucrats use AI technology compared to when bureaucrats do not.

*H*1.b: The perceived efficiency of bureaucratic decisions made under problematic conditions increases when bureaucrats use AI technology compared to when bureaucrats do not.

*H*1.c: The perceived fairness of bureaucratic decisions made under problematic conditions increases when bureaucrats use AI technology compared to when bureaucrats do not.

3. Methodology

3.1. Research Design

This study is empirical, explanatory, and experimental. This study uses a vignette factorial survey as the primary method for data collection. A vignette factorial survey is appropriate for this study because vignettes are capable of presenting more dynamic and relatable situations than a simple question in a survey can, which is relevant because of the ambiguous understanding of AI technology by the public and the relevance of estimating the effects of AI usage on perceptions of relatable and salient bureaucratic decisions. The survey is administered through the Qualtrics website using a Design XM subscription (see Appendix A). Sampling is non-random. Intervention is random. Randomized intervention through Qualtrics

MX places respondents into either the control group (no AI) or the intervention group (AI). Respondents subsequently receive a different survey depending on their group.

3.2. Participants

This study principally targets young academics in Bachelor's and Master's programs who are studying at the University of Leiden (UL) and the University of Amsterdam (UVA). The population of this study is 74,701 students, as there are 33,701 students at UL and 41,000 students at UVA. This study initially contacted a total of 1,444 students to participate in the study. This is the sample frame. Of 1,444 people, 115 students decided to participate, making the response rate of this study 8%. Throughout the study period, 19 students did not complete the survey, resulting in a dropout rate of 18%. Thus, the final number of participating students in this study is 96. The study randomly assigned participants to two treatment groups: no AI (n=45) and AI (n=51). The study recruited students for the study during the period between November 26, 2023, and December 5, 2023.

This study's sample consisted of 96 participants with an age range of 15-30 years old (96%), who possess at least an undergraduate or graduate degree (88%), are primarily Caucasian (81%), male (male: 59%, female: 40%, other: 4%, prefer not to say: 2%), and will vote (88%). The sample is somewhat representative of the target population. Most students studying at UL and UVA are between the ages of 15 and 30, which is consistent with the demographics of this study's sample. The ratio of female to male students at UL and UVA is 60:40, whereas the gender distribution in this study's sample is skewed more towards males.[4] Furthermore, approximately 65-75% of UL and UVA students are likely Caucasian, whereas this study's sample consists of a higher percentage of Caucasians (81%). The non-representativeness of this

[4] This study derives the information from 2024 World University Rankings data (Times Higher Education, 2023a; Times Higher Education, 2023b).

study's sample is due to the non-random sampling method used. The homogeneity of the sample, however, presents a notable benefit. Homogenous samples increase validity because "on average, [they] yield estimates with clearer, albeit narrower, generalizability and, therefore, provide more accurate accounts of population effects and subpopulation differences" (Jager et al., 2017). Thus, this study can generalize the findings to young, Caucasian, and male students at UL and UVA.

The desired sample size is determined using confidence intervals, estimated population size, and margin of error. A correct sample size reduces the margin of error and allows for more precise and reliable inference. Furthermore, increasing the sample size allows for greater statistical power, which reduces the chance of type II errors and increases the potential to capture weak or heterogeneous effects. Based on a maximum margin of error of 10%, a confidence interval of 95%, a standard deviation of 0.5, and a population of 74,701, the minimum sample size after removing invalid survey responses should be n=96. It would not be realistic to decrease the margin of error anymore, as a margin of error of 5% with similar parameters would require 383 respondents. With a response rate of 8%, a sample size of 383 would require an approximate sample frame of 5,000 students. Gaining access to group chats that consist of this many students would be difficult based on this study's reach and resources. Therefore, this study uses the smallest margin of error possible that allows for a realistically attainable sample size to improve internal validity within the constraints of this study.

This study's sample method involves digitally sampling potential participants from student chat groups. The chat groups that this study samples are as follows: BSc Politics, Psychology, Law, and Economics (UVA), BSc Psychology (UVA), MSc Public Administration (UL), The Student Experience Housing Complex, and the DC van Hall Student Housing Complex. The study provides the link to the survey in a text message where potential participants are encouraged to participate to help contribute to the field of public administration and study the topic of legitimacy in government decision-making. The criteria for inclusion are being a student at UVA or UL. By targeting chat groups that include solely students currently enrolled in undergraduate and graduate educational programs at UL and UVA, this study can ensure that the respondents are highly educated and that other demographic characteristics are consistent with the target population.

This study uses the aforementioned sample frame for operational benefits such as convenience and cost-effectiveness. The target population of this study provides the opportunity to measure the perceived legitimacy of citizens who are highly educated. Highly educated individuals are likelier to vote (Ahearn et al., 2022). Moreover, younger adults are more likely to vote on AI-related issues in the future (Hogenhout & Takahashi, 2022). This study does not monetarily incentivize students nor provide other incentives for participating in the survey. Therefore, volunteer bias or self-selection bias is present in this study as respondents voluntarily elect to take the survey and are probably already interested in the topic. Volunteer bias is more likely in this study due to low response rates (Catalog of Bias, 2018). However, self-selection bias also presents a benefit, as this type of bias results in participation due to topic interest, which is often related to the quality of data collected (Brüggen et al., 2011).

3.3. Variables

The primary explanatory variable in this research is AI usage in bureaucratic decision-making. Other explanatory variables included in this study are gender, employment, ethnicity, level of education, familiarity with AI, and opinion of AI. The primary outcome variable is perceived legitimacy. Alternative outcome variables are perceived efficiency and fairness, which are assumed to mediate perceived legitimacy.[5]

[5] For a full list of variables see Table 2.

3.3.1. AI

Scholars first used the term AI in the 1950s to identify machines capable of displaying behavior in response to environmental stimuli in a human-like intelligent fashion (Russell & Norvig, 2019). Agreeing upon a universally accepted definition of AI is difficult because of its many applications of intelligence (Medaglia et al., 2021; McCorduck, 2004). The breadth of AI technology and its ambiguous definition means that public perceptions of AI often involve "biased and irrational [beliefs]" that are either reductive or hyperbolic (Brauner et al., 2023, p. 1). Among 28 countries, almost two-thirds of respondents believe that they understand AI but also have trouble identifying services that utilize AI (IPSOS, 2022). Generally, less than a quarter of people have a good or expert understanding of AI (Fritsch et al., 2022). To effectively capture the effect of AI usage on perceived legitimacy, this study must properly operationalize the often ambiguous concept of AI.

AI technology can be divided into two categories -- augmented intelligence and automated intelligence.[6] Since AI algorithms today are principally utilized as "decisional aides," assisting human decision-making, this study will operationalize AI as augmented intelligence (Alon-Barkat & Busuioc, 2022, p. 154). Thus, the AI intervention group receives a survey featuring vignettes in which AI technology aids a human operator in decision-making. The no AI control group takes a survey featuring vignettes in which a human decides without the assistance of AI technology. Specifically, this research design uses three applications of AI technology: an AI-operated CCTV camera, a machine-learning algorithm that employs AI technology, and an AI-powered risk-assessment model (see Appendix A, S5-7). The descriptions of AI in these

---

[6] Automation "involves replacing people with machines in the performance of certain tasks…without human intervention or guidance." Augmentation "connotes adding machines to a work environment" (Spence, 2022, p. 244).

scenarios are simplistic but represent concrete examples of what would otherwise be the ambiguous concept of AI. The survey never explains AI as a term to the respondents because AI is generally not well understood. Thus, if AI usage is to be measured, this study joins AI as a concept with somewhat understandable technology – a camera, an algorithm, and a model. Furthermore, these AI applications are implemented in this study because they correspond to existing AI technologies.

3.3.2. Legitimacy

The primary outcome variable in this research is perceived legitimacy. Perceived legitimacy is operationalized and measured by survey questions that evaluate whether bureaucratic decisions under certain conditions are "desirable, proper, or appropriate" (Suchman, 1999, p. 575). The mediator variables of perceived efficiency and perceived fairness are measured similarly. The question format is as follows: How much do you agree with the following statement?: The [police officer/tax auditor/health inspector] makes a [a desirable, proper, or appropriate/efficient/fair] decision in this scenario" (see Appendix A). Respondents answer on a 5-point ordinal Likert scale [strongly agree, agree, neither agree nor disagree, disagree, strongly disagree].

3.4. Data Collection

The method used to collect data for this study is a factorial survey administered through the online Qualtrics MX platform, a survey website. All respondents receive traditional survey questions on demographics and potential causally relevant variables such as familiarity with AI and opinion of AI. Demographic questions are in multiple-choice format. An example of such a question is: "How old are you? (see Appendix A, Q37). The survey asks questions about potential causally relevant variables on a 5-point Likert scale. An example of such a question is:

“How much do you agree with the following statement?: I am familiar with AI technology (see Appendix A, Q33).

Survey length is strongly associated with non-response rates and non-completion of the survey (Galešić & Bošnjak, 2009). The ideal survey length is between 10 and 15 minutes (Revilla & Höhne, 2020). Qualtrics predicts that the survey will take 11.3 minutes to complete. Tourangeau and Yan (2007) highlight potential problems with surveys that engage in sensitive topics or include sensitive questions – non-response, attrition, and inaccurate responses. People often consider demographics, gender, income, and other demographic questions sensitive (Rosenfeld et al., 2015). Therefore, answer options for specific questions regarding age, for instance, are provided in ranges instead of string format to avoid potentially implicating anonymity (SurveyMonkey, 2023). Options for ethnicity come from categories used by the United States federal government (Office of Institutional Research, 2023). The survey asks demographic questions after the other questions. The survey does this to minimize the survey's invasiveness. The survey also communicates to respondents that it is anonymous, reducing the potential for social desirability bias (Nikolopoulou, 2023).

The experimental element of this study lies in its use of vignettes. A vignette is a “short, carefully constructed description…a systematic combination of characteristics" (Atzmüller & Steiner, 2010, p. 129). Experimental vignette surveys are fundamentally comparative – exploring how people respond differently to scenarios (Aguinis & Bradley, 2014). This design allows the researcher to determine how the changed variable affects the evaluation of the scenarios (McDonald, 2019). All respondents evaluate three vignettes involving a street-level bureaucrat deciding under problematic conditions. The three vignettes involve three different street-level bureaucrats – a police officer, a tax auditor, and a health inspector – under three conditions. The three conditions correspond to the three principal-agent dilemma conditions – information asymmetry, goal ambiguity, and moral hazard (see Table 1). When a respondent takes the survey, Qualtrics MX randomly assigns them to either the control group (no AI) or the intervention group (AI). The intervention group receives vignettes in which a bureaucrat uses AI in the decision-making process and the control group depicts a bureaucrat deciding without the aid of technology.

Table 1

Vignette Design

	Vignette 1	**Vignette 2**	**Vignette 3**
Street-Level Bureaucrat	Police Officer	Tax Auditor	Health Inspector
Principal-Agent Condition	Information Asymmetry	Goal Ambiguity	Moral Hazard
Treatment (AI Usage)	AI-Operated CCTV Camera	Machine Learning Algorithm	AI Risk Assessment Model

Note. This table demonstrates the elements in the three vignettes used in this study. The street-level bureaucrat row shows what government agent each of the three vignettes represents. The principal-agent problem condition row shows the type of problematic conditions that each of the three vignettes represents (see Section 2.2.). The treatment row shows the application of AI technology that the bureaucrat uses in each vignette – this is only applicable in the intervention group (AI).

This research paper employs a factorial survey design and asks respondents to judge the scenarios presented to them to estimate the effect of AI usage in bureaucratic decision-making on the perceived legitimacy, efficiency, and fairness of those decisions (Taylor, 2005). These judgments capture "normative judgments," what "ought to be" (Jasso, 2006, pp. 335, 352). These normative judgments correspond to definitions of normative or traditional legitimacy. Normative legitimacy closely informs how this study defines legitimacy, making the vignette research design particularly valuable (Suchman, 1995).

Furthermore, the benefit of a vignette factorial survey design is that it can present more realistic scenarios than could be captured by an ordinary survey question. (Atzmüller & Steiner, 2010). The vignettes represent relatable citizen-state interactions. From a methodological standpoint, "the proximity of an event to the audience" can influence the event's salience (Cobb & Elder, 1972). Thus, presenting vignettes that involve scenarios in which the respondents might find themselves can motivate them to consider their answers more carefully. The relatable scenarios and the vignette research design favor preferential decision-making over perceptual decision-making, requiring closer deliberation by the respondents and resulting in more meaningful responses (Dutilh & Rieskamp, 2015).

The public often considers the government an ambiguous "black box" (John, 1998, p. 39). When citizens interact with government institutions, they often do so through street-level bureaucrats (Subramony, 2017; Rainey et al., 2021; Lipsky, 1980). Using street-level bureaucrats to represent government decision-making in the vignettes allows the respondents to provide their assessment of the institutions through the proxy of understandable human agents. The potential benefit of relatability is emphasized by presenting vignettes that involve three types of street-level bureaucrat – a police officer, a tax auditor, and a health inspector. Especially in countries such as the Netherlands, which have been developing community policing programs for over 30 years, it is likely that those studying in the Netherlands would encounter or interact with police (Punch et al., 2004). Although not every person interacts with a tax auditor, almost every person files taxes annually. Although citizen-state interactions with health inspectors are likely scarce, many people eat at restaurants and the health inspector's decision in vignette 3 would affect them. These research design choices are an improvement from Starke and Lünich's (2020) research, in which respondents evaluate budgetary decisions made by the EU Commission and the EU Parliament at an organizational level. In contrast, this study employs relatable scenarios depicting decisions by individual bureaucrats familiar to the participants.

Moreover, the evaluations of perceived legitimacy provided by respondents are likely to be more salient, as the presented bureaucratic decisions would directly affect them in real life. Vignette 1 represents a situation in which a police officer pulls over somebody who has been drinking. The respondent will relate to the person being pulled over or perhaps to another driver on the road who is glad that the police officer prevented an intoxicated driver from causing a motor vehicle accident. Vignette 2 represents a situation in which a tax auditor does not issue a fine to someone who commits tax fraud. If the respondent is averse to paying taxes, they might relate to the person who avoided paying taxes. Alternatively, they might relate to the scenario because of a sense of justice about the tax auditor allowing tax fraud. Vignette 3 depicts a health inspector who decides to approve the health inspection of a restaurant despite health code violations, which results in multiple people contracting food poisoning. The respondents would likely relate to this decision if they had ever eaten at a restaurant or had food poisoning. Referring back to Starke and Lünich's (2020) research, the measurable impact of EU budgetary decisions is plausibly not as noticeable as a police officer's decision to prevent a drunk person from driving, allow a person to avoid paying taxes, or inadvertently cause severe food poisoning. Hence, this study designs the vignettes to depict decisions salient to the survey respondents.

Besides presenting relatable scenarios, salient decisions, and familiar street-level bureaucrats, the vignettes also depict the problematic conditions under which bureaucrats make decisions. Conditions that exacerbate the principal-agent problem are a particularly suitable setting to study the effects of AI usage on bureaucratic decision-making. Firstly, because many bureaucrats rely on a principal – the citizen they serve or their direct manager – the principal-agent problem is nearly ubiquitous in bureaucratic decision-making. Moreover, in situations of information asymmetry, goal ambiguity, or moral hazard, the bureaucrat is *a priori* placed under a problematic condition that will elicit a normative judgment by the survey respondents. All survey respondents in this study consider the principal-agent conditions problematic (information asymmetry: 88%, goal ambiguity: 94%, moral hazard: 99%). Subsequently, when a bureaucrat makes a decision in the vignettes, the respondent is already primed for the problematic nature of the decision. From a methodological standpoint, testing the effect of AI usage on perceptions of bureaucratic decision-making is more likely to elicit a deliberate evaluation by survey participants when the study designs the situations in each vignette so that the stakes of each bureaucratic decision are higher and preemptively understood to be problematic.

3.5. Data Analysis

Data is cleaned and prepared in Microsoft Excel. This study uses Stata 18 for statistical analysis. This analysis will aim to provide enough statistical evidence to affirm a statistically significant relationship between the variables, even though proving true causality is unachievable (Toshkov, 2016). Spearman’s rank correlation and the ordered logistic model, statistical models, are used to estimate the effect of the explanatory variable and other potential causally relevant variables on the outcome variable. A requisite assumption of the ordered logit model is the proportionality of odds. This study tests and verifies this assumption using an approximate likelihood-ratio test of proportionality of odds across response categories.

This study uses Spearman's rank correlation because it is an easily understood non-parametric test suited for ordinal variables. Unlike the ordered logistic model that measures log odds of placing into a higher category of an ordinal variable, Spearman's rank correlation measures the presence of a potentially non-linear monotonic relationship. This test is suitable to discover if AI usage in bureaucratic decision-making will have any effect, even non-linear, on perceptions of legitimacy. The ordered logistic model is a multivariate regression model that is especially well-suited to providing precise coefficients that can measure the odds ratio for different categories of an ordinal outcome variable. The ordered logit model is suited to this study's analysis because the survey used in this study measures most variables on an ordinal Likert scale. Since this study aims to ascertain the effect of AI usage on perceived legitimacy, an ordered logit test will provide insights into the odds of a bureaucratic decision being ranked higher or lower on an ordinal Likert scale if a bureaucrat uses AI in decision-making.

3.6. Validity and Reliability

The validity of vignettes depends on their design and the questioning. This research design changes only one variable per intervention group, operationalizes terms well, and asks neutral questions on a Likert scale to ensure vignette validity (Payton & Gould, 2022; Su & Steiner, 2018). Other potential threats to this study's validity are collider and selection bias. Collider bias could be present, considering the sample consists of highly educated students with similar characteristics. Thus, the sample composition could distort the effect of the explanatory variable on the outcome variable. Collider bias could lead to problems with causal inference to the target population. As aforementioned, non-random sampling comprises the external validity of this study. However, since the target population is relatively similar to the sample, this should not entirely undermine generalizability. Additionally, self-selection bias is present as respondents elect to participate. The research design minimizes such bias by controlling for potential causally relevant variables that capture population characteristics. Finally, the digital format of the survey guarantees concealed allocation. Thus, the respondents were unaware of their treatment group, minimizing the potential weakening of the treatment effect (Toshkov, 2016).

The survey design and the quality-control functions on the Qualtrics MX platform ensure reliability. The Qualtrics MX platform maintains data integrity by preventing respondents from taking the survey twice. This study further establishes the reliability of the results by asking a red-herring question (see Appendix A, Q35). This question asked the respondents to select “green” from the list of multiple-choice options. This study excludes results from the analysis if respondents did not answer correctly. This type of question ensures that respondents are deliberate in their responses and that each respondent is a human being, not a bot. Finally, this study minimizes information bias by using a digital survey platform, which accurately and consistently measures survey responses.

3.7. Ethical Considerations

There are no real ethical concerns related to this study. This study ensures informed consent in two ways. Firstly, the respondents were informed about survey anonymity, how their responses will be used, and their participation in the study is voluntary and can be withdrawn at any time (see Appendix A, S1). This study also establishes informed consent by asking respondents if they voluntarily consent to the survey (see Appendix A, Q1). This study excludes any surveys completed without an affirmative response to this question.

3.8. Results of Pilot Study

This study recruited respondents for the pilot study between November 23, 2023, and November 25, 2023. The pilot study had eight respondents. The pilot test was conducted three days before the study's survey and identified potential issues with the data collection (see Appendix B). All respondents agreed that the survey is professional and its questions are easily understood, logically ordered, take the correct amount of time to complete (Strongly Agree = 88%, Agree = 12%), and are not offensive (Strongly Agree = 100%). The results of this test informed this study's final methodology by correcting significant errors in Qualtrics' randomization mechanism.

4. Analysis

4.1. Data Preparation

Survey data is exported as a delimited .csv data file into Microsoft Excel. If a survey response is incomplete, that data point is removed. The data on the perceived legitimacy, efficiency, and fairness of the three vignettes is exported into 18 columns – nine for the control group and nine for the treatment group. The data is subsequently transformed by combining the eighteen columns into nine columns and creating a new column for the intervention group (no AI or AI). Moreover, any irrelevant data is discarded from the data pool. To ensure reliability variables are meticulously coded (see Table 2).

Table 2

Variables and Codes

Variable Name	Variable Type	Categories	Coded
group	binary	no ai, ai	0, 1
male	binary	not male, male	0, 1
caucasian	binary	not caucasian, caucasian	0, 1
asian	binary	not asian, asian	0, 1
vote	binary	no, yes	0, 1
gender	ordinal	female, male, other, prefer not to say	0, 1, 2, 3
age	ordinal	15-30, 30-45, 45+	1, 2, 3
education	ordinal	high school, bachelor, master, phd	1, 2, 3, 4

ethnicity	ordinal	caucasian, asian, hispanic, black, two or more, other	1, 2, 3, 4, 5, 6
employ	ordinal	unemployed, part time, full time	0, 1, 2
vignette	ordinal	vignette 1, vignette 2, vignette 3	1, 2, 3
familiar	ordinal	strongly disagree, disagree, neither agree nor disagree, agree, strongly agree	1, 2, 3, 4, 5
favorable			
leg_avg			
eff_avg			
fair_avg			
police_avg			
tax_avg			
health_avg			
measure_avg			
leg_police			
eff_police			
fair_police			
leg_tax			
eff_tax			
fair_tax			
leg_health			
eff_health			
fair_health			

Note. This table shows the list of variables used in this study. The table is organized by variable name, variable type, categories, and coding schema. The variables *male*, *caucasian*, and *Asian* are dummy variables. The labels *police*, *tax*, and *health*, correspond to the three vignettes used in the study (vignette 1 = police, vignette 2 = tax, vignette 3 = health). The labels *leg_avg*, *eff_avg*, and *fair_avg* correspond to the three outcome variables of this study, perceived legitimacy, efficiency, and fairness. The *average* variables measure the mean of the specific variable (e.g., *leg_avg* measures the average legitimacy evaluation across all three vignettes). Other variables provide additional insights into average evaluations per vignette (e.g., *police_avg* measures the average evaluation of vignette 1), specific evaluations of legitimacy, efficiency, and fairness per vignette (e.g., *leg_police* represents the legitimacy evaluation of vignette 1), and the average of all variables (e.g., *measure_avg*). Not every variable is used in this study's analysis.

4.2. Descriptive Statistics

The dispersion of *leg_avg*, *eff_avg*, and *fair_avg* are not asymptotically normally distributed. The non-normal shape, value distribution, and measurements of skewness and kurtosis for each variable demonstrate a non-normal distribution (see Table 3). This study uses Kernel density plots to visualize the distribution of observations (see Figure 1-3). The plots show that the data has multiple flat peaks and uncentered values.[7] Furthermore, the three variables are all platykurtic (kurtosis < 3). The variable *fair_avg* is more leptokurtic (kurtosis = 2.466) than *leg_avg* and *eff_avg*. The values of *leg_avg*, *eff_avg,* and *fair_avg* are all skewed away from the mean. Furthermore, the *eff_avg* variable is skewed right (skewness = - 0.363), and *the fair_avg* variable skews left (skewness = 0.375) The interquartile range of *eff_avg* is the highest (see Figure 4).

Moreover, the distribution of outcome variables is generally consistent across both treatment groups, as is evidenced by tests of similarity (see Table 6). The only exception to this is the dummy variable *caucasian*, which demonstrates a statistically significant dissimilarity between groups (p = 0.020). There are 10 more non-Caucasian respondents in the no AI control group than the AI intervention group. This difference between treatment groups indicates potential sampling bias if the variables *caucasian* exerts a statistically significant effect on the outcome variables. Based on Spearman's tests, the variable *caucasian* exerts a non statistically significant effect on *leg_avg*, *eff_avg*, and *fair_avg*, thus this difference does not affect this

[7] Measures of normality differ depending on the statistical software that is used. In Stata 18, a normal distribution is indicated by a skewness of 0 and a kurtosis of 3. Other assumptions of a normal distribution are a bell-curve, single peak, and values centered around the mean.

study's finding in any meaningful way (*leg_avg*: p = 0.416, *eff_avg*: p = 0.367 , *fair_avg*: p = 0.499).

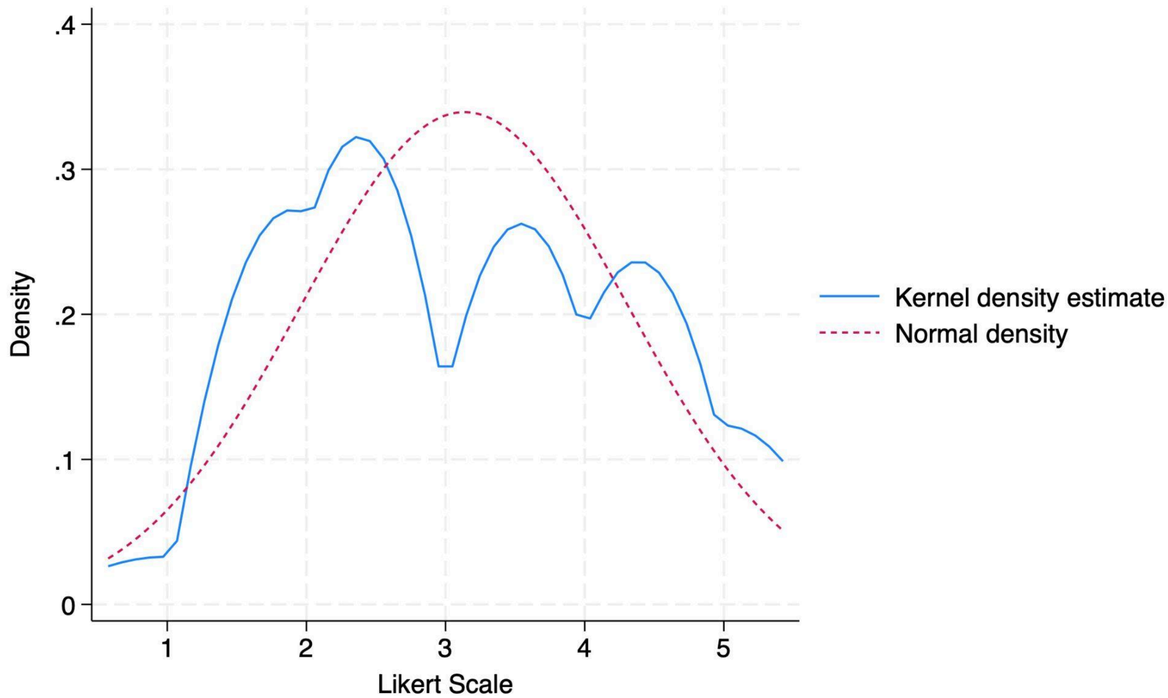
Table 3

Descriptive Statistics Table of Outcome Variables: leg_avg, eff_avg, and fair_avg

Variable	Min	Max	Mean	SD	Variance	Skewness	Kurtosis
leg_avg	1	5	3.135417	1.175359	1.381469	0.1658249	1.854612
eff_avg	1	5	3.822917	1.123854	1.263048	-0.3632029	1.879111
fair_avg	1	5	2.90625	1.006067	1.01271	0.3753649	2.465592

Figure 1

A Kernel Density Plot Overlaid with a Normal Distribution of Outcome Variable: leg_avg

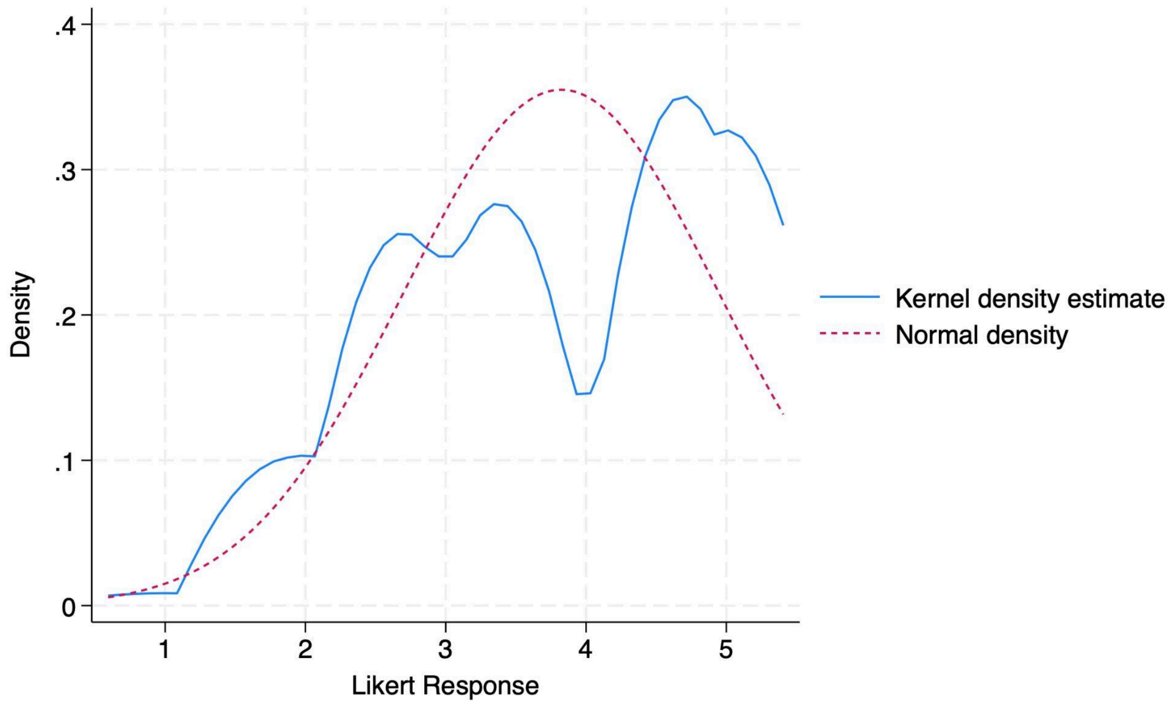


Note. This kernel density plot shows the smoothed distribution of data for *leg_avg* by coded Likert response (see Table 2). The plot is overlaid with a normal distribution. The y-axis shows the probability differential for *leg_avg*.[8]

[8] This is the probability of a point occurring between two values x1 and x2, represented by the area under the curve between those two points.

Figure 2

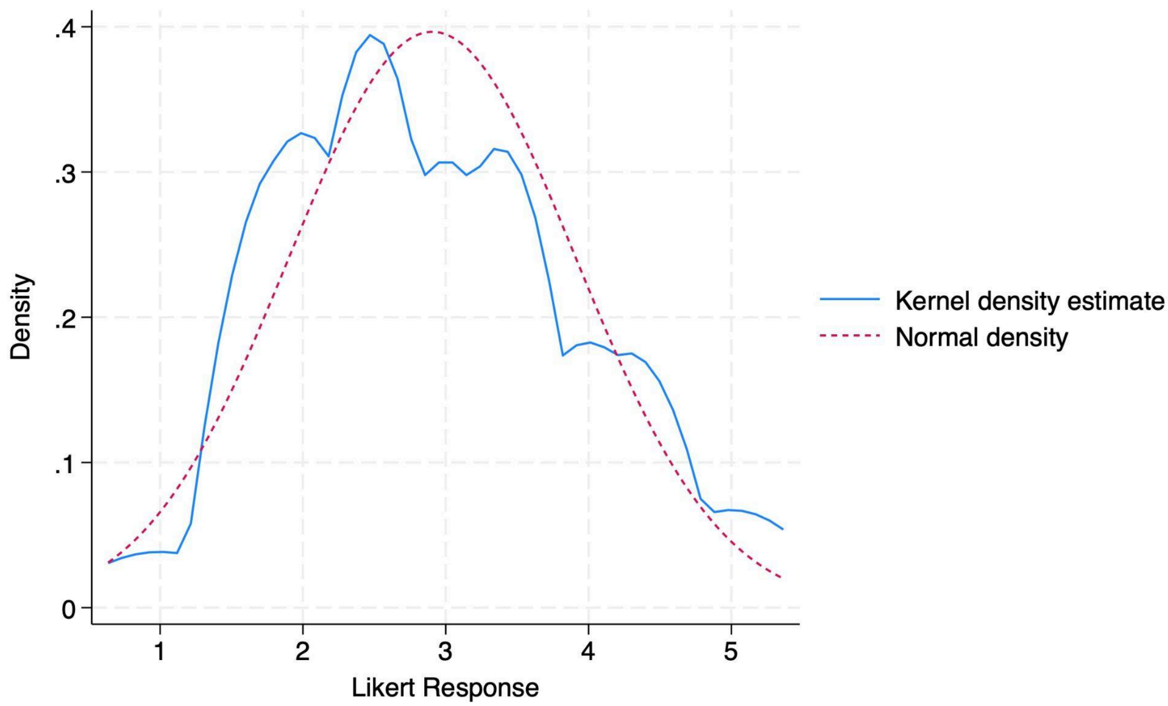
A Kernel Density Plot Overlaid with a Normal Distribution of Outcome Variable: eff_avg



Note. This kernel density plot shows the smoothed distribution of data for *eff_avg* by coded Likert response (see Table 2). The plot is overlaid with a normal distribution. The y-axis shows the probability differential for *eff_avg*.

Figure 3

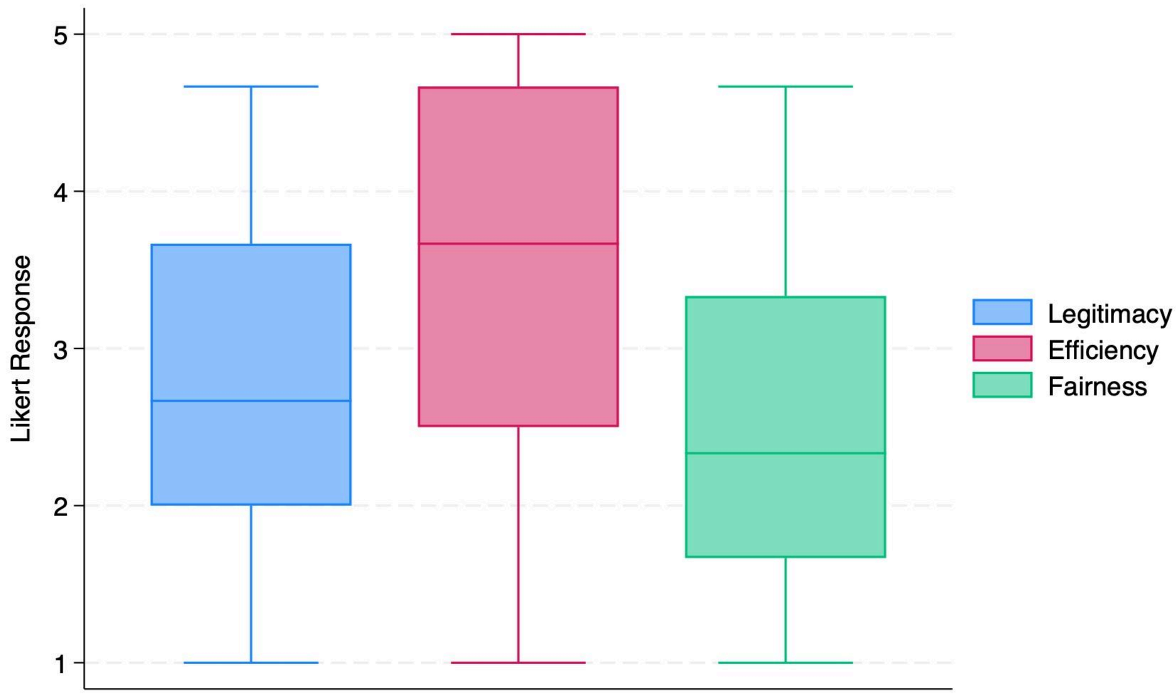
A Kernel Density Plot Overlaid with a Normal Distribution of Outcome Variable: fair_avg



Note. This kernel density plot shows the smoothed distribution of data for *fair_avg* by coded Likert response (see Table 2). The plot is overlaid with a normal distribution. The y-axis shows the probability differential for *fair_avg*.

Figure 4

A Comparative Box Plot of Outcome Variables: leg_avg, eff_avg, and fair_avg



Note. This comparative box plot shows the interquartile range, median, and minimum and maximum values for each of the three main outcome variables – *leg_avg*, *eff_avg*, and *fair_avg*.[9]

[9] The interquartile range is shown by the boxes. The solid line in each box represents the median value. The whiskers show the maximum and minimum values excluding any outliers.

Table 4

A Comparison of Explanatory Variables by Treatment Group

	group		
Variable	**No AI** 51 (53.1%)	**AI** 45 (46.9%)	**Test**
familiar	3.000 (1.510)	2.467 (1.517)	0.088
favorable	2.294 (1.540)	1.778 (1.396)	0.09
caucasian			
0	14 (27.5%)	4 (8.9%)	0.02
1	37 (72.5%)	41 (91.1%)	
asian			
0	45 (88.2%)	44 (97.8%)	0.073
1	6 (11.8%)	1 (2.2%)	
vote			
0	9 (17.6%)	3 (6.7%)	0.105
1	42 (82.4%)	42 (93.3%)	
education			
1	9 (17.6%)	3 (6.7%)	0.22
2	27 (52.9%)	30 (66.7%)	
3	15 (29.4%)	11 (24.4%)	
4	0 (0.0%)	1 (2.2%)	
employment			
1	20 (39.2%)	23 (51.1%)	0.248
2	16 (31.4%)	15 (33.3%)	
3	15 (29.4%)	7 (15.6%)	
gender			
0	23 (45.1%)	16 (35.6%)	0.271
1	25 (49.0%)	26 (57.8%)	
2	2 (3.9%)	0 (0.0%)	
3	1 (2.0%)	3 (6.7%)	
male			
0	26 (51.0%)	19 (42.2%)	0.391
1	25 (49.0%)	26 (57.8%)	
age			
1	48 (94.1%)	44 (97.8%)	0.566

2	2 (3.9%)	1 (2.2%)
3	1 (2.0%)	0 (0.0%)

Note. This table demonstrates the mean, standard deviation (in parentheses), and p-value for tests that measure the distribution of the explanatory variables between the control and treatment groups. The analysis uses a pooled t-test for continuous variables (*familiar* and *favorable*) and Pearson's test for the categorical variables.[10] The coded values of the variables are displayed based on the coding schemes for each variable (see Table 2).

4.3. Statistical Analyses

This study tests the effect of the explanatory variable *group* on outcome variables *leg_avg*, *eff_avg*, and *fair_avg* with three separate ordered logistic models. An ordered logistic regression of all potentially causal explanatory variables with *leg_avg* demonstrates a statistically significant and positive coefficient between *leg_avg* and the variables *group* and *male* (*group*: coefficient = 3.087, $p < 0.001$)(*male*: coefficient = 0.849, $p = 0.044$)(see Table 5).[11] The proportional odds assumption for this model is met ($p = 0.183$).[12] The predicted probabilities chart of *leg_avg* and *group* visualizes the likelihood of a respondent's evaluation of perceived legitimacy belonging to the Likert scale response category (see Figure 5). Albeit slightly above the predicted probability threshold of 0.5, this graph demonstrates a shift towards negative Likert scale responses if the respondent is in the no AI control group and vice versa if the respondent is in the AI intervention group.

---

[10] Although the variables *familiar* and *favorable* are ordinal, they are measured as continuous variables in this test to more easily capture the consistency of these variables between treatment groups.
[11] Ordered logistic model coefficients in this case can be interpreted as follows: For each one-unit increase in group, the log-odds of moving to a higher likert scale category (strongly disagree…strongly agree) increase by the coefficient value.
[12] As aforementioned, this analysis uses the approximate likelihood-ratio test of proportionality of odds across response categories to measure the proportional odds assumption.

Furthermore, an ordered logistic regression of all potentially causal explanatory variables with *eff_avg* also demonstrates a statistically significant and positive coefficient between *eff_avg* and the variables *group* and *asian* (*group*: coefficient = 3.170, $p < 0.001$)(*asian*: coefficient = 2.367, $p = 0.021$)(see Table 5). The proportional odds assumption for this model is met ($p = 0.215$). The predicted probabilities chart of *eff_avg* and *group* shows that being in the AI intervention group (*group* = 1) increases the probability of strongly agreeing that a bureaucratic decision is efficient (see Figure 6). An ordered logistic regression of all potentially causal explanatory variables with *fair_avg* also demonstrates a statistically significant and positive coefficient between *fair_avg* and the variables *group* (*group*: coefficient = 3.170, $p < 0.001$)(see Table 5). The proportional odds assumption for this model is also met ($p = 0.059$). The predicted probabilities chart of *fair_avg* and *group* shows that being in the no AI control group (*group* = 0) increases the probability of disagreeing that a bureaucratic decision is fair (see Figure 7).

Table 5:

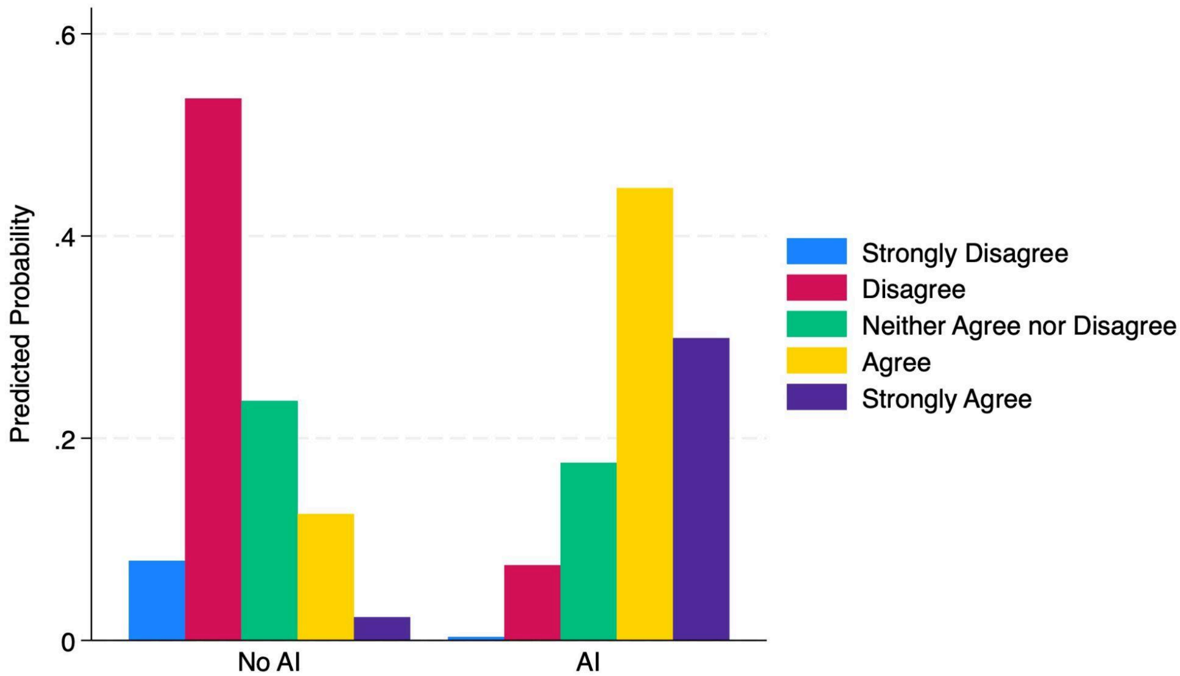
Ordered Logistic Models of Outcome Variables

		leg_avg		
Measure	**Coefficient**	**Std. err.**	**z**	**P > | z |**
group	3.087477	0.5212415	5.92	0
male	0.8492319	0.4219144	2.01	0.044
caucasian	0.0647244	0.6516944	0.1	0.921
asian	1.465657	0.9855491	1.49	0.137
vote	-0.5455115	0.6146065	-0.89	0.375
age	0.0159663	0.8339202	0.02	0.985
employment	-0.151576	0.2687302	-0.56	0.573
familiar	-0.2395566	0.1585169	-1.51	0.131
favorable	-0.0971521	0.1622408	-0.6	0.549
/cut1	-3.740921	1.505398		
/cut2	-0.5070987	1.412523	**Pseudo R^2** = 0.2118	
/cut3	0.9702409	1.417767		

Measure	Coefficient	Std. err.	z	P > | z |
/cut4	3.108249	1.457752		
		eff_avg		
Measure	**Coefficient**	**Std. err.**	**z**	**P > | z |**
group	3.170281	0.5344953	5.93	0
male	0.6363036	0.4589346	1.39	0.166
caucasian	0.1762525	0.6616051	0.27	0.79
asian	2.366874	1.026858	2.3	0.021
vote	0.5203536	0.6380593	0.82	0.415
age	-1.204809	0.8162305	1.48	0.14
employment	0.1131245	0.2866638	0.39	0.693
familiar	0.033068	0.1680413	0.2	0.844
favorable	-0.2273671	0.1616141	1.41	0.159
/cut1	-4.611692	1.822237		
/cut2	-1.681393	1.462253	**Pseudo R^2** = 0.2343	
/cut3	0.4464723	1.449183		
/cut4	1.780639	1.466715		
		fair_avg		
Measure	**Coefficient**	**Std. err.**	**z**	**P > | z |**
group	2.695605	0.5103645	5.28	0
male	0.6410738	0.4241763	1.51	0.131
caucasian	-0.3980098	0.6387367	-0.62	0.533
asian	0.5409137	0.9774161	0.55	0.58
vote	-0.2445358	0.6509902	-0.38	0.707
age	-0.1794243	0.818977	-0.22	0.827
employment	0.0792707	0.2614149	0.3	0.762
familiar	-0.0354538	0.1585551	-0.22	0.823
favorable	-0.1288586	0.1583927	-0.81	0.416
/cut1	-3.209133	1.489928		
/cut2	-0.0852478	1.405698	**Pseudo R^2** = 0.1590	
/cut3	2.048991	1.437029		
/cut4	3.963284	1.487252		

Figure 5

Predicted Probabilities Chart by Treatment Group: leg_avg



Note. This predicted probability chart shows the probability of each respondent's evaluation of legitimacy as measured on the Likert scale.[13]

[13] A threshold value of 0.5 is used to determine whether an observation is normal. As such, predicted probabilities above 0.5 are associated with an observation transitioning to a higher ordinal category, and those below 0.5 are associated with transitioning to a lower ordinal category. A predicted probability value of 0.55, for instance, denotes a 55% chance of belonging to the higher category and a 45% chance of belonging to the lower category.

Figure 6

Predicted Probabilities Chart by Treatment Group: eff_avg

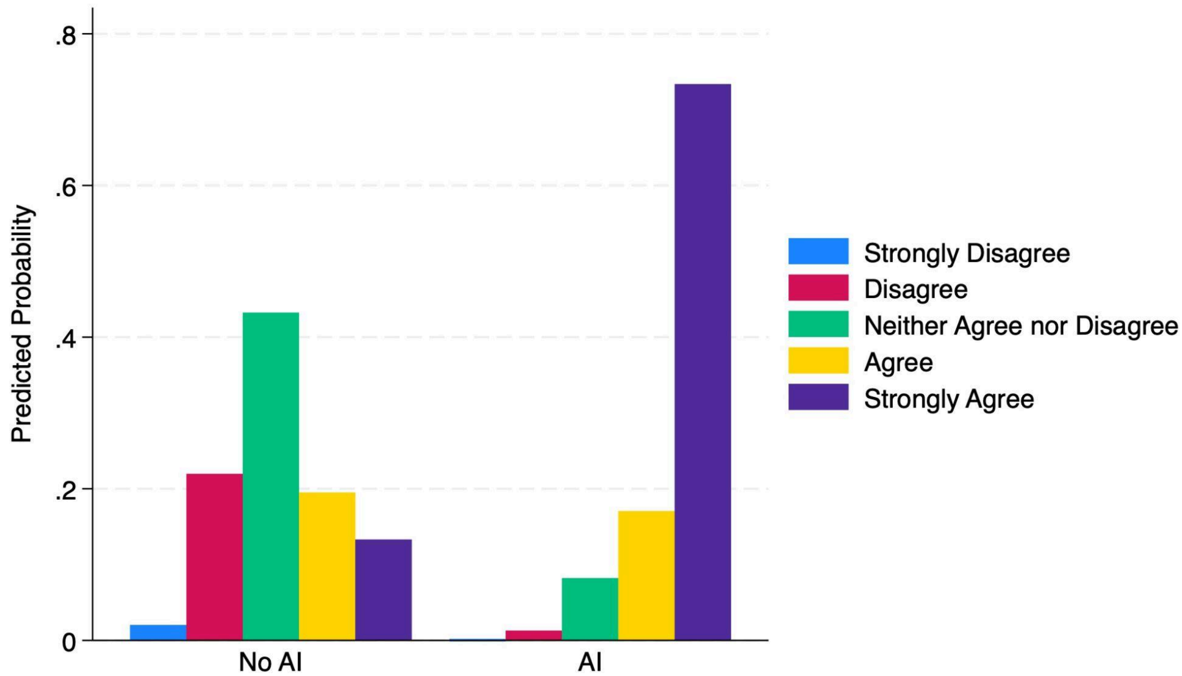
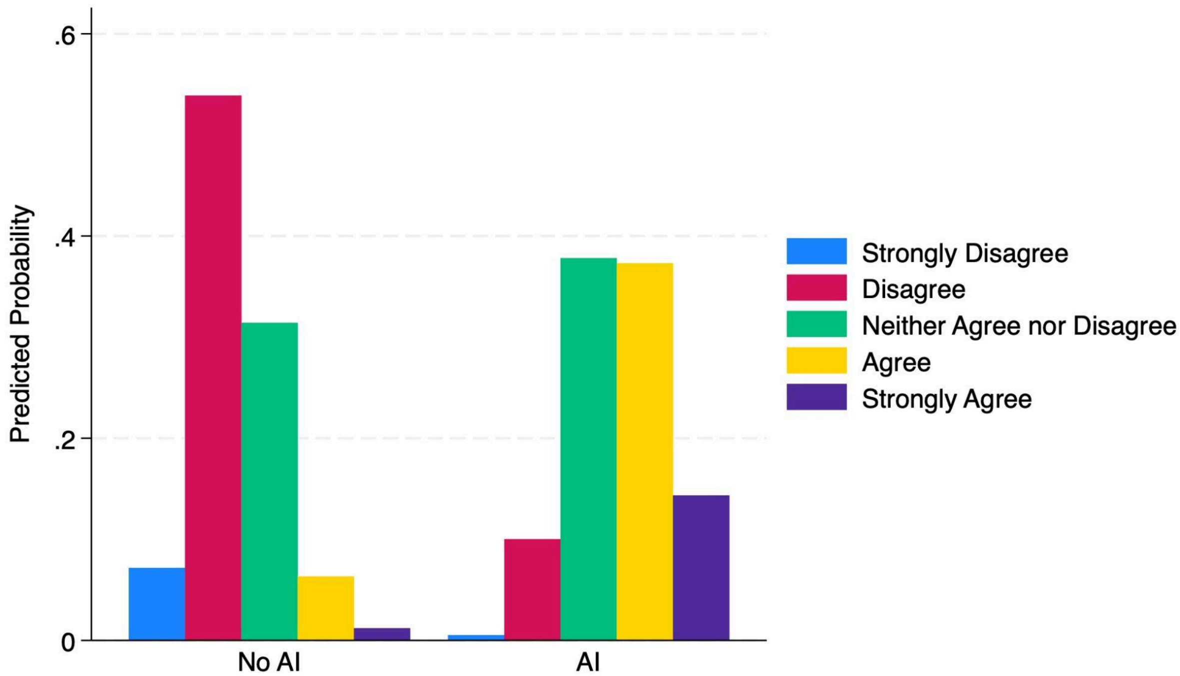


Figure 7

Predicted Probabilities Chart by Treatment Group: fair_avg

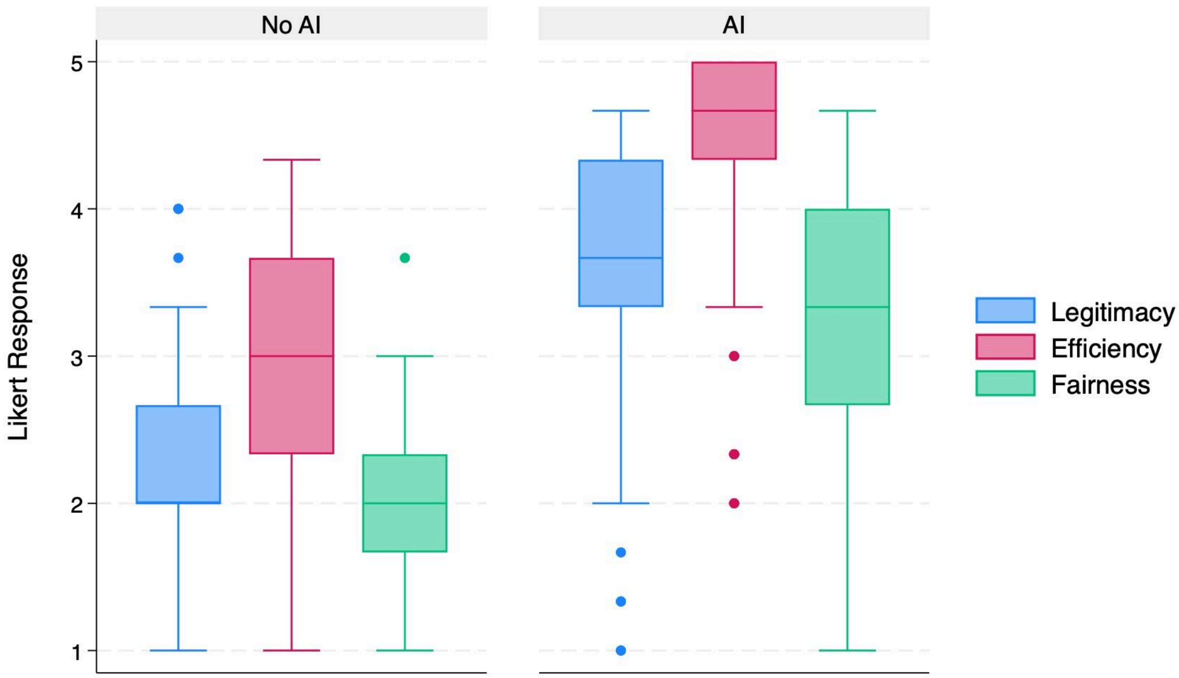


This study also uses Spearman's rank correlation to determine the effect of AI usage in bureaucratic decision-making on the outcome variables. The test finds a positive monotonic relationship between AI usage in bureaucratic decision-making and perceived legitimacy (ρ = 0.591), perceived efficiency (ρ = 0.651), and perceived fairness (ρ = 0.597). These results are statistically significant ($p < 0.001$). The positive effect of AI usage in bureaucratic decision-making on perceptions of legitimacy, efficiency, and fairness is consistent across all outcome variables (see Figure 8). There are outliers in the AI group that contradict the general trend to evaluate bureaucratic decisions more positively. The most notable trend among these outliers is that those who rate bureaucratic decisions as less legitimate than the mean for *leg_avg* (M < 3.13) rate decisions as more efficient than the mean for *eff_avg* (M > 3.82). Furthermore, the interquartile range of *eff_avg* is far narrower than that of *fair_avg*. This trend reverses in the no AI group, where evaluations of *fair_avg* and *leg_avg* are confined to a smaller interquartile range.

Furthermore, the positive effect of AI usage in bureaucratic decision-making on perceptions of legitimacy, efficiency, and fairness is consistent across all three vignettes (see Figure 9). Even though respondents rated the principal-agent conditions represented in each vignette as equally problematic (information asymmetry: 88%, goal ambiguity: 94%, moral hazard: 99%), the perceived legitimacy, efficiency, and fairness of the third vignette is consistently evaluated lowest of all vignettes in both the control group and the intervention group. Both ordered logit models and Spearman's tests indicate a statistically significant negative relationship between *vignette* and *leg_avg* (coefficient = -0.835, $\rho = -0.374$, $p < 0.001$), *eff_avg* (coefficient = -0.582, $\rho = -0.262$, $p < 0.001$), and *fair_avg* (coefficient = -0.576, $\rho = -0.255$, $p < 0.001$).

Figure 8

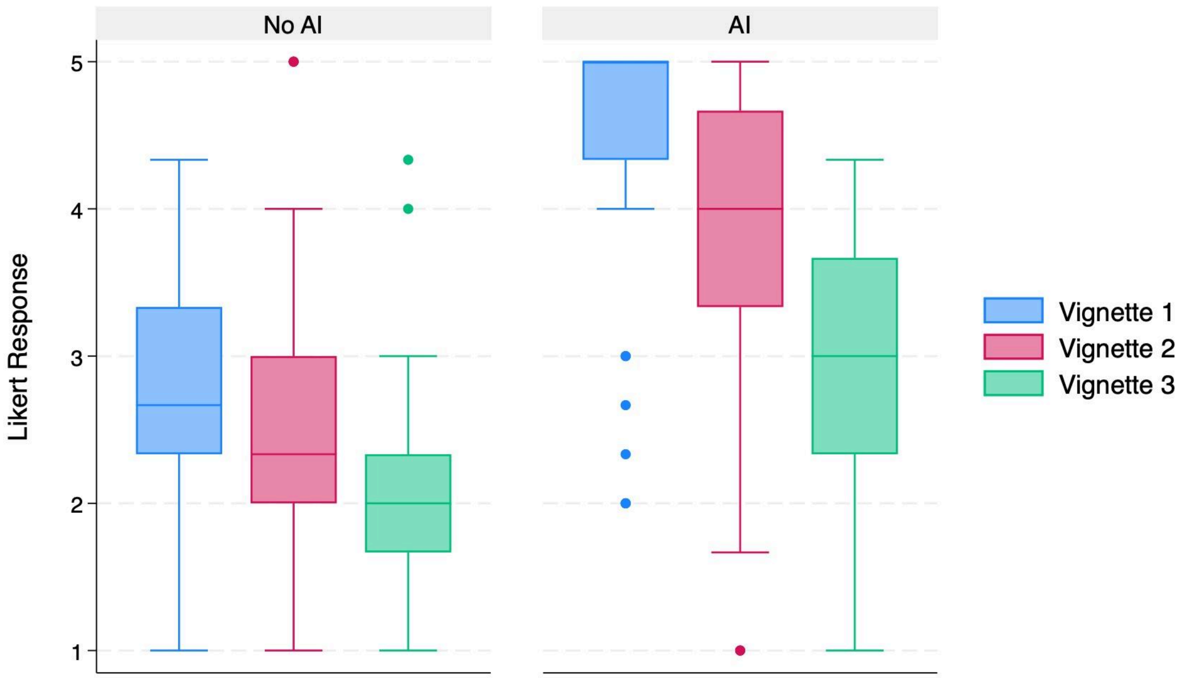
Box Plot of Outcome Variables by Treatment Group



Note. This comparative box plot shows the difference of median, interquartile range, outliers, and minimum and maximum values of *leg_avg*, *eff_avg*, and *fair_avg* by *group*.

Figure 9

Box Plot of Average Vignette Evaluation by Treatment Group



Note. This comparative box plot shows the difference of median, interquartile range, outliers, and minimum and maximum values of the evaluations of the three vignettes in the study by *group*.

This study also tests the alternative outcome variables, perceived efficiency and fairness, for mediating effects through structural equation modeling. The model is not a perfect fit, suggesting other uncaptured relationships that contribute to the model ($\chi 2 = 18$, $p < 0.001$). The statistically significant positive coefficient for *eff_avg* (coefficient = 0.3760158, $p < .001$) indicates that higher levels of perceived efficiency are associated with higher perceived legitimacy. The statistically significant positive coefficient for *fair_avg* (coefficient = 0.6409776, $p < 0.001$) indicates that higher levels of perceived fairness are associated with higher perceived legitimacy. Furthermore, the statistically significant positive relationship of *eff_avg* (coefficient = 1.395207,

$p < 0.001$) and *fair_avg* (coefficient = 1.132898, $p < 0.001$) on *the group* indicates that the treatment group influences perceived efficiency and fairness.

4.4. Hypothesis Testing

*H*1.a: The perceived legitimacy of bureaucratic decisions made under problematic conditions increases when bureaucrats use AI technology compared to when bureaucrats do not.

*H*1.b: The perceived efficiency of bureaucratic decisions made under problematic conditions increases when bureaucrats use AI technology compared to when bureaucrats do not.

*H*1.c: The perceived fairness of bureaucratic decisions made under problematic conditions increases when bureaucrats use AI technology compared to when bureaucrats do not.

This study employs an ordered logistic model and Spearman's rank correlation to examine the hypotheses relating to the primary explanatory variable – AI usage in bureaucratic decision-making – and the outcome variables encompassing perceived legitimacy, efficiency, and fairness. The analysis provides compelling evidence, indicating a robust association between AI usage in bureaucratic decision-making and increased perceived legitimacy, efficiency, and fairness of decisions. Spearman's test provides a statistically significant and moderately positive monotonic relationship between AI usage and perceived legitimacy ($\rho = 0.591$, $p < 0.001$). Furthermore, a statistically significant and strong positive monotonic relationship is observed between AI usage and perceived efficiency ($\rho = 0.651$, $p < 0.001$), as well as fairness ($\rho = 0.597$, $p < 0.001$).

According to the ordered logistic model, being in the AI treatment group is associated with an increase in the log odds of respondents rating decisions as more legitimate (coefficient = 3.087, $p < 0.001$), efficient (coefficient = 3.170, $p < 0.001$), and fair (coefficient = 2.696, $p < 0.001$). The results from Spearman's test and the ordered logistic model offer robust evidence of a positive association between AI usage in bureaucratic decision-making and perceived legitimacy, efficiency, and fairness. All these associations are statistically significant, leading to the failure to reject the alternative hypotheses at a 5% significance level ($\alpha = 0.05$).

5. Discussion

5.1. Interpretation and Implications

This study focuses on the effect of AI usage in bureaucratic decision-making on perceived legitimacy. In pursuit of this goal, this study considers perceived efficiency and perceived fairness mediating variables to perceived legitimacy. That is to say, if there is an increase in these mediating variables, this translates to increased perceived legitimacy. The analysis finds a strong positive monotonic relationship between the mediating variables and perceived legitimacy (efficiency: $\rho = 0.760$, fairness: $\rho = 0.795$). Structural equation modeling determines if these variables act as mediators to perceived legitimacy. A positive and statistically significant coefficient for perceived efficiency (coefficient = 0.3760158, $p < 0.001$) and fairness (coefficient = 0.6409776, $p < 0.001$) reveal that increased perceptions of efficiency and fairness align with enhanced perceptions of legitimacy. There are significant positive relationships between perceived efficiency (coefficient = 1.395207, $p < 0.001$) and fairness (coefficient = 1.132898, $p < 0.001$) in the treatment group. These findings collectively suggest that both perceived efficiency and fairness serve as mediating variables. This is consistent with existing findings on the relationship between efficiency, fairness, and legitimacy (Melamed, 2012; Franck, 1998; Farrar, 2022; Jeong & Kim, 2019).

When comparing the interquartile range of these measurements between treatment groups, the findings reveal that respondents were much more decisive and consistent in their evaluations of fairness and legitimacy in the no AI control group than the AI intervention group, as indicated by the smaller interquartile range (see Figure 8). In contrast, evaluations of efficiency demonstrated an opposite change between treatment groups. This plausibly implies a greater consensus among respondents about the effect of AI usage in bureaucratic decision-making on perceived efficiency than on perceived legitimacy or fairness. The predicted probabilities chart of the ordered logistic model between perceived efficiency and AI usage in bureaucratic decision-making shows that being in the AI intervention group predicts that 75% of the values will fall in the "strongly agree" category on the Likert scale (see Figure 6). Furthermore, AI usage in bureaucratic decision-making exerts the largest effect on perceived efficiency, as demonstrated by the coefficient in the ordered logistic test (coefficient = 3.170) and Spearman's test ($\rho = 0.651$). One implication of the narrower interquartile range of perceived efficiency in the AI intervention group and the comparatively strong coefficient and correlation values is that AI usage in bureaucratic decision-making exerts a more recognized and strong effect on perceived efficiency than on measures of fairness or legitimacy. A plausible explanation for this remarkable effect is the perception of AI as a neutral entity capable of efficient decisions (Boyd & Crawford, 2012; IPSOS, 2022; Alon‐Barkat & Busuioc, 2022; Araujo et al., 2020). AI usage in decision-making increases perceptions of efficiency, even if a human bureaucrat ultimately makes the decision.

For some respondents, there seemed to be a trade-off between legitimacy and efficiency. Outliers in the AI intervention group rated bureaucratic decisions as significantly less legitimate than the mean. They rated the same decisions as significantly more efficient than the mean (see Figure 8). Even when decisions were deemed more illegitimate by the outlier respondents, they still perceived AI usage as contributing to the efficiency of the decision. This seemingly indicates that, at least for some respondents, efficiency as a measure does not influence and perhaps even decreases perceptions of legitimacy. An explanation for these outliers is that these respondents might recognize efficiency as an instrumental value that empowers bureaucrats to execute illegitimate decisions. This is consistent with the enablement thesis, which argues that AI technology empowers bureaucrats to utilize discretion (Buffat, 2013). Subsequently, if bureaucrats are perceived by respondents as illegitimate or biased, then increasing the efficiency of these bureaucrats would empower them to exert their biases. These outlier respondents' deviation from the norm makes sense since perceptions of legitimacy are subjective and depend on socioeconomic and cultural factors (Brandt et al., 2020). Thus, perhaps to these outlier respondents, certain uncaptured demographic variables influence their understanding of efficiency as an instrument for illegitimate behavior.

In a similar vein, this study finds that being of Asian ethnicity increases the log odds of evaluating a bureaucratic decision as efficient (coefficient = 2.367, $p = 0.021$). According to a report by the International Institute of Communications, AI technology is generally perceived positively and is readily used in Southeast Asian countries (2020). China, one of the most prominent countries in Asia, demonstrates a willingness to employ AI technology to surveil its people (Davies, 2021). As a result, people in Asia are more optimistic about the technology and its potential to contribute to productive capacity (Sells, 2023). Thus, in keeping with the understanding that perceptions of legitimacy and its mediating variables are subjective, being Asian might be associated with higher perceptions of efficiency for these reasons. Another explanatory variable that affected evaluations of the vignettes was gender.

The ordered logistic model revealed a statistically significant coefficient for being male (coefficient = 0.849, $p = 0.044$), indicating a positive association with the outcome variable. This suggests that being male is associated with an increase in the log odds of evaluating a bureaucratic decision as legitimate. While the effect size is modest, the statistical significance implies that this association is unlikely to be due to random chance. This finding aligns with existing studies that have found that men report higher familiarity with AI, plausibly explaining this result (IPSOS, 2022).

Next, since respondents rated the principal-agent conditions represented in each vignette as equally problematic (information asymmetry: 88%; goal ambiguity: 94%; moral hazard: 99%), one would expect that measures of legitimacy, efficiency, and fairness would be somewhat consistent. However, in both treatment groups, the third vignette is rated lowest in all measures. On one hand, this could be due to factors such as vignette design, which implicates more victims in the poor decision made by the bureaucrat in vignette 3. However, visualizing the distribution of measurements by vignette shows an incremental decline in evaluations between vignettes 1 and 3. When the relationship between vignettes and measurements is tested, the results indicate that the vignette order exerts a statistically significant negative effect on the average measurement of legitimacy, efficiency, and fairness ($p < 0.001$). Even though the survey completion time falls within the recommended range, these findings point to the presence of survey fatigue (Revilla & Höhne, 2020).[14] However, these associations are generally weak and do not necessarily undermine the findings of this study.

Furthermore, since the vignettes represent three different types of bureaucrats, the findings of this study imply that AI usage in bureaucratic decision-making will exert a positive effect on perceptions of legitimacy despite potential confounders inherent in the type of bureaucrat. For instance, as aforementioned, understandings of legitimacy depend on socioeconomic and cultural factors. Thus, there might be potential preconceived notions about

---

[14] Survey fatigue is the "lack of motivation to participate in assessments [which can] impact response behavior" (DiLeonardo et al., 2022).

police officers, tax auditors, or health inspectors that might influence the respondents' perceptions of each vignette. Such preconceived notions are not explicitly captured in this study. However, what is notable is that if one disregards the potential survey fatigue, it is vignette 1 that elicits the most positive evaluations of legitimacy, efficiency, and fairness. This, despite the fact that the police officer is probably one of the more contentious types of bureaucrats represented in the vignettes. This can be explained by the fact that this study's sample is composed primarily of Caucasian individuals, who on average perceive the police as more effective than other ethnic groups do (Morin & Stepler, 2016).

Moreover, vignette 3 provides a new understanding of automation bias. Automation bias presents itself as a remarkable conundrum for bureaucratic discretion, as when bureaucrats attempt to maximize their resources and minimize stress they are likely to heavily lean on AI technology to do so – especially in problematic situations. In the intervention group, the bureaucrat in vignette 3 fits the definition of automation bias, which is an "overreliance on algorithmic advice" (Alon-Barkat & Busuioc, 2022, p.153). Although evaluations of legitimacy, efficiency, and fairness in vignette 3 are lower than in vignettes 1 and 2, these evaluations are much higher in the intervention group than in the control group (see Figure 9). These findings contradict what the field assumes about automation bias, that automation bias is necessarily a negative phenomenon. However, the findings of this study seem to indicate that reliance – even over-reliance – on AI technology can plausibly provide a way for bureaucrats to shirk the responsibility of their decision-making and improve perceptions of legitimacy.

One implication of analyzing the results by vignette is that AI usage in bureaucratic decision-making exerts a statistically significant ($p < 0.001$) and notable positive effect on perceptions of legitimacy, efficiency, and fairness regardless of the problematic condition that the bureaucrat finds themself in (see Figure 9). As aforementioned, respondents consistently rated the conditions of information asymmetry, goal ambiguity, and moral hazard as problematic. This is reflected in the evaluations of the three vignettes, which are generally negative and in the lower range of Likert responses. These findings point to the possibility that implicating AI usage in bureaucratic making has the ability to alleviate potential detrimental consequences to perceived legitimacy, efficiency, and fairness stemming from misused bureaucratic discretion under problematic conditions. Based on the theory of the principal-agent dilemma, such conditions would elicit a strong negative response since the bureaucrat is contradicting the direction of their principal – regardless of what tools the bureaucrat uses to make the decision. However, this is evidently not the case, as using AI technology changes the evaluation in a significant way.

This study finds a statistically significant ($p < 0.001$) positive relationship between AI usage in bureaucratic decision-making and perceptions of legitimacy, efficiency, and fairness. The positive effect of AI usage in bureaucratic decision-making on perceptions of legitimacy, efficiency, and fairness is consistent across all outcome variables (see Figure 8). The positive effect of AI usage in bureaucratic decision-making on perceptions of legitimacy, efficiency, and fairness is consistent across all three vignettes (see Figure 9). Specifically, the central findings of this study indicate that AI usage in bureaucratic decision-making – especially under problematic conditions – positively affects perceived legitimacy, both directly and through mediating measures of efficiency and fairness. This contradicts findings that there is little difference in the perceived legitimacy between human and AI-assisted decision-making (Starke & Lünich, 2020). Buffat's (2013) curtailment thesis partly explains the effect of AI usage in bureaucratic decision-making on perceptions of legitimacy. Per the curtailment thesis, AI usage in decision-making can co-opt the discretion of bureaucrats. The vignettes in this research present bureaucrats as principally self-interested, biased, and willing to make decisions that are not always in the best interest of their principals. Thus, in the vignettes where the bureaucrats utilize AI technology, the discretion and subjectivity of an inherently biased decision-making process are plausibly reduced, leading to a more favorable evaluation of these scenarios.

The findings that AI usage increases perceived legitimacy, even in detrimental situations, demonstrate the curtailment thesis put forth by Buffat (2013). Although Buffat (2013) suggests that technology will take away bureaucratic discretion in practice, this study suggests that AI technology can make citizens perceive less discretion. This is similar to what Jorna and Wagenaar (2007) conclude – the technology does not reduce bureaucratic discretion but merely obscures it. This is evidenced by the fact that although the vignettes are consistent across the treatment groups in all ways except the implication of AI technology in decision-making, perceptions are drastically different. In other words, a feasible explanation for this is that when bureaucrats utilize AI technology to make their decisions this alleviates some of their blame in the decision because bureaucrats share the responsibility for the decision between themselves and the technology.

It is plausible to assume that bureaucrats will not only use AI to make their jobs easier but will shirk their responsibilities for making poor decisions by shifting blame to the technology they employ. This is consistent with research on bureaucratic decision-making, which highlights the fact that bureaucrats wish to reduce uncertainty, minimize conflict, protect themselves, maximize agency and personal resources, and are fundamentally self-interested and willing to do what they can to minimize their workload and stress (Lipsky, 1980; Maynard-Moody & Musheno, 2003; Prottas, 1979; Harrits & Møller, 2014; Epp et al., 2014; Dubois, 2010l;

Hinterleitner & Wittwer, 2022; Gajduschek, 2003). Thus, if AI usage can obscure bureaucratic discretion and subsequently provide a buffer against negative judgment, it is reasonable to assume that there is an incentive for bureaucrats to rely on AI in decision-making to reap this benefit.

The findings of this study seem to evidence the utility associated with AI usage's capability to obscure bureaucratic discretion – it can provide a buffer against negative perceptions, it can increase perceptions of efficiency, and it can potentially reduce prejudiced evaluations of certain bureaucrats like police officers. The findings of this research paper complicate Buffat's (2013) theses, provide potential positive benefits of automation bias, evidence Jorna and Wagenaar's (2007) identification of the technology's ability to obscure discretion, and build on research by Starke and Lünich (2020) by demonstrating that in problematic scenarios AI-assisted decision making fares better than human decision making alone.

5.2. Limitations

A perfect research design is not expected, but a design that sufficiently answers this study's research question. This is especially true since legitimacy depends on context (Adams et al., 2019). The perceived legitimacy of decision-making in the vignettes might be higher in other contexts – this is not tested in the research at hand. Moreover, public opinion on specific issues is unstable, and research fails to reproduce consistent results; people can be asked the same question a few months later and will provide entirely different answers (Dye, 1998). Perceived legitimacy, in particular, is known to change over time (Risse‑Kappen & Stollenwerk, 2018). The research design is also limited in its reliance on the assumption that the government will adopt AI technology. It is not evident that this will be the case, especially as many governments are passing legislation to restrict the use of such technology (European Parliament, 2023). There is a notable lack of technology uptake by governments due to the absence of necessary infrastructure or the desire to maintain government budgets by restricting efficiency (Devarajan, 2022).

Moreover, because of the complexity of the issues involved, it is important to consider that public decision-making on the topic of AI usage in government should not be expected to be rational (John, 1998, p. 34). There are many factors that can potentially affect voting such as rational voter ignorance and apathy (Downs, 1957). Furthermore, voters face few negative individual downsides from elections, making them less motivated to vote in their interests (Brennan & Lomasky, 1993). Although a majority of the respondents in this study reported that they would vote in the upcoming election (88%), it is not evident that opinions on something as specific as the scenarios presented in the vignettes of this study will translate to broad and often ambiguous electoral platforms of political candidates. It is not clear that the opinions voiced by the respondents will directly carry over to electoral outcomes or meaningful political decisions that will affect governance. Thus, we cannot infer that the respondents of this study will or will not vote in favor of AI adoption, for instance.

Furthermore, this study cannot generalize the effect of AI usage in bureaucratic decision-making on perceived legitimacy due to the non-random sample. The low external validity of the research design means that these findings, with a margin of error of 10%, can only potentially infer about populations similar to the sample frame – Caucasian, educated, young people studying in the Netherlands at UL and UVA.

6. Conclusion

6.1. Executive Summary

Bureaucrats play an integral part in governance. Therefore, the government empowers bureaucrats with significant discretion in their work. Under certain conditions, however, bureaucrats misuse their discretion to make inefficient, potentially biased decisions consequential to the citizens they serve and the institutions they represent. Poor decisions on behalf of bureaucrats can hamper the perceived legitimacy of government organizations. This presents a salient problem for public administrators everywhere, as legitimacy is necessary to the functioning of government institutions. AI technology is a potential solution to the problematic effect of poor bureaucratic decision-making on perceived legitimacy, as the technology is known for its neutrality, efficiency, and influence on bureaucratic discretion. This study aimed to discover how the perceived legitimacy of bureaucratic decisions made under problematic conditions changes when bureaucrats use AI technology. The research question of this study is: What is the effect of AI technology usage by bureaucrats under problematic conditions on the perceived legitimacy of bureaucratic decisions? This study found a statistically significant ($p < 0.001$) positive relationship between AI usage in bureaucratic decision-making and perceived legitimacy and mediating variables. These results suggest that AI usage can plausibly alleviate the impact of consequential bureaucratic decisions on perceived legitimacy by obscuring bureaucratic discretion. Furthermore, AI usage in bureaucratic decision-making seemingly exerts a stronger positive effect on perceived efficiency than on perceived legitimacy and fairness.

An alternative causal explanation for the results of this study is the characteristics inherent in the sample. It is possible that the AI group appears to support problematic bureaucratic decisions more than the non-AI group because of uncaptured characteristics that are shared among students of UVA and UL. An example might be AI use. One study finds that almost two-thirds of students pursuing higher education use AI-based tools. Males in particular were more likely to use such AI tools (Von Garrel & Mayer, 2023). This study's sample is evidently composed of male students pursuing higher education. One would assume that the variables used in this study that measure familiarity and favorability of AI technology would capture AI usage by the respondents. However, AI technology is a notoriously ambiguous concept. As such, perhaps students are unaware that tools such as Chat GPT are based on AI. This assumption is certainly consistent with findings in the IPSOS report (2022), which shows that around two-thirds of people had trouble identifying technologies that utilize AI. Thus, this sample's composition might have influenced the effect of AI usage in bureaucratic decision-making on perceived legitimacy.

6.2. Future Research

Future research would benefit from a longitudinal study of perceived legitimacy in government decision-making that employs AI, which would capture any changes in perceptions over time. Furthermore, testing similar variables to this study in different contexts and at varied levels of analysis can provide additional insights into the persistence of the AI effect on perceived legitimacy. Moreover, this research focuses on street-level bureaucrats in situations relatable to the average citizen. Perhaps in more formal or unfamiliar settings, the results would differ. An evident improvement to this study is the use of randomized sampling from a larger target population, which would improve the external validity of the research's findings.

Initially, the research aimed to use two interventional groups, some-AI and all-AI, like Starke and Lünich (2020). Due to constraints in acquiring an adequate sample size for two interventional groups, this study only uses one treatment group – AI. However, as Katzenbach and Ulbright (2019) mention, "the *degree of automation* matters greatly because the legitimacy of governance regimes relies on the responsibility and accountability of a human decision-maker" (p. 8). Therefore, it would be valuable to include more treatment groups that specify the extent of AI usage to capture the association between the degree of AI usage in decision-making and perceived legitimacy.

Although this study's results indicate a strong positive association between AI usage in bureaucratic decision-making and perceptions of legitimacy, and former research finds that people generally favor AI decision-making over human decision-making, it does not mean that decisions made entirely by AI will prove to be superior over human decisions in practice. For instance, respondents' evaluations of problematic outcomes due to AI decision-making alone are notably different (Gerlich, 2023). This is also evidenced in Starke and Lünich's (2020) research, which finds that decisions made by AI with no human oversight detrimentally impacted perceived legitimacy. Moreover, in a review of automated vehicle collisions, Hidalgo et al. (2021) find that when AI technology decides on problematic outcomes, citizens judge the decision more severely than if the decision was made by humans alone. In other words, the protective obscuring effect of AI technology is plausibly eliminated if AI, with the approval of human administrators, is permitted to make decisions on its own. More research is therefore needed to explore this obscuring phenomenon in the context of fully autonomous AI technology making governmental decisions.

6.3. Policy Recommendations

As Denhardt (1999) remarks, "Public administrators live in a goldfish bowl, their every movement scrutinized by an often critical public," meaning that negative consequences resulting from the implementation of AI technology will fall on those public managers who drafted the policy (p. 15). Political preferences are malleable and can change (Risse-Kappen & Stollenwerk, 2018; Gelrich, 2023). Thus, it is prudent for public administrators who consider implementing AI usage to influence the public discourse on AI technology and emphasize the potential benefits and efforts to mitigate any potential harm from its implementation. Furthermore, public perceptions about the use of AI can result in non-adoption (Kieslich et al., 2021). Thus, managing public perceptions is important if public managers wish to implement AI technology in the execution of their tasks.

Thus, to avoid potential backlash, if a public manager decides to use AI technology, especially in fully autonomous functions, public managers should implement auditing systems to screen for potential bias or unintended consequences and make the results of these audits open to public scrutiny (Ma & Agarwal, 2007). Moreover, it is helpful to establish regulations on the use of AI, as AI technology is constantly evolving, meaning that normative standards of use are not always applicable (Ma & Agarwal, 2007). Just because the public approves of using AI technology in government decision-making in a certain context or for a particular function does not mean this approval extends to all contexts and tasks. Thus, implementing auditing systems and establishing clear guidelines using AI technology could mitigate the negative impact of government AI usage.

References

Aberbach, J. D., Putnam, R. D., & Rockman, B. A. (1981). Bureaucrats and Politicians in Western Democracies. *Harvard University Press EBooks*. https://doi.org/10.4159/9780674020047

Adams, N. P., Scherz, A., & Schmelzle, C. (2019). Legitimacy beyond the state: institutional purposes and contextual constraints. *Critical Review of International Social and Political Philosophy*, *23*(3), 281–291. https://doi.org/10.1080/13698230.2019.1565710

Agarwal, P. K. (2018). Public Administration Challenges in the World of AI and Bots. Public Administration Review, 78(6), 917–921.

Aguinis, H., & Bradley, K. J. (2014). Best practice recommendations for designing and implementing experimental vignette methodology studies. *Organizational Research Methods*, *17*(4), 351–371. https://doi.org/10.1177/1094428114547952

Ahearn, C., Brand, J., & Zhou, X. (2022). How, and For Whom, Does Higher Education Increase Voting? *Research in Higher Education*, *64*, 574–597. https://doi.org/10.1007/s11162-022-09717-4

Akerlof, G. A. (1970). The market for "Lemons": quality uncertainty and the market mechanism. *Quarterly Journal of Economics*, *84*(3), 488. https://doi.org/10.2307/1879431

Alexander, L. (2018, July 2). Do Google's "unprofessional hair" results show it is racist? *The Guardian*. https://www.theguardian.com/technology/2016/apr/08/does-google-unprofessional-hair-results-prove-algorithms-racist-

Alon‐Barkat, S., & Busuioc, M. (2022). Human–AI interactions in public sector decision making: "Automation bias" and "Selective adherence" to algorithmic advice. *Journal of*

Public Administration Research and Theory, *33*(1), 153–169. https://doi.org/10.1093/jopart/muac007

Alva, M. (2021, March 8). *Impact of AI on government decision-makers*. KPMG. https://info.kpmg.us/news-perspectives/technology-innovation/thriving-in-an-ai-world/ai-adoption-government.html

Ansoff, I. (1979). *Strategic Management*. John Wiley & Son.

Araujo, T., Helberger, N., Kruikemeier, S., & De Vreese, C. H. (2020). In AI we trust? Perceptions about automated decision making by artificial intelligence. *AI & SOCIETY*, *35*(3), 611–623. https://doi.org/10.1007/s00146-019-00931-w

Arrow, K. J. (1984). The economics of agency. In *Institute for Mathematical Studies in the Social Sciences, Stanford Univ. eBooks*. https://apps.dtic.mil/sti/pdfs/ADA151436.pdf

Atzmüller, C., & Steiner, P. (2010). Experimental vignette studies in survey research. *Methodology: European Journal of Research Methods for the Behavioral and Social Sciences*, *6*(3), 128–138. https://doi.org/10.1027/1614-2241/a000014

Autoriteit Persoonsgegevens. (2020, July 17). *Werkwijze Belastingdienst in strijd met de wet en discriminerend*. https://autoriteitpersoonsgegevens.nl/actueel/werkwijze-belastingdienst-in-strijd-met-de-wet-en-discriminerend

Baker, A. (2010, May 12). New York Minorities More Likely to Be Frisked. *The New York Times*. https://www.nytimes.com/2010/05/13/nyregion/13frisk.html\

Bannister, F., & Connolly, R. (2014). ICT, public values and transformative government: A framework and programme for research. Government Information Quarterly, 31(1), 119–128.

Beckman, L., Rosenberg, J. H., & Jebari, K. (2022). Artificial intelligence and democratic legitimacy. The problem of publicity in public authority. *AI & Society*. https://doi.org/10.1007/s00146-022-01493-0

Blom-Hansen, J., Baekgaard, M., & Serritzlew, S. (2020). How bureaucrats shape political decisions: The role of policy information. *Public Administration*, *99*(4), 658–678. https://doi.org/10.1111/padm.12709

Bottou, L. (2013). From machine learning to machine reasoning. *Machine Learning*, *94*(2), 133–149. https://doi.org/10.1007/s10994-013-5335-x

Bovens, M., & Zouridis, S. (2002). From Street‐Level to System‐Level Bureaucracies: How Information and Communication Technology is Transforming Administrative Discretion and Constitutional Control. *Public Administration Review*, *62*(2), 174–184. https://doi.org/10.1111/0033-3352.00168

Boyd, D., & Crawford, K. (2012). Critical Questions for Big Data. *Information, Communication & Society*, *15*(5), 662–679. https://doi.org/10.1080/1369118x.2012.678878

Brandt, M. J., Kuppens, T., Spears, R., Andrighetto, L., Autin, F., Babinčák, P., Badea, C., Bae, J., Batruch, A., Becker, J. C., Bocian, K., Bodroža, B., Bourguignon, D., Bukowski, M., Butera, F., Butler, S. E., Chryssochoou, X., Conway, P. M., Crawford, J. T., . . . Zimmerman, J. L. (2020). Subjective status and perceived legitimacy across countries. *European Journal of Social Psychology*, *50*(5), 921–942. https://doi.org/10.1002/ejsp.2694

Brantingham, P. J., Mohler, G., & MacDonald, J. M. (2022). Changes in public–police cooperation following the murder of George Floyd. *PNAS Nexus*, *1*(5). https://doi.org/10.1093/pnasnexus/pgac189

Brauner, P., Hick, A., Philipsen, R., & Ziefle, M. (2023). What does the public think about artificial intelligence?—A criticality map to understand bias in the public perception of AI. *Frontiers in Computer Science*, *5*. https://doi.org/10.3389/fcomp.2023.1113903

Brennan, G., & Lomasky, L. (1993). Democracy and Decision: The Pure Theory of Electoral Preference. In *Cambridge University Press eBooks*. Cambridge University Press. https://doi.org/10.1017/cbo9781139173544

Brückner, M., Ciccone, A., & Tesei, A. (2012). Oil price shocks, income, and democracy. *The Review of Economics and Statistics*, *94*(2), 389–399. https://doi.org/10.1162/rest_a_00201

Brüggen, E., Wetzels, M., De Ruyter, K., & Schillewaert, N. (2011). Individual differences in motivation to participate in online panels. *International Journal of Market Research*, *53*(3), 369–390. https://doi.org/10.2501/ijmr-53-3-369-390

Brundage, M. (2018, February 20). *The Malicious Use of Artificial intelligence: Forecasting, Prevention, and Mitigation*. arXiv.org. https://arxiv.org/abs/1802.07228

Brunson, R. K. (2007). "Police Don't Like Black People": African-American Young Men's Accumulated Police Experiences. Criminology and Public Policy, 6(1), 71–101. https://doi.org/10.1111/j.1745-9133.2007.00423.x

Bryson, J. M., Crosby, B. C., & Bloomberg, L. L. (2014). Public value governance: Moving beyond traditional public administration and the new public management. *Public Administration Review*, *74*(4), 445-456. https://doi.org/10.1111/puar.12238

Buabbas, A. J., Miskin, B., Alnaqi, A., Ayed, A. K., Shehab, A. A., Syed-Abdul, S., & Uddin, M. (2023). Investigating Students' Perceptions towards Artificial Intelligence in Medical Education. *Healthcare*, *11*(9), 1298. https://doi.org/10.3390/healthcare11091298

Buffat, A. (2013). Street-Level bureaucracy and E-Government. *Public Management Review*, *17*(1), 149–161. https://doi.org/10.1080/14719037.2013.771699

Bullock, J., Huang, H., & Kim, K. (2022). Machine intelligence, bureaucracy, and human control. *Perspectives on Public Management and Governance*, *5*(2), 187–196. https://doi.org/10.1093/ppmgov/gvac006

Busch, P. A., & Henriksen, H. Z. (2018). Digital discretion: A systematic literature review of ICT and street-level discretion. *Information Polity*, *23*(1), 3–28. https://doi.org/10.3233/ip-170050

Busuioc, M. (2021). Accountable artificial intelligence: Holding algorithms to account. *Public Administration Review*, *81*(5), 825-836.

Catalog of Bias. (2018, February 23). *Volunteer Bias*. Catalog of Bias. https://catalogofbias.org/biases/volunteer-bias/

Cellan-Jones, B. R. (2014, December 2). Stephen Hawking warns artificial intelligence could end mankind. *BBC News*. https://www.bbc.com/news/technology-30290540

Charles, V., Rana, N. P., & Carter, L. (2022). Artificial Intelligence for data-driven decision making and governance in public affairs. *Government Information Quarterly*, *39*(4), 101742. https://doi.org/10.1016/j.giq.2022.101742

Chowdhury, G. (2003). Natural language processing. *Annual Review of Information Science and Technology*, *37*(1), 51–89. https://doi.org/10.1002/aris.1440370103

Christensen, R. K., John Szmer, and J. M. Stritch. (2012). Race and Gender Bias in Three Administrative Contexts: Impact on Work Assignments in State Supreme Courts. *Journal of Public Administration Research and Theory, 22* (4), 625–48. https://doi.org/10.1093/jopart/mus020

Clark, J. R., & Lee, D. R. (2001). The Optimal Trust in Government. *Eastern Economic Journal*, *27*(1), 19–34. http://www.jstor.org/stable/40326012

Cobb, R., & Elder, C. (1972). *Participation in American Politics*. Boston: Allyn and Bacon.

Cobbe, J. (2019). Administrative law and the machines of government: judicial review of automated public-sector decision making. *Legal Studies*, *39*(4), 636-655.

Common Crawl - Get started. (2023). https://commoncrawl.org/get-started

Compton, M. E., Young, M., Bullock, J., & Greer, R. O. (2022). Administrative errors and race: Can technology mitigate inequitable administrative outcomes? *Journal of Public Administration Research and Theory*, *33*(3), 512–528. https://doi.org/10.1093/jopart/muac036

Concerned Researchers. (2021, December 9). On recent research auditing commercial facial analysis technology. *Medium*. https://medium.com/@bu64dcjrytwitb8/on-recent-research-auditing-commercial-facial-analysis-technology-19148bda1832

Crawford, V. P., & Guasch, J. L. (1983). The theory of contracts and agency. *American Journal of Agricultural Economics*, *65*(2), 373–379. https://doi.org/10.2307/1240895

Davies, D. (2021, January 5). Facial Recognition and Beyond: journalist ventures inside China's "Surveillance State." *NPR*. https://www.npr.org/2021/01/05/953515627/facial-recognition-and-beyond-journalist-ventures-inside-chinas-surveillance-sta

De Boer, N., & Raaphorst, N. (2021). Automation and discretion: explaining the effect of automation on how street-level bureaucrats enforce. *Public Management Review*, *25*(1), 42–62. https://doi.org/10.1080/14719037.2021.1937684

De Fine Licht, K., & De Fine Licht, J. (2020). Artificial intelligence, transparency, and public decision making. *AI & Society*, *35*(4), 917–926. https://doi.org/10.1007/s00146-020-00960-w

Denhardt, R. B. (1999). *Theories of public organization*. Wadsworth Publishing Company.

Devarajan, S. (2022, March 9). *Why hasn't technology improved government effectiveness?* Brookings. https://www.brookings.edu/articles/why-hasnt-technology-improved-government-effectiveness/

Diakopoulos, N. (2014). Algorithmic accountability reporting: On the investigation of black boxes.

DiLeonardo, A., Lauricella, T., & Schaninger, B. (2021, May 10). *Survey fatigue? Blame the leader, not the question*. McKinsey & Company. https://www.mckinsey.com/capabilities/people-and-organizational-performance/our-insights/the-organization-blog/survey-fatigue-blame-the-leader-not-the-question

DiMaggio, P., & Powell, W. W. (1983). The Iron Cage revisited: institutional isomorphism and collective rationality in organizational fields. *American Sociological Review*, *48*(2), 147. https://doi.org/10.2307/2095101

Downs, A. (1957). An Economic Theory of Political Action in a Democracy. *Journal of Political Economy*, *65*(2), 135–150. https://doi.org/10.2307/1827369

Dubois V. (2010). *The bureaucrat and the poor: Encounters in French welfare offices*. Burlington, VT: Ashgate.

Dutilh, G., & Rieskamp, J. (2015). Comparing perceptual and preferential decision making. *Psychonomic Bulletin & Review*, *23*(3), 723–737. https://doi.org/10.3758/s13423-015-0941-1

Dwivedi, Y. K., Hughes, L., Ismagilova, E., Aarts, G., Coombs, C., Crick, T., Duan, Y., Dwivedi, R., Edwards, J. S., Eirug, A., Galanos, V., Ilavarasan, P. V., Janssen, M., Jones, P., Kar, A. K., Kizgin, H., Kronemann, B., Lal, B., Lucini, B., . . . Williams, M. D. (2021). Artificial Intelligence (AI): Multidisciplinary perspectives on emerging challenges, opportunities, and agenda for research, practice and policy. *International Journal of Information Management*, *57*, 101994. https://doi.org/10.1016/j.ijinfomgt.2019.08.002

Dye, T. R. (1998). *Understanding public policy*.

Edwards, L., & Veale, M. (2017). Slave to the Algorithm? Why a Right to Explanation is Probably Not the Remedy You are Looking for. *Duke Law & Technology Review*, *16*(18), 18–84. https://doi.org/10.2139/ssrn.2972855

Epp C. R., Maynard-Moody S., Haider-Markel D. P. (2014). *Pulled over: How police stops define race and citizenship*. Chicago, IL: The University of Chicago Press.

Erman, E., & Furendal, M. (2022). Artificial intelligence and the political legitimacy of global governance. *Political Studies*, 003232172211266. https://doi.org/10.1177/00323217221126665

Eubanks, V. (2018). *Automating inequality: How High-Tech Tools Profile, Police, and Punish the Poor*. St. Martin’s Press.

European Parliament. (2019, July 4). *2019 European election results*. europarl.europa.eu. https://www.europarl.europa.eu/election-results-2019/en

European Parliament. (2023, August 6). *EU AI Act: first regulation on artificial intelligence*. https://www.europarl.europa.eu/news/en/headlines/society/20230601STO93804/eu-ai-act-first-regulation-on-artificial-intelligence

Farrar, J. (2022). *Fairness, legitimacy, and tax compliance*. https://api.semanticscholar.org/CorpusID:248981372

Favier, S., & Wijzenbeek, B. (2023). 2023 NUF Factsheet International Students. In *NUFFIC* (NUF2023/9). https://www.nuffic.nl/sites/default/files/2023-08/2023%20NUF%20factsheet%20international%20students%20EN%20mei%202023.pdf

Federal Bureau of Investigation. (2019). Arrests: Table 43 [Dataset; Webpage]. In *Crime in the U.S. 2019*. Federal Bureau of Investigation: Uniform Crime Data. https://ucr.fbi.gov/crime-in-the-u.s/2019/crime-in-the-u.s.-2019/topic-pages/tables/table-43

Ferguson, A. G. (2017). The rise of big data policing. In *The Rise of Big Data Policing*. New York University Press.

Finkelstein, A. (2014). *Moral hazard in health insurance*. https://doi.org/10.7312/fink16380

Franck, T. M. (1998). Legitimacy and fairness. In *Oxford University Press eBooks* (pp. 25–46). https://doi.org/10.1093/acprof:oso/9780198267850.003.0002

Fritsch, S., Blankenheim, A., Wahl, A., Hetfeld, P., Maaßen, O., Deffge, S., Kunze, J., Rossaint, R., Riedel, M., Marx, G., & Bickenbach, J. (2022). Attitudes and perception of artificial intelligence in healthcare: A cross-sectional survey among patients. *DIGITAL HEALTH*, *8*, 205520762211167. https://doi.org/10.1177/20552076221116772

Gajduschek, G. (2003). Bureaucracy: Is It Efficient? Is It Not? Is That The Question? *Administration & Society*, *34*(6), 700–723. https://doi.org/10.1177/0095399702239171

Galešić, M., & Bošnjak, M. (2009). Effects of questionnaire length on participation and indicators of response quality in a web survey. *Public Opinion Quarterly*, *73*(2), 349–360. https://doi.org/10.1093/poq/nfp031

Gau, J. M., & Brunson, R. K. (2009). Procedural Justice and Order Maintenance Policing: A study of Inner‐City Young men's perceptions of police legitimacy. Justice Quarterly, 27(2), 255–279. https://doi.org/10.1080/07418820902763889

Georgieva, I., Lantta, T., Lickiewicz, J., Pekara, J., Wikman, S., Loseviča, M., Raveesh, B. N., Mihai, A., & Lepping, P. (2021). Perceived Effectiveness, Restrictiveness, and Compliance with Containment Measures against the Covid-19 Pandemic: An International Comparative Study in 11 Countries. *International Journal of Environmental Research and Public Health*, *18*(7), 3806. https://doi.org/10.3390/ijerph18073806

Gerlich, M. (2023). Perceptions and Acceptance of Artificial Intelligence: A Multi-Dimensional Study. *Social Sciences*, *12*(9), 502. https://doi.org/10.3390/socsci12090502

Gilley, B. (2009). *The right to rule: How States Win and Lose Legitimacy*.

Goddard, K., Roudsari, A. V., & Wyatt, J. (2012). Automation bias: a systematic review of frequency, effect mediators, and mitigators. *Journal of the American Medical Informatics Association*, *19*(1), 121–127. https://doi.org/10.1136/amiajnl-2011-000089

Hang, H., & Chen, Z. (2022). How to realize the full potentials of artificial intelligence (AI) in digital economy? A literature review. *Journal of Digital Economy*, *1*(3), 180–191. https://doi.org/10.1016/j.jdec.2022.11.003

Harrits, G. S., & Cecchini, M. (2021). The Professional Agency Narrative—Conceptualizing the role of professional knowledge in frontline work. *Journal of Public Administration Research and Theory*, *32*(1), 41–57. https://doi.org/10.1093/jopart/muab021

Harrits G. S., Møller M. Ø. (2014). Prevention at the front line: How home nurses, pedagogues, and teachers transform public worry into decisions on special efforts. *Public Management Review*, *16*, 447-480.

Hartl, B., Hofmann, E., Gangl, K., Hartner-Tiefenthaler, M., & Kirchler, E. (2015). Does the sole description of a tax authority affect tax evasion? - the impact of described coercive and legitimate power. *PLOS ONE*, *10*(4), e0123355. https://doi.org/10.1371/journal.pone.0123355Brückner, M., Ciccone, A., & Tesei, A. (2012b). Oil price shocks, income, and democracy. *The Review of Economics and Statistics*, *94*(2), 389–399. https://doi.org/10.1162/rest_a_00201

Helsby, J., Carton, S., Joseph, K., Mahmud, A., Park, Y., Navarrete, A., ... & Ghani, R. (2018). Early intervention systems: Predicting adverse interactions between police and the public. *Criminal justice policy review*, *29*(2), 190-209.

Hern, A. (2018, January 12). Google's solution to accidental algorithmic racism: ban gorillas. *The Guardian*. https://www.theguardian.com/technology/2018/jan/12/google-racism-ban-gorilla-black-people

Hidalgo, C. A., Orghian, D., Canals, J. A., De Almeida, F., & Cantero, N. M. (2021). *How humans judge Machines*.

Hinterleitner, M., & Wittwer, S. (2022). Serving quarreling masters: Frontline workers and policy implementation under pressure. *Governance*, *36*(3), 759–778. https://doi.org/10.1111/gove.12692

Hogenhout, L., & Takahashi, T. (2022). A Future with AI. In *Unite.un.org*. The United Nations. https://unite.un.org/sites/unite.un.org/files/a_future_with_ai-final_report.pdf

Holtom, B. C., Baruch, Y., Aguinis, H., & Ballinger, G. A. (2022). Survey response rates: Trends and a validity assessment framework. *Human Relations*, *75*(8), 1560–1584. https://doi.org/10.1177/00187267211070769

How to determine sample size in research. (2022, April 1). Qualtrics.

https://www.qualtrics.com/uk/experience-management/research/determine-sample-size/?rid=ip&prevsite=en&newsite=uk&geo=CZ&geomatch=uk

Hunter, J. D., & Bowman, C. (1996). The State of Disunion: 1996 Survey of American Political Culture. In *Institute for Advanced Studies in Culture* (Summary Report. Charlottesville: The Post-Modernity Project). University of Virginia. https://s3.amazonaws.com/iasc-prod/uploads/pdf/d6c71cdd653aa15c3b8e.pdf

Ingrams, A., Kaufmann, W., & Jacobs, D. (2021). In AI we trust? Citizen perceptions of AI in government decision making. *Policy & Internet*, *14*(2), 390–409. https://doi.org/10.1002/poi3.276\

International IDEA. (2024). *Comparison of Voter Turnout by Region*. idea.int. https://www.idea.int/data-tools/data/region-compare?database_theme=293&question_id=9188&typeA=political_entity®ionA=142733&typeB=political_entity®ionB=142733

International Institute of Communications. (2020). Artificial Intelligence in the Asia-Pacific Region. In *International Institute of Communications*. TRCP. https://www.iicom.org/wp-content/uploads/IIC-AI-Report-2020.pdf

IPSOS. (2022). *Global Opinions and Expectations About Artificial Intelligence: A Global Intelligence Survey*. https://www.ipsos.com/sites/default/files/ct/news/documents/2022-01/Global-opinions-and-expectations-about-AI-2022.pdf

Jackson, J., & Gau, J. M. (2016). Carving up concepts? Differentiating between trust and legitimacy in public attitudes towards legal authority. In *Springer eBooks* (pp. 49–69). https://doi.org/10.1007/978-3-319-22261-5_3

Jackson, J., Huq, A. Z., Bradford, B., & Tyler, T. R. (2012). Going outside the law: the role of the state in shaping attitudes to private acts of violence. *Social Science Research Network*. https://doi.org/10.2139/ssrn.1984957

Jager, J., Putnick, D. L., & Bornstein, M. H. (2017). More Than Just Convenient: The Scientific Merits of Homogenous Convenience Samples. *Monographs of the Society for Research in Child Development*, *82*(2), 13–30. https://doi.org/10.1111/mono.12296

Janssen, M., & Van Der Voort, H. (2016). Adaptive governance: Towards a stable, accountable and responsive government. *Government Information Quarterly*, *33*(1), 1–5. https://doi.org/10.1016/j.giq.2016.02.003

Jasso, G. (2006). Factorial survey methods for studying beliefs and judgments. *Sociological Methods & Research*, *34*(3), 334–423. https://doi.org/10.1177/0049124105283121

Jensen, M. C., & Meckling, W. H. (1976). Theory of the Firm: Managerial Behavior, Agency Costs and Ownership Structure. *Journal of Finance Economics*, *3*(4), 305–360. https://www.sfu.ca/~wainwrig/Econ400/jensen-meckling.pdf

Jeong, Y., & Kim, T. (2019). Between Legitimacy and Efficiency: An Institutional Theory of Corporate giving. *Academy of Management Journal*, *62*(5), 1583–1608. https://doi.org/10.5465/amj.2016.0575

John, P. (1998). *Analysing public policy*. A&C Black.

Johnson, C., Dowd, T. J., & Ridgeway, C. L. (2006). Legitimacy as a social process. *Annual Review of Sociology*, *32*(1), 53–78. https://doi.org/10.1146/annurev.soc.32.061604.123101

Jorna, F., & Wagenaar, F. (2007). The Iron Cage Strengthened: Discretion and Digital Discipline. *Public Administration*, *85*(1), 189–214. https://doi.org/10.1111/j.1467-9299.2007.00640.x

Kappmeier, M., & Fahey, K. (2022). Trust and Legitimacy: Policing Among Racial Groups. *New Zealand Journal of Psychology*, *51*(2), 35–42. https://www.psychology.org.nz/application/files/8616/6631/4272/Kappmeier_35-42.pdf

Katzenbach, C., & Ulbricht, L. (2019). Algorithmic governance. *Internet Policy Review*, *8*(4). https://doi.org/10.14763/2019.4.1424

Kieslich, K., Lünich, M., & Marcinkowski, F. (2021). The threats of Artificial Intelligence Scale (TAI). *International Journal of Social Robotics*, *13*(7), 1563–1577. https://doi.org/10.1007/s12369-020-00734-w

Klette, R. (2014). *Concise Computer Vision: An Introduction into Theory and Algorithms*. Springer Science & Business Media.

Kolodner, J. L. (1992). An introduction to case-based reasoning. *Artificial Intelligence Review*, *6*(1), 3–34. https://doi.org/10.1007/bf00155578

Koteen, J. (1997). *Strategic management in public and nonprofit organizations: Managing Public Concerns in an Era of Limits*. Praeger.

Latham, G. P., & Locke, E. A. (2002). Building a practically useful theory of goal setting and task motivation: A 35-year odyssey. *American Psychologist, 57*(9), 705–717. https://doi.org/10.1037/0003-066X.57.9.705

Lawler, E. L. (1971). Pay and organization effectiveness: A psychological view. *NY: McGraw Hill*.

Lee, P. (2018, November 26). The US Army is developing unmanned drones that can decide who to kill. *Business Insider*. https://www.businessinsider.com/the-us-army-is-developing-unmanned-drones-that-can-decide-who-to-kill-2018-4?international=true&r=US&IR=T

Leiden University. (2022). *Facts and figures*. https://www.universiteitleiden.nl/en/about-us/facts-and-figures

Levitov, A. (2016). Normative legitimacy and the state. In *Oxford University Press eBooks*. https://doi.org/10.1093/oxfordhb/9780199935307.013.131

Linders, S., & Peters, B. G. (1990). An institutional approach to the theory of policy-making: the role of guidance mechanisms in policy formulation. *Journal of Theoretical Politics, 2*, 59-83.

Lipsky, Michael. (1980). *Street Level Bureaucracy: Dilemmas of the Individual in Public Services*. Russell Sage Foundation. http://www.jstor.org/stable/10.7758/9781610447713

Lipsky, Michael (1969). Toward a Theory of Street-Level Bureaucracy (IRP Discussion Papers No. 48-69) (p. 45). Madison, WI: Institute for Research on Poverty (IRP), University of Wisconsin. http://www.irp.wisc.edu/publications/dps/pdfs/dp4869.pdf

Lipsky, A. (2011). Evaluating the strength of faith: Potential comparative advantages of faith-based organizations providing health services in sub-Saharan Africa. *Public Administration and Development*, *31*(1), 25–36. https://doi.org/10.1002/pad.586

Ma, M., & Agarwal, R. (2007). Through a Glass Darkly: Information technology design, identity verification, and knowledge contribution in online communities. *Information Systems Research*, *18*(1), 42–67. https://doi.org/10.1287/isre.1070.0113

March, J. G., & Olsen, J. P. (1989). Rediscovering institutions: The organizational basis of politics. New York: Free Press.

Martin, K., & Waldman, A. E. (2022). Are algorithmic decisions legitimate? The Effect of process and outcomes on perceptions of legitimacy of AI decisions. *Journal of Business Ethics*, *183*(3), 653–670. https://doi.org/10.1007/s10551-021-05032-7

Maynard-Moody S., Musheno M. (2003). *Cops, Teachers, Counselors: Stories from the Front Lines of Public Service.* The University of Michigan Press. https://doi.org/10.3998/mpub.11924

Melamed, D. (2012). The effects of legitimacy and power on perceptions of fairness. *Sociological Focus*, *45*(2), 125–142. https://doi.org/10.1080/00380237.2012.657119

McCorduck P. (2004). *Machines who think. A personal inquiry into the history and prospects of artificial intelligence*. Taylor & Francis.

McDonald, P. (2019). How factorial survey analysis Improves our understanding of employer preferences. *Schweizerische Zeitschrift Für Soziologie*, *45*(2), 237–260. https://doi.org/10.2478/sjs-2019-0011

McGaughey, E. (2021). Will robots automate your job away? Full employment, basic income and economic democracy. *Industrial Law Journal*, *51*(3), 511–559. https://doi.org/10.1093/indlaw/dwab010

McGregor, D., & Cutcher-Gershenfeld, J. (2006). *The Human Side of Enterprise*. McGraw Hill Professional.

McKinsey. (2020, November 17). *The state of AI in 2020*. McKinsey & Company. https://www.mckinsey.com/capabilities/quantumblack/our-insights/global-survey-the-state-of-ai-in-2020\

Medaglia, R., Gil-García, J. R., & Pardo, T. A. (2021). Artificial intelligence in government: Taking stock and moving forward. *Social Science Computer Review*, *41*(1), 123–140. https://doi.org/10.1177/08944393211034087

Mennerick L. A. (1974). Client typologies: A method of coping with conflict in the service worker-client relationship. *Work and Occupations, 1*, 396-418.

Mergel, I., Edelmann, N., & Haug, N. (2019). Defining digital transformation: Results from expert interviews. *Government Information Quarterly*, *36*(4), 101385. https://doi.org/10.1016/j.giq.2019.06.002

Miller, S. M., & Keiser, L. R. (2020). Representative bureaucracy and attitudes toward automated decision making. *Journal of Public Administration Research and Theory*, *31*(1), 150–165. https://doi.org/10.1093/jopart/muaa019

Moore, G. E. & Intel.Com. (1965). Cramming more components onto integrated circuits. *Electronics Magazine*. http://cva.stanford.edu/classes/cs99s/papers/moore-crammingmorecomponents.pdf

Morin, R., & Stepler, R. (2023, May 10). *The Racial Confidence Gap in Police Performance.* Pew Research Center's Social & Demographic Trends Project. https://www.pewresearch.org/social-trends/2016/09/29/the-racial-confidence-gap-in-police-performance/

Nair, C. S., Adams, P., & Mertova, P. (2008). Student engagement: the key to improving survey response rates. *Quality in Higher Education*, *14*(3), 225–232. https://doi.org/10.1080/13538320802507505

National Institute of Justice. (2013, January 9). *Race, Trust and Police Legitimacy*. National Institute of Justice. https://nij.ojp.gov/topics/articles/race-trust-and-police-legitimacy

Nicas, J. (2018, February 7). How YouTube drives people to the internet's darkest corners. *WSJ*. https://www.wsj.com/articles/how-youtube-drives-viewers-to-the-internets-darkest-corners-1518020478

Niklasson, B., Christiansen, P., & Öhberg, P. (2020). Speaking truth to power: political advisers' and civil servants' responses to perceived harmful policy proposals. *Journal of Public Policy*, *40*(3), 492–512. https://doi.org/10.1017/s0143814x18000508

Nikolopoulou, K. (2023, March 24). *What is social desirability bias? | Definition & Examples*. Scribbr. https://www.scribbr.com/research-bias/social-desirability-bias/

Noble, S. U. (2018). *Algorithms of Oppression: How Search Engines Reinforce Racism*. New York: NYU Press.

Nordesjö, K., Ulmestig, R., & Denvall, V. (2020). Coping with tensions between standardization and individualization in social assistance. *Nordic Social Work Research*, *12*(4), 435–449. https://doi.org/10.1080/2156857x.2020.1835696

Nye, J. S., Jr, Júnior, J. S. N., Zelikow, P. D., & King, D. C. (1997). *Why people don't trust government*. Harvard University Press.

OECD Survey on Drivers of Trust in Public Institutions. (2021). In *OECD*. OECD. https://www.oecd-ilibrary.org/sites/b407f99c-en/1/2/1/index.html?itemId=/content/publication/b407f99c-en&_csp_=c12e05718c887e57d9519eb8c987718b&itemIGO=oecd&itemContentType=book

Office of Institutional Research. (2023, February 28). *Race/Ethnicity FAQs.* Office of Institutional Research. https://provost.tufts.edu/institutionalresearch/race-ethnicity-faq/

Organization For Economic Co-Operation And Development. (2004). *Compliance Risk Management: Managing and Improving Tax Compliance*. OECD Publishing. https://www.oecd.org/tax/administration/33818656.pdf

Organization For Economic Co-Operation And Development. (2005). *Performance-related Pay Policies for Government Employees* [Print]. OECD Publishing.

Oxford English Dictionary. (n.d.) Citation. In *OED.com dictionary.* Retrieve November 20, 2023, from https://www.oed.com/dictionary/fair_adj?tab=meaning_and_use#4657049.

Pachucki, M. C., & Breiger, R. L. (2010). Cultural holes: beyond relationality in social networks and culture. *Annual Review of Sociology*, *36*(1), 205–224. https://doi.org/10.1146/annurev.soc.012809.102615

Parasuraman, R., & Manzey, D. (2010). Complacency and Bias in Human Use of Automation: an attentional integration. *Human Factors*, *52*(3), 381–410. https://doi.org/10.1177/0018720810376055

Payton, K. S., & Gould, J. B. (2022). Vignette Research Methodology: an essential tool for quality improvement collaboratives. *Healthcare*, *11*(1), 7. https://doi.org/10.3390/healthcare11010007

Pierre, J., & Peters, B. G. (2017). The shirking bureaucrat: a theory in search of evidence? *Policy and Politics*, *45*(2), 157–172. https://doi.org/10.1332/030557317x14845830916703

Phelps, Edmund S. (1972). The Statistical Theory of Racism and Sexism. *The American Economic Review, 62*(4), 659–661.

Plant, J. F. (2011). Carl J. Friedrich on Responsibility and Authority. *Public Administration Review*, *71*(3), 471–482. http://www.jstor.org/stable/23017504

Pratt, J. W., & Zeckhauser, R. (1985). *Principals and agents: The Structure of Business*. Harvard Business Review Press.

Priddy K. L., Keller P. E. (2005). *Artificial neural networks: An introduction*. SPIE Press.

Prottas J. M. (1980). People-processing: The street-level bureaucrat in public service bureaucracies. *Contemporary Sociology, 9*, 715. https://doi.org/10.2307/2065531

Punch, M., Van Der Vijver, K., & Zoomer, O. (2004). Dutch “COP”: Developing Community Policing in the Netherlands. In *Office of Justice Programs* (No. 195131). U.S. Department of Justice. https://ojp.gov/ncjrs/virtual-library/abstracts/dutch-cop-developing-community-policing-netherlands

Qualtrics. (2021, September 27). *How to ask sensitive survey questions*. Qualtrics. https://www.qualtrics.com/blog/how-to-get-the-truth-when-asking-survey-questions-about-sensitive-topics/

Rainey, H. G., Fernandez, S., & Malatesta, D. (2021). Analyzing the Environment of Public Organizations. *Understanding and Managing Public Organizations,* 77–98. John Wiley & Sons.

Rainey, H. G., & Steinbauer, P. (1999). Galloping Elephants: Developing Elements of a Theory of Effective Government Organizations. *Journal of Public Administration Research and Theory*, *9*(1), 1–32. https://doi.org/10.1093/oxfordjournals.jpart.a024401

Revilla, M., & Höhne, J. K. (2020). How long do respondents think online surveys should be? New evidence from two online panels in Germany. *International Journal of Market Research*, *62*(5), 538–545. https://doi.org/10.1177/1470785320943049

Rice, D. R., & Zorn, C. (2021). Corpus-based dictionaries for sentiment analysis of specialized vocabularies. *Political Science Research and Methods*, *9*(1), 20-35.

Risse‐Kappen, T., & Stollenwerk, E. (2018). Legitimacy in areas of limited statehood. *Annual Review of Political Science*, *21*(1), 403–418. https://doi.org/10.1146/annurev-polisci-041916-023610

Ritchie, H., Mathieu, E., Roser, M., & Ortiz-Ospina, E. (2023, April 13). *Internet*. Our World in Data. https://ourworldindata.org/internet

Roig-Franzia, M. (2020, December 2). Miles Taylor spoke out against Trump as ‘Anonymous.’ Now he’s gone public and is hiding out. *Washington Post*. https://www.washingtonpost.com/lifestyle/style/miles-taylor-anonymous-trump-profile/2020/11/30/ece1d5c0-28df-11eb-92b7-6ef17b3fe3b4_story.html

Rosenfeld, B., Imai, K., & Shapiro, J. N. (2015). An empirical validation study of popular survey methodologies for sensitive questions. *American Journal of Political Science*, *60*(3), 783–802. https://doi.org/10.1111/ajps.12205

Roth, B. R. (2000). Legal Legitimacy and Recognition of Governments: A Doctrinal guide. In *Oxford University Press eBooks* (pp. 121–199). https://doi.org/10.1093/acprof:oso/9780199243013.003.0005

Rothkopf, D. (2022). *American Resistance: The Inside Story of How the Deep State Saved the Nation*. Hachette UK.

Rubington, E., & Weinberg, M. S. (2011). *The study of social Problems: Seven Perspectives*. Oxford University Press, USA.

Russell, S., & Norvig, P. (2019). *Artificial intelligence: A Modern Approach*. Pearson Higher Education.

Samuelson, R. J. (1995). *The good life and its discontents: The American Dream in the Age of Entitlement, 1945-1995*. Crown.

Sartori, L., & Bocca, G. (2022). Minding the gap(s): public perceptions of AI and socio-technical imaginaries. *AI & SOCIETY*, *38*(2), 443–458. https://doi.org/10.1007/s00146-022-01422-1

Savage, N. (2022). Breaking into the black box of artificial intelligence. *Nature*. https://doi.org/10.1038/d41586-022-00858-1

Scanlon, J. M., Kusano, K. D., Daniel, T., Alderson, C., Ogle, A., & Victor, T. (2021). Waymo simulated driving behavior in reconstructed fatal crashes within an autonomous vehicle operating domain. *Accident Analysis & Prevention*, *163*, 106454. https://doi.org/10.1016/j.aap.2021.106454

Scharpf, FW (1999) Governing in Europe. Oxford: Oxford University Press.

Schmidt, VA (2013) Democracy and legitimacy in the European Union revisited: input, output and "throughput." Political Studies 61, 2–22. https://doi.org/10.1111/j.1467-9248.2012.00962.x

Schmidt, VA and Wood, M (2019) Conceptualizing throughput legitimacy: procedural mechanisms of accountability, transparency, inclusiveness and openness in EU governance. Public Administration, padm.12615. 97, 727–740. https://doi.org/10.1111/padm.12615

Schoon, E. W. (2022). Operationalizing legitimacy. *American Sociological Review*, *87*(3), 478–503. https://doi.org/10.1177/00031224221081379

Scottish Government. (2015). *Characteristics of recent and established EEA and Non-EEA migrants in Scotland: Analysis of the 2011 Census* (No. 9781785441974). https://www.gov.scot/publications/characteristics-recent-established-eea-non-eea-migrants-scotland-analysis-2011-census/pages/3/

Sells, K. (2023, September 20). *AI's Growing Influence in Asia Pacific*. https://www.linkedin.com/pulse/ais-growing-influence-asia-pacific-kass-sells/

Shafiq, M., Gu, Z., Cheikhrouhou, O., Alhakami, W., & Hamam, H. (2022). The Rise of "Internet of Things": Review and open research issues related to detection and prevention of IoT-Based security attacks. *Wireless Communications and Mobile Computing*, *2022*, 1–12. https://doi.org/10.1155/2022/8669348

Shafritz, J. M., & Hyde, A. C. (1992). *Classics of public administration*. Thomson Brooks/Cole.

Simon, H. A. (2013). *Administrative Behavior, 4th Edition*. Simon and Schuster.

Skelly, A. C., Dettori, J. R., & Brodt, E. D. (2012). Assessing bias: the importance of considering confounding. *Evidence-based Spine-care Journal*, *3*(01), 9–12. https://doi.org/10.1055/s-0031-1298595

Smola A., Vishwanathan S. V. N. (2008). *Introduction to machine learning*. Cambridge University.

Sparks, J., & Bottoms, A. (1995). Legitimacy and order in prisons. *The British Journal of Sociology*, *46*(1), 45. https://doi.org/10.2307/591622

Sparrow, M. K. (2015). Measuring Performance in a Modern Police Organization. *Office of Justice Programs*, 1–40. https://www.ojp.gov/pdffiles1/nij/248476.pdf

Spence, M. (2022). Automation, Augmentation, Value Creation & the Distribution of Income & Wealth. *Daedalus*, *151*(2), 244–255. https://doi.org/10.1162/daed_a_01913

Starke, C., & Lünich, M. (2020). Artificial intelligence for political decision making in the European Union: Effects on citizens' perceptions of input, throughput, and output legitimacy. *Data & Policy*, *2*. https://doi.org/10.1017/dap.2020.19

Statistics Netherlands. (2022, March 18). CBS introducing new population classification by origin. *Statistics Netherlands*. https://www.cbs.nl/en-gb/news/2022/07/cbs-introducing-new-population-classification-by-origin

Stone, D. A. (2011). *Policy paradox: the art of political decision making*. New York, NY: W.W. Norton.

Stillman, P. G. (1974). The concept of legitimacy. *Polity*, *7*(1), 32–56. https://doi.org/10.2307/3234268

Su, D., & Steiner, P. (2018). An evaluation of experimental designs for constructing vignette sets in factorial surveys. *Sociological Methods & Research*, *49*(2), 455–497. https://doi.org/10.1177/0049124117746427

Subramony, M., & Pugh, S. D. (2015). Services management research: Review, integration, and future directions. *Journal of Management, 41*(1), 349–373. https://doi.org/10.1177/0149206314557158

Subramony, M. (2017). Service Organizations and Their Communities: Perspectives and New Directions for Management Research. *Academy of Management Perspectives*, *31*(1), 28–43. https://doi.org/10.5465/amp.2015.0125

Suchman, M. C. (1995). Managing Legitimacy: strategic and institutional approaches. *Academy of Management Review*, *20*(3), 571. https://doi.org/10.2307/258788

Sun, T. Q., & Medaglia, R. (2019). Mapping the challenges of Artificial Intelligence in the public sector: Evidence from public healthcare. *Government Information Quarterly*, *36*(2), 368–383. https://doi.org/10.1016/j.giq.2018.09.008

SurveyMonkey. (2023). *Demographic Survey questions: Why & how to ask*.SurveyMonkey. https://www.surveymonkey.com/mp/gathering-demographic-information-from-surveys/#:~:text=Place%20age%20range%20and%20other,the%20end%20of%20the%20survey.

Svara, J. H. (2001). The Myth of the Dichotomy: Complementarity of Politics and Administration in the Past and Future of Public Administration. *Public Administration Review*, *61*(2), 176–183. https://doi.org/10.1111/0033-3352.00020

Taylor, B. J. (2005). Factorial Surveys: Using vignettes to study professional judgement. *British Journal of Social Work*, *36*(7), 1187–1207. https://doi.org/10.1093/bjsw/bch345

Taylor, M (2018, October 28). Opinion | I Am Part of the Resistance Inside the Trump Administration. *The New York Times*. https://www.nytimes.com/2018/09/05/opinion/trump-white-house-anonymous-resistance.html

Thesaurus. (2023). Fairness. In *Thesaurus*. https://www.thesaurus.com/browse/fairness

Thrun, S. (2010). Toward robotic cars. *Communications of the ACM*, *53*(4), 99–106. https://doi.org/10.1145/1721654.1721679

Times Higher Education (THE). (2023a, September 25). *University of Amsterdam*. https://www.timeshighereducation.com/world-university-rankings/university-amsterdam

Times Higher Education (THE). (2023b, November 16). *Leiden University*. https://www.timeshighereducation.com/world-university-rankings/leiden-university

Toshkov, D. (2016). *Research design in political science*. Red Globe Press.

Tourangeau, R., & Yan, T. (2007). Sensitive questions in surveys. *Psychological Bulletin*, *133*(5), 859–883. https://doi.org/10.1037/0033-2909.133.5.859

Tummers, L., Bekkers, V., Vink, E., & Musheno, M. (2015). Coping during Public Service Delivery: A conceptualization and Systematic Review of the literature. *Journal of Public Administration Research and Theory*, *25*(4), 1099–1126. https://doi.org/10.1093/jopart/muu056

Tyler, T. (2009). Legitimacy and criminal justice: The benefits of self-regulation. *Ohio State Journal of Criminal Law*, *7*, 307–359. https://openyls.law.yale.edu/bitstream/handle/20.500.13051/2412/Legitimacy_and_criminal_justice.pdf?sequence=2&isAllowed=y

Tyler, T. (2011). Trust and legitimacy: Policing in the USA and Europe. *European Journal of Criminology*, *8*(4), 254–266. https://doi.org/10.1177/1477370811411462

Universiteit van Amsterdam. (2022, April 20). *Facts and figures*. University of Amsterdam. https://www.uva.nl/en/about-the-uva/about-the-university/facts-and-figures/facts-and-figures.html

Vaish, A., Großmann, T., & Woodward, A. L. (2008). Not all emotions are created equal: The negativity bias in social-emotional development. *Psychological Bulletin*, *134*(3), 383–403. https://doi.org/10.1037/0033-2909.134.3.383

Vedung, E. (2015). Autonomy and street-level bureaucrats' coping strategies. *Nordic Journal of Studies in Educational Policy*, (2), 28643. https://doi.org/10.3402/nstep.v1.28643

Viola, P., & Jones, M. J. (2004). Robust real-time face detection. *International journal of computer vision*, *57*, 137-154.

Von Garrel, J., & Mayer, J. (2023). Artificial Intelligence in studies—use of ChatGPT and AI-based tools among students in Germany. *Humanities and Social Sciences Communications*, *10*(1). https://doi.org/10.1057/s41599-023-02304-7

Vroom, V. H. (1964). Work and motivation.

Weber, M. (1985). Gesammelte Aufsätze zur Wissenschaftslehre, ed. Johannes Winckelmann. *Tübingen, Germany: JCB Mohr [Paul Siebeck]*. https://verlagsarchivweb.ub.uni-potsdam.de/494/html/WL.pdf

Weber, M. (1919). *Politics as a vocation*. https://archive.org/details/weber_max_1864_1920_politics_as_a_vocation/page/n1/mode/2up

Wenger, Jeffrey B., and Vicky M. Wilkins. (2009). At the Discretion of Rogue Agents: How Automation Improves Women's Outcomes in Unemployment Insurance. *Journal of Public Administration Research and Theory, 19* (2), 313–33. https://doi.org/10.1093/jopart/mum044

Wooldridge M. (2009). *An introduction to multiagent systems*. John Wiley & Sons.

Wortley, S., Hagan, J., & Macmillan, R. (1997). Just Des(s)erts? The Racial Polarization of Perceptions of Criminal Injustice. Law & Society Review, 31(4), 637. https://doi.org/10.2307/3053983

Wu, M., Zhao, K., & Fils-Aime, F. (2022). Response rates of online surveys in published research: A meta-analysis. *Computers in Human Behavior Reports*, *7*, 100206. https://doi.org/10.1016/j.chbr.2022.100206

Young, M., Bullock, J., & Lecy, J. D. (2019). Artificial Discretion as a Tool of Governance: A framework for understanding the impact of artificial intelligence on public administration. *Perspectives on Public Management and Governance*. https://doi.org/10.1093/ppmgov/gvz014

Zacka, B. (2017). *When the state meets the street*. https://doi.org/10.4159/9780674981423

Zerilli, J. (2021). *A Citizen's Guide to Artificial Intelligence*. MIT Press.

Zouridis, S., Van Eck, M., & Bovens, M. (2019). Automated Discretion. In *Discretion and the quest for controlled freedom* (pp. 313–329). Palgrave Macmillan.

Appendix A. Survey

Start of Block: Baseline

S1: Informed Consent:

This survey is anonymous and the only data collected are the answers you provide. The information is used solely for academic purposes. Your participation is voluntary and you can withdraw your consent at any time. Thank you for your contribution!

Q1: Do you voluntarily consent to taking this survey?

o Yes (1)

o No (2)

Q2: Do you agree with the following statement describing your relationship to the government?:

"I consent to the existence of the government and obey the law so that the government can protect basic rights and promote the common good of society through its many services."

o Yes (1)

o No (2)

Q3: Do you believe the following scenario is problematic?:

"A government agent makes a risky decision because they do not bear the cost of this decision."

o Yes (1)

o No (2)

Q4: Do you believe the following scenario is problematic?:

"A government agent knows more than citizens and can use this advantage in decision making to exert biases or shirk responsibilities."

o Yes (1)

o No (2)

Q5: Do you believe the following scenario is problematic?:

"A government agent is asked to pursue multiple goals that are unclear or conflict with each other."

o Yes (1)

o No (2)

Q6: How much do you agree with the following statement?:

"I expect a police officer to make desirable, proper, or appropriate decisions."

o Strongly Disagree (1)

o Disagree (2)

o Neither agree nor disagree (3)

o Agree (4)

o Strongly agree (5)

Q7: How much do you agree with the following statement?:

"I expect a police officer to make efficient decisions."

o Strongly Disagree (1)

o Disagree (2)

- o Neither agree nor disagree (3)
- o Agree (4)
- o Strongly agree (5)

Q8: How much do you agree with the following statement?:

"I expect a police officer to make fair decisions."

- o Strongly Disagree (1)
- o Disagree (2)
- o Neither agree nor disagree (3)
- o Agree (4)
- o Strongly agree (5)

Q9: How much do you agree with the following statement?:

"I expect a tax auditor to make desirable, proper, or appropriate decisions."

- o Strongly Disagree (1)
- o Disagree (2)
- o Neither agree nor disagree (3)
- o Agree (4)
- o Strongly agree (5)

Q10: How much do you agree with the following statement?:

"I expect a tax auditor to make efficient decisions."

- o Strongly Disagree (1)

- o Disagree (2)
- o Neither agree nor disagree (3)
- o Agree (4)
- o Strongly agree (5)

Q11: How much do you agree with the following statement?:

"I expect a tax auditor to make fair decisions."

- o Strongly Disagree (1)
- o Disagree (2)
- o Neither agree nor disagree (3)
- o Agree (4)
- o Strongly agree (5)

Q12: How much do you agree with the following statement?:

"I expect a health inspector to make desirable, proper, or appropriate decisions."

- o Strongly Disagree (1)
- o Disagree (2)
- o Neither agree nor disagree (3)
- o Agree (4)
- o Strongly agree (5)

Q13:How much do you agree with the following statement?:

"I expect a health inspector to make efficient decisions."

- o Strongly Disagree (1)
- o Disagree (2)
- o Neither agree nor disagree (3)
- o Agree (4)
- o Strongly agree (5)

Q14: How much do you agree with the following statement?:

"I expect a health inspector to make fair decisions."

- o Strongly Disagree (1)
- o Disagree (2)
- o Neither agree nor disagree (3)
- o Agree (4)
- o Strongly agree (5)

End of Block: Baseline

Randomization Protocol: Block (NO AI) or Block (AI).

Start of Block: NO AI

S2: Review the following statement:

"A person has three alcoholic drinks in the evening before driving to a friend's house. A police officer spots the vehicle and runs the person's license plate. The officer discovers that the person has been arrested in the past for driving under the influence of alcohol. The police officer has no current evidence to suggest that the person has been drinking and driving. Under the pretext of a routine traffic stop, the officer pulls the person over and questions them about the amount of alcoholic drinks they have had. The person admits to drinking earlier that evening. The police officer arrests the person for driving under the influence."

Q15: How much do you agree with the following statement?:

"The police officer makes a desirable, proper, or appropriate decision in this scenario."

o Strongly Disagree (1)

o Disagree (2)

o Neither agree nor disagree (3)

o Agree (4)

o Strongly agree (5)

Q16: How much do you agree with the following statement?:

"The police officer makes an efficient decision in this scenario."

o Strongly Disagree (1)

o Disagree (2)

o Neither agree nor disagree (3)

o Agree (4)

o Strongly agree (5)

Q17: How much do you agree with the following statement?:

"The police officer makes a fair decision in this scenario."

o Strongly Disagree (1)

o Disagree (2)

o Neither agree nor disagree (3)

o Agree (4)

o Strongly agree (5)

S3: Review the following statement:

"The federal government issues a goal to tax auditors to reduce the total amount of fraudulent tax filings. A managing supervisor at a federal tax agency issues a goal to tax auditors to process tax forms as fast as possible to reduce waiting times. A person with a history of tax-fraud has filed their income incorrectly. As a result, the person paid less taxes than they are legally required to. A tax auditor determines that following up on this case will significantly slow down the processing time of tax forms. In an attempt to satisfy their supervisor's goal, the tax auditor decides not to issue the person a warning or fine."

Q18: How much do you agree with the following statement?:

"The tax auditor makes a desirable, proper, or appropriate decision in this scenario."

o Strongly Disagree (1)

o Disagree (2)

o Neither agree nor disagree (3)

o Agree (4)

o Strongly agree (5)

Q19: How much do you agree with the following statement?:

"The tax auditor makes an efficient decision in this scenario."

o Strongly Disagree (1)

o Disagree (2)

o Neither agree nor disagree (3)

o Agree (4)

o Strongly agree (5)

Q20:How much do you agree with the following statement?:

"The tax auditor makes a fair decision in this scenario."

o Strongly Disagree (1)

o Disagree (2)

o Neither agree nor disagree (3)

o Agree (4)

o Strongly agree (5)

S4: Review the following statement:

"A health inspector performs a scheduled inspection on a restaurant. The health inspector is concerned about the unhygienic way that the food is being handled in the restaurant's kitchen. The owner of the restaurant ensures the health inspector that the issues will be corrected. The health inspector does not expect to eat at the restaurant themselves. Based on the type of restaurant, the owner, and other factors, the health inspector determines that dining at the restaurant is unlikely to result in food-poisoning. Despite the restaurant's minor infractions of the health code, the health inspector decides to approve the restaurant's health inspection. One week later, several people who dine at the restaurant are diagnosed with severe food poisoning."

Q21: How much do you agree with the following statement?:

"The health inspector makes a desirable, proper, or appropriate decision in this scenario."

o Strongly Disagree (1)

o Disagree (2)

o Neither agree nor disagree (3)

o Agree (4)

o Strongly agree (5)

Q22: How much do you agree with the following statement?:

"The health inspector makes an efficient decision in this scenario."

o Strongly Disagree (1)

o Disagree (2)

o Neither agree nor disagree (3)

o Agree (4)

o Strongly agree (5)

Q23: How much do you agree with the following statement?:

"The health inspector makes a fair decision in this scenario."

o Strongly Disagree (1)

o Disagree (2)

- o Neither agree nor disagree (3)
- o Agree (4)
- o Strongly agree (5)

End of Block: NO AI

Start of Block: AI

S5: Review the following statement:

"A person has three alcoholic drinks in the evening before driving to a friend’s house. An AI-operated CCTV camera spots the vehicle and runs the person’s license plate. The person has been arrested in the past for driving under the influence of alcohol. The information is relayed to a nearby police officer. The police officer has no current evidence to suggest that the person has been drinking and driving. Under the pretext of a routine traffic stop, the officer pulls the person over and questions them about the amount of alcoholic drinks they have had. The person admits to drinking earlier that evening. The police officer arrests the person for driving under the influence."

Q24: How much do you agree with the following statement?:

"The police officer makes a desirable, proper, or appropriate decision in this scenario."

- o Strongly Disagree (1)
- o Disagree (2)
- o Neither agree nor disagree (3)
- o Agree (4)

o Strongly agree (5)

Q25: How much do you agree with the following statement?:

"The police officer makes an efficient decision in this scenario."

o Strongly Disagree (1)

o Disagree (2)

o Neither agree nor disagree (3)

o Agree (4)

o Strongly agree (5)

Q26: How much do you agree with the following statement?:

"The police officer makes a fair decision in this scenario."

o Strongly Disagree (1)

o Disagree (2)

o Neither agree nor disagree (3)

o Agree (4)

o Strongly agree (5)

S6: Review the following statement:

"The federal government issues a goal to tax auditors to reduce the total amount of fraudulent tax filings. A managing supervisor at a federal tax agency issues a goal to tax auditors to process tax forms as fast as possible to reduce waiting times. A person with a history of tax-fraud has filed their income incorrectly. As a result, the person paid less taxes than they are legally required to. A tax auditor uses a machine-learning algorithm that employs AI technology to determine that following up on this case will significantly slow down the processing time of tax forms. In an attempt to satisfy their supervisor's goal, the tax auditor decides not to issue the person a warning or fine."

Q27: How much do you agree with the following statement?:

"The tax auditor makes a desirable, proper, or appropriate decision in this scenario."

o Strongly Disagree (1)

o Disagree (2)

o Neither agree nor disagree (3)

o Agree (4)

o Strongly agree (5)

Q28: How much do you agree with the following statement?:

"The tax auditor makes an efficient decision in this scenario."

o Strongly Disagree (1)

o Disagree (2)

o Neither agree nor disagree (3)

o Agree (4)

o Strongly agree (5)

Q29: How much do you agree with the following statement?:

"The tax auditor makes a fair decision in this scenario."

- o Strongly Disagree (1)
- o Disagree (2)
- o Neither agree nor disagree (3)
- o Agree (4)
- o Strongly agree (5)

S7: Review the following statement:

"A health inspector performs a scheduled inspection on a restaurant. The health inspector is concerned about the unhygienic way that the food is being handled in the restaurant's kitchen. The owner of the restaurant ensures the health inspector that the issues will be corrected. The health inspector does not expect to eat at the restaurant themselves. The health inspector uses an AI-powered risk-assessment model to determine that dining at the restaurant is unlikely to result in food-poisoning. Despite the restaurant's minor infractions of the health code, the health inspector decides to approve the restaurant's health inspection. One week later, several people who dine at the restaurant are diagnosed with severe food poisoning."

Q30: How much do you agree with the following statement?:

"The health inspector makes a desirable, proper, or appropriate decision in this scenario."

- o Strongly Disagree (1)
- o Disagree (2)
- o Neither agree nor disagree (3)
- o Agree (4)
- o Strongly agree (5)

Q31: How much do you agree with the following statement?:

"The health inspector makes an efficient decision in this scenario."

- o Strongly Disagree (1)
- o Disagree (2)
- o Neither agree nor disagree (3)
- o Agree (4)
- o Strongly agree (5)

Q32: How much do you agree with the following statement?:

"The health inspector makes a fair decision in this scenario."

- o Strongly Disagree (1)
- o Disagree (2)
- o Neither agree nor disagree (3)
- o Agree (4)
- o Strongly agree (5)

End of Block: AI

Start of Block: Confounder Block

Q33: How much do you agree with the following statement?:

"I am familiar with AI technology."

- o Strongly disagree (1)

- o Disagree (2)
- o Neither agree nor disagree (3)
- o Agree (4)
- o Strongly agree (5)

Q34: How much do you agree with the following statement?:

"I have a favorable opinion of AI technology."

- o Strongly disagree (1)
- o Disagree (2)
- o Neither agree nor disagree (3)
- o Agree (4)
- o Strongly agree (5)

End of Block: Confounder Block

Start of Block: Demographic Block

Q35: Please select "Green" below:

- o Red (1)
- o Blue (2)
- o Green (3)

Q36: What gender do you identify as?

- o Male (1)

- o Female (2)
- o Other (3)
- o Prefer not to say (4)

Q37: How old are you?

- o 0 - 15 years old (1)
- o 15 - 30 years old (2)
- o 30 - 45 years old (3)
- o 45 + years old (4)

Q38: What is the highest degree or level of education that you have completed?

- o High School (1)
- o Bachelor's Degree (2)
- o Master's Degree (3)
- o Ph.D or Higher (4)
- o Trade School (5)

Q39: What is your ethnicity?

- o Caucasian (1)
- o Black (2)
- o Hispanic (3)
- o Asian (4)
- o Two or More (5)

o Other (6)

Q40: What country are you a citizen of? (Select one option)

▼ Afghanistan (1) ... Other (1358)

Q41: Are you employed?

o Employed Full Time (1)

o Employed Part Time (2)

o Retired (3)

o Unemployed (4)

Q42: Are you registered to vote?

o Yes (1)

o No (2)

Q43: Will you vote in the upcoming election?

o Yes (1)

o No (2)

End of Block: Demographic Block

Appendix B. Pilot Test

Q1: Were you briefed on informed consent in the survey?

o Yes (1)

o No (2)

Q2: I understand the questions asked in the survey:

o Strongly Disagree (1)

o Disagree (2)

o Neither agree nor disagree (3)

o Agree (4)

o Strongly agree (5)

Q3: The order of the questions in the survey makes sense:

o Strongly Disagree (1)

o Disagree (2)

o Neither agree nor disagree (3)

o Agree (4)

o Strongly agree (5)

Q4: The questions asked in the survey are offensive:

o Strongly Disagree (1)

o Disagree (2)

o Neither agree nor disagree (3)

- o Agree (4)
- o Strongly agree (5)

Q5: The survey takes the right amount of time to complete:

- o Strongly Disagree (1)
- o Disagree (2)
- o Neither agree nor disagree (3)
- o Agree (4)
- o Strongly agree (5)

Q6: The survey feels professional:

- o Strongly Disagree (1)
- o Disagree (2)
- o Neither agree nor disagree (3)
- o Agree (4)
- o Strongly agree (5)

Q7: Are there any technical issues with the survey?

- o Yes (1) ____
- o No (2)