Received: 19 February 2024 | Revised: 31 July 2024 | Accepted: 8 August 2024
DOI: 10.1111/padm.13030

ORIGINAL ARTICLE

public administration
WILEY

Algorithmic formalization: Impacts on administrative processes

Antonio Cordella | Francesco Gualdi

London School of Economics and Political Science, London, UK

Correspondence
Francesco Gualdi, London School of Economics and Political Science, Houghton Street, WC2A 2AE London, UK.
Email: f.gualdi@lse.ac.uk

Abstract

This paper investigates the influence of algorithms on the administrative processes within public organizations, utilizing the foundational theory of formalization from Walsh and Dewar (1987) as a framework. Introduces the concept of "algorithmic formalization", a new form of formalization induced by the adoption of algorithms, which fundamentally alters administrative workflows. Focusing on COMPAS algorithm used in the US judiciary for risk assessment, the paper illustrates how the algorithm serves multiple roles – as code, channel, and standard – systematizing administrative processes related to risk assessment and judicial decisions. By delving into COMPAS case study, the research sheds light on the novel concept of algorithmic formalization, emphasizing its significant repercussions for analyzing and applying algorithmic administrative processes.

Questa ricerca discute l'impatto degli algoritmi sui processi amministrativi delle organizzazioni pubbliche, utilizzando la teoria della formalizzazione di Walsh e Dewar (1987) come framework. In questa ricerca introduciamo il concetto di "formalizzazione algoritmica": un nuovo tipo di formalizzazione indotta dall'adozione degli algoritmi, che altera radicalmente i flussi di lavoro nella pubblica amministrazione. Studiando l'algoritmo COMPAS, utilizzato nel settore giudiziario statunitense per la valutazione del rischio, la ricerca illustra come l'algoritmo svolga molteplici ruoli – codice, canale e standard – sistematizzando i processi amministrativi relativi alla valutazione del rischio e alle decisioni giudiziarie. Lo studio del caso COMPAS permette di introdurre il concetto innovativo di "formalizzazione algoritmica", sottolineando le ripercussioni significative di tale concetto nell'analisi e nell'applicazione dei processi amministrativi algoritmici.

1 | INTRODUCTION

The study of administrative processes within the domain of public administration has produced significant insights into the critical functions these processes serve in the generation and delivery of public services (Fayol, 1949; Gulick, 1937). Scholarly investigations offer a thorough analysis of the impact of these processes on the operational efficiency and overall effectiveness of public sector organizations.

One of the core findings within this body of research is the demonstrable positive association between administrative processes standardization and organizational efficiency (Lim & Tang, 2008). The implementation of standardized processes has been shown to rationalize operations, ensuring the efficient utilization of resources and the effective delivery of public services (Denhardt & Denhardt, 2000; Hood, 1991). Standardized administrative processes are instrumental in the reduction of superfluous activities and the strategic allocation of resources, thereby amplifying the general efficiency of public service provision (Hughes, 2017; Pollitt & Bouckaert, 2017). Standardization contributes to achieve effectiveness in the processes it mediates (Scholta et al., 2020): for example, the adoption of standardized technological processes enables the public sector to allocate more resources on the design and delivery of public services, which leads to a more functional and effective service provision (Fishenden & Thompson, 2012).

However, standardization of administrative processes, while streamlining procedures and improving efficiency and effectiveness, can also have several negative impacts (Lee, 2024): for instance, standardization can lead to increased costs due to rigid adherence to standardized protocols which may not always align with local needs or citizens-specific requirements (Kwon, 2008). Moreover, tensions and challenges can arise when organizations attempt to standardize processes, where the loss of contextual flexibility can impede responsiveness and innovation (Brunsson et al., 2012) and reduce the effectiveness of administrative responses to local or tailored needs (Waugh & Streib, 2006).

Public administration transparency and accountability are also significantly impacted by administrative processes standardization. Research underscores the importance of standardization in facilitating scrutiny and enabling the holding of public officials to account (Hood, 1991), but it also emphasizes the paradoxical effects that lead to reduced democratic engagement and lessened accountability due to an overemphasis on efficiency and output measures (Christensen & Lægreid, 2002).

Also, the literature consistently highlights the importance of standardized administrative processes to ensure equitable treatment and impartiality across all the interactions between public administrations and citizens (Cordella, 2007; Guy & McCandless, 2012).

Overall, the body of research on administrative processes in public administration provides compelling evidence of the importance of studying the way in which administrative processes are standardized to fully appreciate the impacts they have on the operation of public sector organizations.

In recent years, algorithmic systems have been largely adopted by public sector organizations to standardize and automate administrative processes (Ammitzbøll Flügge et al., 2021; Medaglia et al., 2021; Meijer et al., 2021) and to increase their efficiency (de Sousa et al., 2022). Also, algorithms have been deployed to enhance the extensiveness and enforcement of the mechanism governing administrative processes, framing administrative processes into the algorithmic code (Wenzelburger et al., 2024). However, these very valuable contributions have so far devoted only limited attention to the way in which the algorithms' specific technological characteristics transform the administrative processes they are expected to inform. In this paper, we contribute to the study of administrative processes shedding light on how algorithmic systems shape these processes by framing their interactions and the mechanisms by which they produce, process, and analyze information. Drawing from Walsh and Dewar (1987), we refer to this framing as formalization of administrative processes.

In the conceptual framework of Walsh and Dewar (1987), formalization of administrative processes is intricately defined through three dimensions: code, channel, and standard. Code simplifies complex organizational activities into manageable formulae, streamlining processes and making them easier to implement. Channel reduces variability in human performance by establishing clear, predefined communication pathways, ensuring efficient information flow. Standard sets performance benchmarks, providing a basis for evaluating actions and determining appropriate rewards or punishments. Together, these elements of formalization work synergistically to systematize administrative process, enabling organizations to operate with greater clarity, consistency, and accountability.

Considering the growing significance of algorithms in framing interactions by structuring information analysis and processing, this paper poses the following research question:

How do algorithmic systems formalize administrative processes by redefining their code, channel, and standard?

The paper addresses this question introducing "algorithmic formalization" as a distinctive form of formalization within administrative processes. Our contribution enhances the literature by detailing how algorithms manage data structuring and aggregation to render data computable and how algorithmic computation introduces new forms of code, channels, and standards in formalizing administrative processes.[1]

To illustrate these concepts, the paper analyses the algorithmic system utilized in the United States judiciary to formalize the administrative process related to the assessment of offenders' likelihood to recidivate. The case study offers new insights into the impacts of algorithmic systems on administrative processes. By doing so, the paper contributes to the existing debate concerned with the increasing "algorithmization" of public sector (Wenzelburger et al., 2024) that triggers further questions related to multifaced impacts of algorithmic adoptions within public administration (McDonald et al., 2022).

2 | ADMINISTRATIVE PROCESSES FORMALIZATION

Formalization guiding the systematic execution of administrative processes within organizations is a pivotal concept in organizational theory and public administration. Formalization roots in Weber's work (Weber, 1922) on bureaucracies' reliance on structured hierarchies, rules, and procedures to govern their operations. This concept, which has evolved from its theoretical roots to adapt to contemporary technological advancements and changes in organizational environments, remains a fundamental aspect of public administration studies (Hansen, 2022).

Weber (1922) laid the groundwork for understanding administrative processes formalization. He described bureaucracy as inherently structured, standardized, and impersonal, characterized by a clear hierarchical structure, strict rules and regulations, extensive division of labour, and decision-making grounded in rational-legal authority. This depiction of bureaucracy essentially frames administrative processes formalization as the development and maintenance of an organized, rule-based approach to management, complete with explicit coordination mechanisms, comprehensive rules, and meticulous documentation. Administrative processes formalization results in the standardization of rules and procedures (Rainey & Bozeman, 2000) setting explicit parameters that govern specific jobs or tasks (Dalton et al., 1980).

2.1 | Administrative processes formalization and algorithms

The advent of AI, driven by the continually evolving frontier of algorithmic computation (Berente et al., 2021; McCorduck, 2004), necessitates new understandings of how digital technologies impact the formalization of administrative processes. Unlike traditional digital technologies, algorithmic computation presents unique influences on administrative processes formalization abstracting the execution of the administrative process into the algorithmic logic. These new dynamics warrant further academic exploration to fully comprehend their implications on government structures and administrative functions (Van der Voort et al., 2019).

Recent literature has discussed the transformative impact of algorithms on administrative processes. Contributions in these fields are mostly geared towards algorithms as instrumental to enhance administrative processes efficiency and responsiveness. Integrating algorithms in public administration makes it possible to automate administrative processes, enhancing decision-making and public service delivery efficiency (Margetts & John, 2023; Pencheva et al., 2020; Young et al., 2019). However, while administrative processes standardization can enhance efficiency (Schiff et al., 2021) and decision-making quality (Criado et al., 2020; Ingrams, 2020), it can also have negative effects (Agarwal, 2018; Sun & Medaglia, 2019; Wirtz et al., 2020). This includes reducing trustworthiness (Grimmelikhuijsen, 2022) and compromising accountability (Busuioc, 2021; Gualdi & Cordella, 2024). Additionally, algorithms reshape administrative processes while simultaneously being shaped by them: this interplay constraints how administrative processes are executed, considerably impacting discretion in public organizations (Bullock et al., 2020).

In summary, the academic discourse underscores that the integration of algorithms into administrative processes is not just a technological upgrade but a significant reconfiguration of the mechanisms underlining the execution of these processes within organizations.

3 | THEORETICAL FRAMEWORK: ADMINISTRATIVE PROCESSES FORMALIZATION

To study the impact of algorithms on administrative processes formalization we build on the three dimensions identified by Walsh and Dewar (1987): code, channel, and standard. These dimensions explain how formalization frames administrative processes by structuring information analysis and processing. Therefore, formalization proves useful in studying how the specific mechanisms by which algorithms standardize information analysis and processing impact administrative processes.

As a code, formalization reduces complex information analysis and processing actions into simple rules, easing communication and coordination. For example, a formalized procedure encapsulates a set of information analysis and processing actions without needing elaborate explanations making their execution easier across different organizational departments.

As a channel, formalization directs and restricts information analysis and processing to create predictable outcomes and behaviors. It reduces variability in human performance by defining appropriate behaviors within specific contexts, delineating the rules of information analysis and processing, and determining the interests that the administrative process seeks to pursue and those it does not.

Lastly, as a standard, formalization sets benchmarks for assessing the actions related to information analysis and processing within the organization. It determines what actions are correct and what are not, thus reducing disputes and facilitating easier coordination among different organization departments. Table 1 offers a synthesis of the definitions utilized, as per Walsh and Dewar (1987).

TABLE 1 Definitions of key dimensions of Code, Channel, and Standard (Walsh & Dewar, 1987).

Dimension	Definition
Code	Formalization as code reduces a complex set of activities to fewer complex formulae.
Channel	Formalization as channel decreases variance in human performance.
Standard	Formalization establishes the standard against which action is compared and punishments are provided.

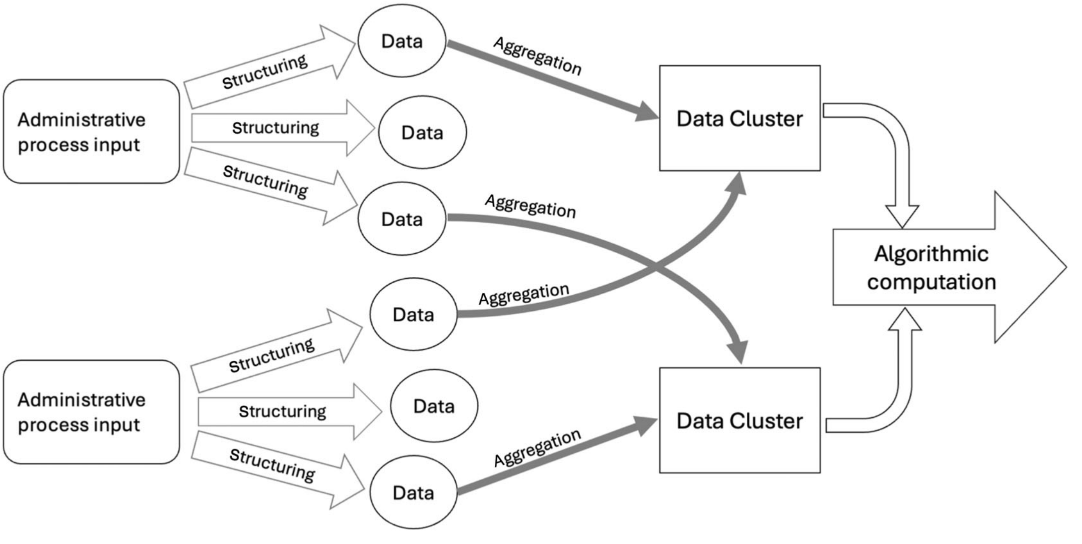

FIGURE 1 Data structuring and aggregation from administrative inputs to algorithmic computation.

3.1 | Algorithmic computing and administrative processes formalization

Algorithms, by analyzing and processing information based on predefined criteria, provide organizations, notably in the public sector, with innovative means to formalize administrative processes (Brynjolfsson & McAfee, 2014). To enable this, algorithms require data to be in format compatible with their computational logic, ensuring that the data can be effectively computed and output be accurately produced. For data to be compatible with algorithmic processing, they must be transformed and structured appropriately (Gillespie et al., 2014). This transformation process decontextualizes and standardizes the data, organizing them into distinct categories or clusters, thereby preparing them for algorithmic computations.

Accordingly, the labeling of decontextualized and standardized data allows for their aggregation into coherent groups or clusters – a process known as data aggregation (Alaimo & Kallinikos, 2017). This process is instrumental in linking various data elements through the algorithmic code, thus forming a coherent data structure. Figure 1 explains how the process of data structuring and aggregation takes place.

Data structuring and aggregation shape the mechanisms by which formalization acts as code, channel and standard, influencing administrative practices and processes (Walsh & Dewar, 1987).

The formalization process is integral in shaping the framework within which algorithmically processed data inform and structure administrative tasks and responsibilities. Figure 2 illustrates how the process of data structuring and aggregation influences the formalization of administrative processes into what we call algorithmic formalization.

Algorithmic formalization refers to the process by which algorithms structure and standardize administrative tasks and procedures through the systematic analysis and processing of data. This involves data structuring, which

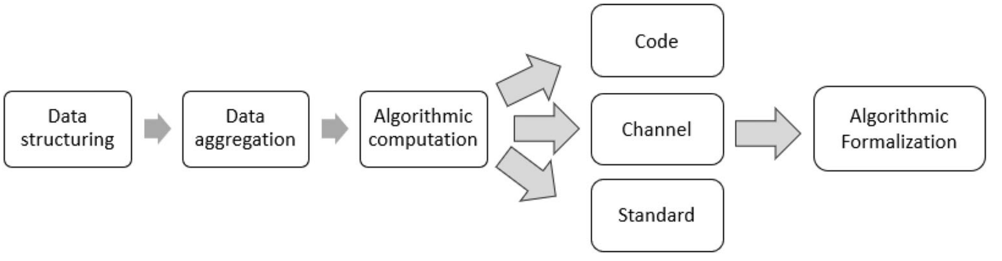

FIGURE 2 Phases of algorithmic formalization through data structuring and aggregation.

organizes raw data into a format suitable for algorithmic processing, and data aggregation, which combines this structured data into coherent groups or clusters. Data structuring, data aggregation, and algorithmic computation convert complex administrative processes into simplified, executable rules (codes), directing the flow of information to ensure consistency and predictability (channels), and setting benchmarks for evaluating performance and compliance (standards). By doing so, algorithmic formalization transforms traditional human-driven administrative processes into automated, rule-based systems. This transformation impacts on administrative performance but also imposes a rigid framework that may limit flexibility and adaptability. Algorithmic formalization thus reshapes how administrative functions are executed, embedding computational logic into the very fabric of organizational operations and influencing how decisions are made, actions are performed, and outcomes are evaluated.

Algorithmic formalization relies on structured and aggregated datasets, employing statistical techniques to infer unknown relationships and generate new, previously unknown outputs. This complex algorithmic computation redefines the codes, channel and standard at the core of administrative processes formalization. Indeed, every data input that is structured and aggregated represents specific categories within the algorithm's framework (Alaimo & Kallinikos, 2017). These categories are crucial for the algorithm's computational needs, influencing the formalization of public sector administrative processes and reflecting on values such as equality, fairness, and impartiality (Bouckaert et al., 2016) the code, channel, and standard carry over.

Understanding the impact of these algorithmically influenced formalizations on administrative processes and interdependencies are vital for comprehending the implications of the adoption of algorithmic systems, such as AI, for public administrations. This includes understanding how data processed by the algorithm are structured and aggregated as well as the intrinsic functions embedded within the algorithmic code that governs this processing. The practices of data structuring and aggregation presume that specific and contextual individual conditions are stripped away by data so to have minimal impact on the algorithmic computation (van Leijen, 2005). This premise rests on the negation of contextual, local, and individual differences, rendering structured and aggregated data ostensibly objective (Au, 2022). Such decontextualised data are reconstituted into formats that cater to the computational requisites of specific algorithms, thereby altering the very fabric of the information analysis and processing that frame administrative interactions and guide the systematic execution of administrative processes.

Within public sector organizations, algorithmic processes structure and aggregate data in line with planned problem-solving logics, imposed by predefined rules and systematic computational operations embedded into the algorithmic code (Conte & De Boor, 2018). This shift delegates tasks, traditionally within the purview of human agents, to algorithmic systems, thereby reshaping fundamental organizational practices (von Krogh, 2018). Furthermore, the governing rules of these algorithmic systems entrench specific patterns of action aligned with the designed code, thereby shaping how administrative processes operate and reconfiguring the distribution of control and authority within organizational functions (Martin, 2019). This transformation of administrative processes into algorithm-driven data processing redesigns the code, channel, and standard which constitute the administrative processes formalization.

To illustrate this argument, this paper analyses and compares how changes in the way in which data are structured and aggregated, and algorithmically processed, impact on administrative processes formalization. Building on Walsh and Dewar's (1987) framework, we posit that the formalization of administrative processes carried over by algorithms impacts the code, channel, and standard which govern these processes. To assess the validity of this assumption, we develop three main propositions that we are going to test in the remainder of the paper.

Proposition 1. *Algorithmic formalization, as a code, reduces complex rules and procedures into simpler formulae, redesigning administrative processes.*

Proposition 2. *Algorithmic formalization, as a channel, automates the production of outputs in the administrative processes, reducing variance in human performance.*

Proposition 3. *Algorithmic formalization, as a standard, contributes to establishing the benchmarks against which action is compared. By analyzing and processing information based on predefined criteria, algorithms minimize the variability often associated with human judgment, ensuring uniformity in administrative actions.*

4 | METHODOLOGY

To explain our argument, the study adopts a qualitative case study method (Yin, 2018). We align with Ospina et al. (2018), who advocate for a rigorous reporting of key methodological decisions in qualitative research in the domain of public administration. In this section, we provide a careful justification for the case study selection.

The case study focuses on the Correctional Offender Management Profiling for Alternative Sanctions (COMPAS) system: the algorithmic system deployed in the United States' judiciary to enhance the quality of risk offenders' assessment. Through the analysis of COMPAS, the paper aims to illustrate how algorithms impact administrative processes formalization. The COMPAS case study presents two challenges, one about novelty, and one about data collection.

A robust literature has discussed the adoption of COMPAS in judicial field, with a renewed and increased interest (Humerick, 2019; Stevenson & Doleac, 2022). However, on the one hand, scholars have mostly paid attention to ethical dimensions, such as fairness (Ávila et al., 2020), discriminations (Hamilton, 2019), and bias (Kirkpatrick, 2017). On the other hand, law scholars have discussed the challenges generated by COMPAS for the actors in the judicial sector (Washington, 2018), or in relation to the trade secret laws utilized by private companies to protect the algorithms they produce (Chander, 2017). Ethical and legal issues have received increasing attention over recent years, and for these reasons, the paper focuses on a different dimension, that is, the impact of COMPAS algorithms on the formalization of administrative processes. COMPAS can provide relevant insights into how the functionalities of the algorithm change the administrative interactions, thereby guiding the systematic execution of administrative processes and their formalization.

The second challenge is about the data collection: we acknowledge that direct access to the algorithm is not possible due to property rights. However, we believe that the secondary data collected offer a thorough illustration of how the algorithm formalizes the public administration information analysis and processing. With the exception of the well-known investigative journalistic work by ProPublica (Larson et al., 2016), research has not delved into the specific characteristics of the technological artifact that constitutes the COMPAS algorithm. Using secondary data, this research delves into the algorithm to illustrate how its functionalities impact administrative processes formalization. Data collection includes secondary sources: reports about COMPAS and COMPAS practitioners' guides released by the private firm that has designed COMPAS, databases with COMPAS raw data made public, judicial sentences, public inquiries, and other gray literature.

The explanatory case study method (Yin, 2018) is chosen because of the revelatory approach of this research (Baxter & Jack, 2012). We have no influence or control over the events that happened, and the focus of the investigation is on past events (Yin, 2018). Explanatory case study is a relevant method to explain how "conditions came to be" (Yin, 2018, p. 238): through the selected case, we aim to illustrate how the algorithmic formalization impacts administrative processes. Following the explanatory case study method, the research question was formulated after a thorough analysis of the academic literature that identified existing gaps (Yin, 2018). The literature review and theoretical background also allowed us to approach some preliminary theoretical propositions associated with the research's focus. In explanatory case study research, it is paramount to develop theoretical propositions before the data collection (Yin, 2018). The three theoretical propositions outlined above will be tested against the case study of COMPAS.

Analyzing the core functionalities of an algorithm in a case study can be approached systematically using methods corroborated in literature. We began the analysis examining existing technical documentation, as suggested by Kitchin (2019), who emphasizes the importance of understanding the conceptual and operational framework of an algorithm. We then analyzed the algorithm's structure breaking down the algorithm as recommended by Cormen et al. (2009) to understand the algorithm's inputs, processes, and outputs. The available source code was overviewed to understand the coding logic behind algorithms (Witten et al., 2011). The data were used to study the practical applications of the algorithm (O'Neil, 2017) examining its real-world implications.

5 | CASE STUDY

Before the widespread adoption of sophisticated digital tools, felons' risk assessment was an administrative process delegated to officers (such as correctional staff and clinical professionals) who executed tasks to produce individualized risk assessments (Turner et al., 2013). Administrative processes to determine risk assessment included actions such as data collection, interviews with felons, data analysis based on static factors, and decisions on sentencing for felons (Kehl et al., 2017). Throughout all stages of the administrative process for risk assessment, officers and professionals primarily depended on their own professional judgment (Kehl et al., 2017).

The US judicial system has increasingly adopted digital tools to assist professionals in the determination of risk assessment. One of the most widely utilized is COMPAS algorithm, a fourth-generation risk assessment AI tool developed by a private business company, Northpointe Inc.,[2] to statistically assess "many of the key risk and need factors in adult correctional populations and to provide information" (Equivant, 2019, p. 2) to guide decisions on constrictive measures. The algorithm predicts the likelihood a felon will reoffend on the basis of standardized risk factors that scientific literature has identified as those most relevant to predict recidivism – that is, "the ability to discriminate between offenders who will and will not recidivate" (Equivant, 2019, p. 7). Over the years, Equivant has continuously refined COMPAS algorithm, and the structure and aggregation of the data processed by the algorithm (risks scales) to improve the accurateness of the automated administrative processes used to make the predictions. Equivant first developed COMPAS Core tool and subsequently deployed the updated version COMPAS-R. The two algorithms rely on different scales which structure data differently in aggregated risk factors to be processed by the algorithms to automatically compute recidivism risk. We will first analyze each version of COMPAS and then discuss how they differently impact on administrative processes formalization.

5.1 | COMPAS Core

COMPAS Core utilizes a combination of scales clustered in Need and Risk scales that are algorithmically processed to produce risk scores used to inform supervision decisions.

The construction of these scales is a crucial process where data are structured and aggregated to make them computable by COMPAS algorithm. These scales have been constructed by gathering data from over 30,000 offenders to select a normative group of 7381 individuals, representative of varied gender, ethnic, and correctional backgrounds. This normative group completed a survey, the responses to which were structured into numerical values and subsequently aggregated into risk scale scores.

Each scale used different datasets. For example, the "Criminal Involvement scale" is based on four questions regarding the offender's past criminal activities and sentences: (a) How many times has the offender been arrested before as an adult or juvenile; (b) How many times has this person been sentenced to jail for 30 days or more; (c) How many times has this person been sentenced (new commitment) to state or federal prison (include current); (d) How many times has this person been sentenced to probation as an adult (Northpointe, 2009, p. 9). Cut-off points are: 1–4 low; 5–7 medium; 8–10 high. Table 2 shows how the Criminal Involvement Scale is constructed and what type of answer is needed.

In contrast, scales that aim to elucidate complex aspects of an offender's behavior, encompassing historical, social, economic, and cultural dimensions, necessitate a more comprehensive dataset. Consequently, such scales involve administering a larger number of questions to offenders to capture this multifaceted information. For instance, the "Vocational/Educational Problems" scale delves into areas such as the offender's educational background, employment history, and acquired skills. As indicated in Table 3, to gather these extensive data, the survey includes 12 questions, employing various response formats including binary (Yes/No), rating scales, and closed-ended options (Northpointe, 2009, p. 17). Cut-off points are: 1–5 low; 6–7 medium; 8–10 high.

The normative group is ranked based on the scores obtained in the different scales. Scores are ranked from lowest to highest values. Subsequently, an algorithm is employed to segment this ranked group into deciles of equal size. This process is crucial for creating clusters represented in the deciles needed to benchmark new offenders against the normative group.

TABLE 2 Questions for the Criminal Involvement Scale – *Source*: Authors' elaboration adapting from Northpointe (2009).

Survey questions	Answers	Type of answer
How many times has the offender been arrested before as an adult or juvenile (criminal arrests only)?		Numerical value
How many times has this person been sentenced to jail for 30 days or more?	0 1 2 3 4 5+	Closed-ended
How many times has this person been sentenced (new commitment) to state or federal prison (include current)?	0 1 2 3 4 5+	Closed-ended
How many times has this person been sentenced to probation as an adult?	0 1 2 3 4 5+	Closed-ended

TABLE 3 Questions for the Vocational/Educational Problems Scale – *Source*: Authors' elaboration adapting from Northpointe (2009).

Survey question	Answers	Type of answer
Did you complete your high school diploma or GED?	Yes No	Binary
What was your final grade completed in school?		Numerical value
What were your usual grades in high school?	A B C D E/F Did Not Attend	Rating
Were you ever suspended or expelled from school?	Yes No	Binary
Did you fail or repeat a grade level?	Yes No	Binary
Do you have a job?	Yes No	Binary
Do you currently have a skill, trade or profession at which you usually find work?	Yes No	Binary
Can you verify your employer or school (if attending)?	Yes No	Binary
How much have you worked or been enrolled in school in the last 12 months?	12 Months Full Time 12 Months Part Time 6+ Months FT 0 to 6 Months PT/FT	Closed-ended
Right now, do you feel you need more training in a new job or career skill?	Yes No	Binary
Right now, if you were to get (or have) a good job, how would you rate your chance of being successful?	Good Fair Poor	Closed-ended
How hard is it for you to find a job ABOVE minimum wage compared to others?	Easier Same Harder Much Harder	Closed-ended

5.1.1 | The composition of deciles within COMPAS Core scales

The deciles in the COMPAS Core scales are then algorithmically computed to categorize offenders into three primary scales' risk clusters: low risk (encompassing the first to fourth deciles), medium risk (fifth to seventh deciles), and high risk (eighth to tenth deciles).

The algorithmic demarcation of decile boundaries in the normative group determines the risk clusters boundaries. Offenders' risk is benchmarked against the risk clusters to determine their risk classification. This process establishes a direct correlation between the individual scores of an offender and the scores of the normative group, structuring into rigid clusters risk assessment.

This aggregation creates a systematic, albeit somewhat artificial, classification system for evaluating offenders' risks, grounding the behavior of the normative group in a standardized context to facilitate the construction of these risk deciles. This presupposes a behavioral consistency in each offender with the algorithmically computed cluster within the normative group.

5.1.2 | The assessment of offender's risk of recidivism with COMPAS Core

Different risk assessments integrate results from COMPAS Core scales using multiple predictive modeling approaches. However, Northpointe has only revealed how the "General Recidivism Risk Scale" is computed using the results of COMPAS Core scales.

The "General Recidivism Risk Scale", designed to predict new offenses within 2 years of assessment, incorporates factors like prior criminal history, criminal associates, drug involvement, and early signs of juvenile delinquency (Northpointe, 2015, p. 27). While the correlation of COMPAS Core scales in this risk algorithm is known, Northpointe has not made public the weights assigned to each scale. However, it is revealed that the algorithm, protected by property rights, uses a linear regression model to calculate the "General Recidivism Risk Score". The formula includes variables such as age, age at first arrest, criminal involvement, vocational education, and drug history, each multiplied by undisclosed weights:

$$\begin{aligned}\text{General Recidivism Risk Score} = {} & (\text{age} * -w) + (\text{age-at-first-arrest} * -w) + (\text{criminal involvement} * w) \\ & + (\text{vocation education} * w) + (\text{drug history} * w)\end{aligned}$$

The "General Recidivism Risk Score" algorithmically processes data formatted into the COMPAS Core scales to determine the offender's risk score. This score is then benchmarked against the normative group's decile distributions. The algorithm computes these data to produce a final assessment to inform judicial decisions. This process is pivotal in channeling administrative processes, as delineated in Northpointe's (2015) report, because individuals scoring above the high-risk threshold in the individual scales are earmarked for intensified treatment programs.

5.2 | COMPAS-R

In 2022, Equivant introduced COMPAS-R, a shorter and more streamlined version of the original COMPAS Core. This new version was developed to provide a quicker and simpler tool for assessing recidivism risk. COMPAS-R reorganized risk factors, adjusting the basic (non-predictive) scales, and focusing on a singular predictive tool: the "General Recidivism Risk" algorithm. This algorithm assesses the likelihood of recidivism based on the results of basic risk scales, with an extended timeframe for recidivism prediction from 2 to 3 years. Additionally, COMPAS-R provides a more detailed analysis of felons' profiles, offering gender-specific predictions.

To create these new scales, Equivant modified COMPAS Core scales using different datasets. Also, scales in COMPAS-R are standardized against different normative groups. For instance, the "Cognitive Behavioral" scale includes 35 risk factors and is based on a normative group of 4314 individuals, while the "Current Violence" scale, with seven risk factors, uses a normative group of 16,011 individuals.

COMPAS-R uses different scales, which requires structuring and aggregating data differently compared to COMPAS Core. For example, the "Legal System Involvement Scale" in COMPAS-R primarily focuses on the criminal history of the individuals being assessed. Four items make up this scale and contribute to the overall score; three of these items are carried over from the COMPAS Core "Criminal Involvement" scale, with one notable modification: the revised version inquires whether the person has ever been sentenced to jail for 30 days or more, differing from the original question in the COMPAS Core scale which asked about the number of times an offender had been arrested. Table 4 shows the composition of the scale, with items, survey questions, answers, and whether the items have been carried over from COMPAS Core. Despite the changes leading to a reduction in the scale's reliability (from 0.782 to 0.714), Equivant considers it to be operating within the "acceptable range of reliability" (Equivant, 2022, p. 16).

Moreover, the "Drug Problems Scale" is constructed by merging three items (use of drugs at time of arrest; benefit from treatment; ever being in treatment) from the "Substance Use scale" in COMPAS Core and adding two new items (current trafficking and current possession). Table 5 shows the items used to build the scale, survey questions, answers, and whether the items have been carried over from COMPAS Core.

Also, "Vocational and Educational Scale" is an update of the scale used in COMPAS Core (Equivant, 2022, p. 116). The scale builds upon questions on individuals' levels of education, skills, and job career. Eight items define the scale and concur to the provision of a score. Compared to the original COMPAS Core scale, three items have been removed (grades in high school; failing or repeating a grade; expulsion from school). Table 6 shows the items used to build the scale, survey questions, answers, and whether the items have been carried over from COMPAS Core.

TABLE 4 Questions for the COMPAS-R Legal System Involvement Scale – *Source*: Authors' elaboration adapting from Equivant (2022).

Item of the scale	Item code	Survey questions	Answers	Present in COMPAS Core
Jail sentence	jail30_R	Has this person ever been sentenced to jail for 30 days or more as an adult (exclude current)?	No = 0 Yes = 1	No – modified to ask whether the person has been sentenced, not how many times
Arrest as an adult	n_prison_R	How many times has this person been sentenced (new commitment) to a state or federal prison as an adult (exclude current)?	0 = 0 1 = 1 2 = 2 3 = 3 4 = 4 5+ = 5	Yes
Probation as an adult	n_probations_R	How many prior times has this person been sentenced to probation as an adult (exclude current)?	0 = 0 1 = 1 2 = 2 3 = 3 4 = 4 5+ = 5	Yes
Prior arrest	t_prev_arrests_R	How many prior times has this person been arrested as an adult or juvenile (including for possession of small amounts of marijuana)?	0 times = 0 1 time = 1 2/3 times = 2 4/5 times = 3 6+ times = 4	Yes – added the marijuana issue

TABLE 5 Questions for the COMPAS-R Drug Problems Scale – *Source*: Authors' elaboration adapting from Equivant (2022).

Item of the scale	Item code	Survey question	Answers	Present in COMPAS Core
Substance use at the time of offense	ad_arrest_R	Were you using alcohol or drugs (including opioids) at the time of the current offense?	No = 0 Yes, alcohol only = 0 Yes, drugs only = 1 Yes, both = 1	Yes – merged alcohol and drugs
Need for treatment	benefit_rx_ad_R	Do you think you would benefit from treatment for alcohol, or drugs, or both?	No = 0 Yes, alcohol only = 0 Yes, drugs only = 1 Yes, both = 1	Yes - merged alcohol and drugs
Drug possession charge	currdrg_poss_R	Is the current charge drug possession?	Not checked in Current Charges table = 0 Checked in Current Charges table = 1	No
Drug trafficking charge	currdrg_traf_R	Is the current charge drug trafficking?	Not checked in Current Charges table = 0 Checked in Current Charges table = 1	No
Treatment records	ever_rx_ad_R	Have you ever been in formal treatment for alcohol or drugs, such as counseling, outpatient, inpatient, or resident?	No = 0 Yes, alcohol only = 0 Yes, drugs only = 1 Yes, both = 1	Yes - merged alcohol and drugs
Opioids charge	op_arrest_R	1 s the current offense opioid related (were opioids involved at the time of offense or arrest)?	No Yes Not scored	No

5.2.1 | The assessment of offender's risk of recidivism with COMPAS-R

COMPAS-R uses a single predictive scale: the "Summative General Recidivism Risk Scale" (Summative GRRS) which introduces significant changes in how recidivism risk is algorithmically computed. Notably, it differentiates norm groups for men and women to benchmark scores, modifies items within the scales, and employs new score ranges for assessing individual scores.

The "Summative GRRS" is based on a logistic regression model, expressed as follows:

$$\begin{aligned}\text{newGRRSraw} = &-0.683(1) + 0.593(\text{logcrimv}) + 0.174(\text{voced6}) + 0.489(\text{drugprob5}) - 0.378(\text{age.1}) - 0.613(\text{logage1})\\ &+ 0.393(\text{logarate})\end{aligned}$$

TABLE 6 Questions for the COMPAS-R Vocational/Educational Scale – *Source*: Authors' elaboration adapting from Equivant (2022).

Item of the scale	Item code	Survey question	Answers	Present in COMPAS Core
Chances of work	chance_success_work_R	Right now, if you were to get (or have) a good job how would you rate your chance of being successful?	Good = 0 Fair = 1 Poor = 2	Yes
Employment or school verification	haveempsclool_R	Do you have a verifiable employer or school?	No = 1 Yes = 0	Yes
Completion of high school	high_school_R	Did you complete your high school diploma, GED, or equivalent credential, or are you currently enrolled in a school or program to obtain such a credential?	No = 1 Yes = 0	Yes – added the current enrollment
Possession of job	job_last_year_R	How much have you worked or been enrolled in school within the last 12 months?	12 months full time = 0 12 months part time = 1 6+ months = 2 0–6 months = 3	Yes
Current job	job_R	Do you currently have a job?	No = 1 Yes = 0	Yes
Request for training	need_training_R	Right now, do you feel you need more training in a new job or career skill?	No = 0 Yes = 1	Yes
Possession of skills	skill_R	Do you have a skill, trade or profession in which you usually find work?	No = 1 Yes = 0	Yes
Finding job above minimum wage	wages_above_min_R	How hard is for you to find a job above minimum wage compared to others?	Easier = 0 Same = 1 Harder = 2 Much harder = 3	Yes

where logcrimv stands for the "Legal Involvement scale", voced6 stands for the "Vocational and Educational scale", drugprob5 stands for "Drug Problems scale"; age1 is the age at the time of the assessment; logage1 is the age the moment of the arrest; logarate is the arrest rate.

To simplify usage, Equivant transformed the logistic regression equation results into a summative scale. Also, to render easier the computation of the summative scale each term of the equation was altered to yield non-negative integer outputs (Equivant, 2022). Consequently, each scale used by the Summative GRRS produces a logistic regression Risk Score, which is then transformed into a specific risk contribution score. Row scores are structured into different risk contribution scores (see Figure 3) that can be easily computed by the Summative GRRS algorithm. For instance, the "Legal Involvement" scale raw risk scores are converted into risk contribution scores ranging from 0 to 15; the "Vocational/Educational" scale yields risk contribution scores from 0 to 13; and the Drug Problem scale from 0 to 5. This process of data structuring and aggregation of raw scores into risk contribution values across different scales is explained by Equivant (2022).

Risk contribution scores are calculated as non-negative values, which may not always be integers, using specific parameters for each scale. To simplify addition in the Summative GRRS, these scores are converted to integer values

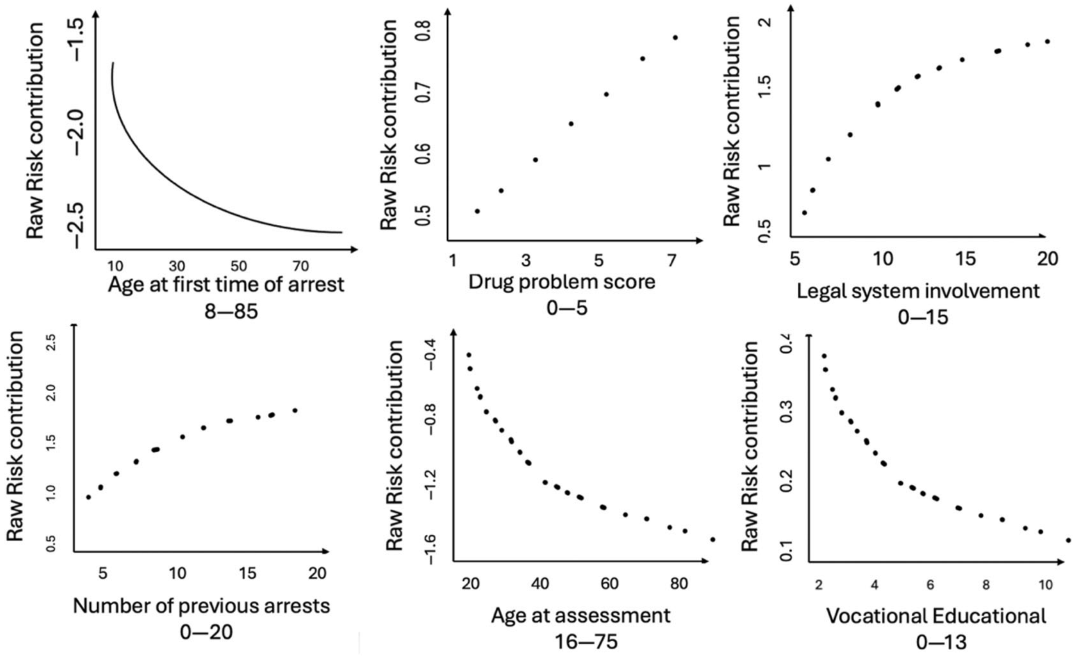

FIGURE 3 Simplified graphic representation of risk scores in different COMPAS-R scales – *Source*: Authors' elaboration using Equivant's publicly available data.

by multiplying them by 10 and rounding to the nearest whole number. Equivant (2022) notes that if the lowest value among the factors is not zero, all scores should be adjusted to ensure the minimum value is zero, thus ensuring consistency across the scale.

The integer scores of each scale are combined to produce a Summative GRRS risk score ranging from 0 to 71. Equivant constructed risk score clusters grouping the scores of the norm group into three segments: the lowest 40%, the subsequent 30%, and the highest 30% for each subgroup. As a result, the risk scores are distributed as follows: for composite males, scores from 0 to 35 are categorized as low risk, 36–40 as medium risk, and 41–71 as high risk; for composite females, scores from 0 to 33 are considered low risk, 34–40 medium risk, and 41–71 high risk.

The means by which data are structured and aggregated, and the logarithmic computation of scores into risk contribution ranging from 0 to 71 and their clustering into low, medium, and high-risk scores significantly redefine the formalization of the risk assessment administrative processes. Equivant acknowledges that the process of rounding and using representative values for intervals in transforming raw scores into contribution scores may introduce distortions in the scoring system (Equivant, 2022). This acknowledgment highlights the importance the algorithmic computation of structured and aggregated raw data has on defining the formalization of administrative processes underlying risk assessments.

6 | DISCUSSION

The COMPAS Core and COMPAS-R algorithms evaluate offenders' data through risk factor scales that employ different data structuring and aggregation methods underpinning the administrative processes to assess recidivism. Each scale utilizes different norm groups and survey questions, converting survey results into numerical values. COMPAS Core and COMPAS-R then integrate these values into their respective scales.

TABLE 7 Differences in construction of the key scales in COMPAS Core and COMPAS-R.

Algorithm	Decile distribution	Cut-off points
COMPAS Core: Criminal Involvement (CrimInv) Scale	Integer values	Low-Risk Group: Aggregates values from 1 to 7 across the first four deciles. Medium-Risk Group: Includes values from 9 to 12. High-Risk Group: Comprises values from 13 to 19.
COMPAS-R: General Recidivism Risk Scale (GenRecidRisk)	Non-integer values	Low-Risk Group: Ranges from −1.30 to −0.40. Medium-Risk Group: Spans from −0.20 to 0.20. High-Risk Group: Extends from 0.40 to 1.90.

COMPAS Core's algorithm correlates offender risk scores with norm group deciles, benchmarking against past offenders' scores in the same decile. The results are then algorithmically computed to produce risk assessment score. COMPAS-R differentiates itself by transforming raw scores into computable, positive integers. This transformation yields distinct aggregated outcomes. COMPAS-R's cut-off points, crucial for defining risk scales clusters, alter the assessment of offenders significantly compared to COMPAS Core. Unique to COMPAS-R is the consistency of its cut-off value across all scales, set uniformly at 1, contrasting with COMPAS Core's variable cut-offs. Despite aiming to evaluate offender risk, COMPAS-R and COMPAS Core's methodologies in structuring, and aggregating data lead to diverse scale's risk scores. Table 7 offers a summary of these differences.

The COMPAS algorithmic assessments of these scales play a crucial role in shaping the administrative decision-making processes in judicial systems, with their outcomes rooted in the scale's scores and the algorithmic computation. These include selecting questions, assigning values to responses, setting inclusion criteria for norm groups, and determining the weight of each question and scale in the risk assessment algorithm.

COMPAS Core's risk assessments depend on how data are structured and aggregated to create scales, the definition of deciles, and the algorithmic computation used for assessing recidivism risk, encapsulated in COMPAS Core's General Recidivism Risk Score. Conversely, COMPAS-R structures survey data and aggregates them into its scales to be computed by logistic regression to predict recidivism likelihood. The scales simplify complex data into clusters for analysis, allowing for the benchmarking of an offender's risk against aggregated data clusters.

Both algorithms develop correlations across different scales through predictive modeling to estimate recidivism risk, benchmarking offender's scores against a historical dataset. Despite similarities, their differences in data structuring, aggregation, and computational methods produce distinct approaches to formalizing administrative processes and informing supervision decisions. The two scales produce inconsistent results (see Figure 4) which highlight the importance of understanding each system's nuances and their impact on risk assessment and supervision recommendations.

The way in which the two different algorithms structure, aggregate, and compute data redesigns the systematic execution of the administrative processes underpinning the formalization of the assessment of felons' risk, as can be seen in Table 8.

6.1 | How COMPAS algorithms introduce a new formalization altering the code, channel, and standard of administrative processes

The COMPAS case illustrates how data structuring, aggregation, and the associated algorithmic computations change the code, channel, and standard of administrative processes. Together, data structuring and aggregation define the

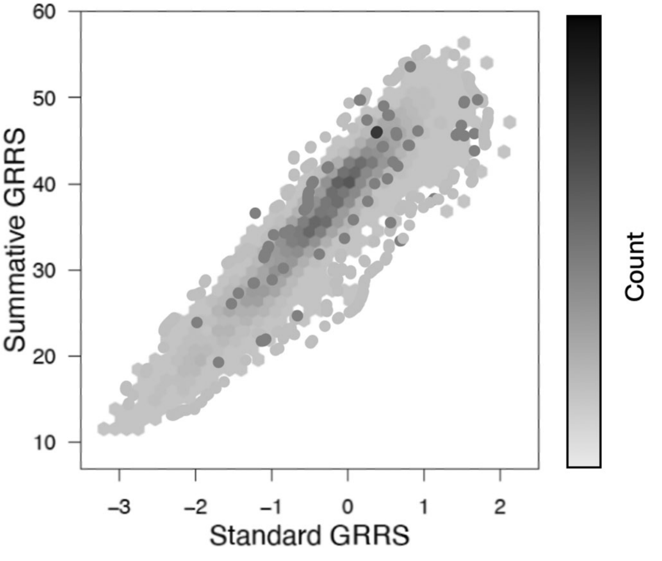

FIGURE 4 Correspondence between COMPAS Core and COMPAS-R risk assessment – *Source*: Authors' graphic elaboration using Equivant's publicly available data.

TABLE 8 Differences in how COMPAS Core and COMPAS-R structure and aggregate data.

	COMPAS Core	COMPAS-R
Data structuring	Surveys' answers are translated into integer numerical values.	New surveys' answers are translated into integer positive numerical values.
Data aggregation	Within the normative group, decile distribution is defined by an algorithm.	Definition of three risk clusters (low, medium, high) through a logistic function.
	Cut-off points vary across different scales.	Cut-off points consistent across logarithmic scales.

values and criteria used by the different COMPAS algorithms to create new code, channel, and standard that formalize the administrative processes of risk assessment.

Prior to COMPAS, felons' risk scores were assessed using a point-based system where each entry carried the same weight. Data were collected following administrative processes as those outlined in systems such as the federal pretrial risk assessment (Administrative Office of the U.S. Courts, 2013). Following the user guide's instructions, officers collected information on offenders and calculated risk scores by adding up the points for each criterion. COMPAS then introduced new administrative processes to generate these risk scores, embedding these procedures within the algorithm's code. This innovation led to the creation of new formalization standards, against which offenders' behaviors are benchmarked, influencing the allocation of rewards and punishments.

The risk assessments generated by this automated process serve as new benchmarks for judges, who use these scores to inform their decisions. Offenders' profiles are analyzed and processed through the algorithm, diverging from the administrative processes traditionally followed, such as those outlined in the federal pretrial risk assessment guide by the Office of Probation and Pretrial Services, overseen by The Criminal Law Committee of the Judicial Conference of the United States.

The COMPAS algorithms produce new benchmarks assessing offenders based on collective experiences encoded into data, ignoring individual circumstances. This leads to the creation of benchmarks that "objectify" personal offenders' circumstances into abstract numbers (Au, 2022), then compared against structured and aggregated data clusters. This objectification arises from the risk scales introducing a new formalization code that simplifies the complex factors defining each individual felon's circumstances into less complex clusters aggregated within the scales. By benchmarking individuals against these simpler clusters, a new formalization channel is created, which narrows the variance in assessing felons' risk.

Assessing felons' risk through COMPAS poses challenges in connecting predictions to individuals' backgrounds. The COMPAS system forecasts an individual's potential harm to American society using an algorithmic computation that simplifies individual circumstances into a new, less complex code. Consequently, human actors in judicial courts rely on decontextualized data, shaped by the code, channel, and standard established by COMPAS algorithms' specific data structuring, aggregation, and computations. This necessitates that judges evaluate cases based on information abstracted by the algorithm, deviating from the nuanced personal data and circumstances traditionally considered under legal norms and rules of judicial processes (United Nations, 2003).

6.2 | The assessment of risk recidivism through COMPAS: Formalizing judicial administrative process

The adoption of algorithms within the judicial system was promoted for their potential to systematize the administrative procedures for assessing individual offenders' risk scores. By structuring and aggregating necessary data, algorithms aim to reduce data inconsistencies and computational errors, thereby increasing the reliability of administrative processes. However, this method creates artificial scales and imposes arbitrary divisions, leading to the generation of data that strips away the context from individual profiles. These profiles are then compared, based on the algorithm's logic, to the scores of other offenders within a predefined normative group.

The process of assessing an individual offender's risk involves algorithmically generated correlations across various COMPAS scales. These correlations reduce complex psychological, behavioral, social, economic, and institutional factors that influence recidivism into numerically weighted risk factors. Thus, COMPAS assessments do not directly reflect an individual's risk of recidivism but rather offer predictions grounded in the administrative processes coded into the algorithm. This allows for the comparison of an offender's data against a series of risk scales, formulated based on criteria that define variables for a normative group, abstracting away from the individual's specific circumstances.

The way COMPAS algorithms analyze and process information significantly affects decisions regarding sentencing or supervision. Research has shown that COMPAS scores influence administrative decisions by introducing cognitive biases, leading decision-makers to align their judgments with the algorithm's outputs (Chouldechova, 2017; Vaccaro, 2019).

The administrative processes by which judges make decisions are informed by algorithm-generated risk scores. While algorithms don't take the final decision, they provide critical data for judges' information analysis and processing, which, being decontextualized, cannot be fully comprehended, or contested. This leads to a new formalization of judicial administrative processes. Algorithms define key attributes of the judicial administrative processes that are impossible to challenge or dispute by human actors. Table 9 provides an explanation of how COMPAS algorithms impact formalization, with the latter broken down into the three main dimensions of code, channel, and standard.

7 | CONTRIBUTION: ALGORITHMIC FORMALIZATION

The formalization introduced by COMPAS fundamentally alters the underpinning administrative processes. The specific characteristics of data structuring, aggregation, and algorithmic computation result in a new algorithmic

TABLE 9 Impact of COMPAS algorithms on formalization as code, channel, and standard.

	Formalization as code	Formalization as channel	Formalization as standard
Impact of COMPAS on administrative processes	COMPAS deconstructs the complexity of offenders' profiles. COMPAS reconstructs offenders' profiles into structured data that are computable to algorithms. COMPAS benchmarks offenders' individual profiles against codified scales and clusters of other offenders.	COMPAS automates the production of offenders' scores releasing an output that reflects new correlations between structured data and codified scales and clusters. COMPAS' output decreases discretion in the administrative processes used to assess felons' risk to reoffend.	COMPAS constraints judges' actions requiring them to perform administrative processes against the algorithmic benchmark they receive. Judges do not have the possibility to understand or dispute the COMPAS score which is the outcome of structured and aggregated data.

formalization acting as a code, a channel, and a standard, each playing a distinct role in shaping administrative processes. To assess the impact of algorithmic formalization on administrative processes, we test the theoretical propositions we have previously presented.

7.1 | Algorithmic formalization as code: Simplifying complexity into computable data

Proposition 1 posits that algorithmic formalization acts as a code reducing the complexity of the context into abstract and simplified data. In the case of COMPAS, the reduction of complex rules and procedures into computable formulae means transforming detailed, nuanced information about felons into structured data that the algorithm can process. This involves data structuring, where raw information is organized into a format suitable for computation, and data aggregation, where these structured data are combined into coherent groups or clusters. By doing so, the unique profiles of felons are decontextualized, stripping away individual conditions to fit the rigid structure required by the algorithm. This simplification is crucial for the algorithm to function effectively, but it also means that some of the nuances of the individual cases are lost. Proposition 1 is hence confirmed, as the algorithmic reduction of complex information into simplified and computable forms fundamentally alters the administrative process in which it is utilized.

7.2 | Algorithmic formalization as channel: Reducing discretion through predefined mechanisms

Proposition 2 posits that algorithmic formalization acting as a channel reduces the discretion within administrative processes by outlining predefined mechanisms for structuring, aggregating, and processing data. In the COMPAS system, the goal of reducing variance is achieved through automation, which standardizes the computation of risk assessments. This channeling effect means that decisions are made based on consistent, algorithm-driven processes rather than on human judgment, which can be mutable. While this increases consistency and predictability, it also diminishes the discretionary power of judges and other officials, potentially oversimplifying complex human behaviors and contexts. Evidence has been provided to support Proposition 2, as the algorithm automates the production of outputs in the administrative processes it mediates, thus reducing variance in the performance carried over by human actors.

7.3 | Algorithmic formalization as standard: Establishing benchmarks for administrative processes

Proposition 3 posits that algorithmic formalization establishes a standard against which felons' profiles and judges' decisions are benchmarked. These standardized administrative processes emerge from the structured, aggregated, and algorithmically processed data, creating benchmarks that are used to assess offenders' risk scores. However, this decontextualized standard can sometimes overlook the unique circumstances of individual cases. By establishing uniform criteria for evaluation, the algorithm ensures consistency but at the risk of applying a one-size-fits-all approach that may not be suitable for every situation. Testing Proposition 3, we can observe that the creation of these decontextualized standards is necessary to ensure uniformity and minimize variability in administrative processes mediated by the algorithm. Accordingly, Proposition 3 is confirmed.

7.4 | From administrative to algorithmic formalization: Evolution and implications

The findings from the analysis of the COMPAS case study suggest that algorithmic formalization possesses unique features distinct from traditional administrative formalization. Algorithmic formalization as a code simplifies complex administrative processes into computable data but involves decontextualizing felons' profiles. As a channel, it reduces variance by automating processes, thereby diminishing human discretion. As a standard, it establishes benchmarks based on decontextualized data, potentially overlooking individual circumstances. Table 10 offers a breakdown of the main differences between traditional formalization and algorithmic formalization, highlighting the specific contributions of this paper.

The analysis of the COMPAS system reveals how the propositions hold true, illustrating the transformative impact of algorithmic formalization on administrative processes. This examination underscores the significant changes in how processes are structured and executed, revealing the complex balance between efficiency, consistency, and the potential loss of nuance and discretion.

The integration of algorithms into administrative processes introduces a new type of formalization, which builds upon, but distinctly evolves from, the foundational principles outlined by Walsh and Dewar (1987). This evolution is marked by several unique attributes. Firstly, the mechanisms that constitute algorithmic formalization significantly diverge from conventional methods, indicating a fundamental shift in how processes are structured and executed. Secondly, this form of formalization is characterized by its dynamic and evolutionary nature, driven by continuous

TABLE 10 Algorithmic formalization compared against formalization in Walsh and Dewar (1987).

Dimension	Formalization in Walsh and Dewar (1987)	Algorithmic formalization
Code	Formalization as code reduces a complex set of activities to fewer complex formulae.	Algorithmic formalization as code reduces complex information into abstract and decontextualized data computable by algorithms. To do so, contextual individual conditions must be stripped from the data.
Channel	Formalization as channel decreases variance in human performance.	Algorithmic formalization as channel automates the execution of tasks decreasing discretion within the administrative processes defining predefined mechanisms used to structure, aggregate and algorithmically process data.
Standard	Formalization establishes the standards against which action is compared, and rewards and punishments are provided.	Algorithmic formalization produces a decontextualized standard that serves as benchmark for the administrative processes.

updates to data inputs and the ongoing refinement and advancement of algorithmic techniques. Highlighting the specific features and implications of this unique form of formalization is crucial for comprehending its influence on the administrative processes within the public sector and the changes it introduces. Specifically, we have identified that algorithmic formalization heavily relies on advanced computational techniques and data processing technologies which require specific data formats. It incorporates sophisticated algorithms and data logics that produce a fundamental shift in processes structuring and execution. Algorithmic formalization is dynamic and evolutionary, driven by continuous updates to data inputs and ongoing advancements in algorithmic techniques. The impact of algorithmic formalization on administrative operations is profound. It reshapes traditional workflows by automating routine tasks and standardizing processes, which enhances efficiency but also requires a re-evaluation of roles and responsibilities within public administrations.

8 | CONCLUSIONS AND OUTLOOK

Research has explored how algorithms influence the structure and workflows of public administrations, introducing an "additional level of complexity" (König & Wenzelburger, 2020), transforming administrative decision-making processes (Bullock et al., 2020), and impacting the discretion, autonomy, and consistency of public administration service delivery (Schiff et al., 2021). However, existing studies have often overlooked the distinct characteristics of algorithms and their direct impact on the administrative processes they mediate. This paper aims to bridge this gap by leveraging the concept of formalization (Walsh & Dewar, 1987) and theorizing algorithmic formalization as a distinctive form of formalization within administrative processes.

8.1 | Implications for scholars and practitioners

This paper enhances the literature by detailing how algorithms manage data structuring and aggregation to render data computable and how algorithmic computation introduces new forms of code, channel, and standard in formalizing administrative processes. Our study offers several contributions to existing literature, including the introduction of the theoretical concept of algorithmic formalization, to better understand the influence of algorithms on the administrative processes of public organizations. By focusing on the specifics of data structuring, aggregation, and algorithmic computation, we underscore the importance of closely examining the technological nuances of algorithms crucial for grasping their subtle effects.

This paper has also important implications for professionals and practitioners. The rise of algorithmic formalization in administrative processes significantly curtails the discretion of human actors reliant on algorithmic data for analyzing and processing information vital to performing administrative tasks and delivering public services. Therefore, in deploying algorithms within public administration, it is essential to critically examine how they structure and aggregate data to facilitate algorithmic computation. The deployment of algorithms in public administration has been linked to problematic, biased, or discriminatory outcomes, adversely affecting the citizens served by these entities. Through an in-depth analysis of how algorithms operate and thereby formalize administrative processes, public administration policymakers can better anticipate, address, and understand these issues.

8.2 | Limitations and future research

Two boundary conditions are worth noting in this study. First, limitations in data collection: the documents examined included information about the adoption of COMPAS algorithms in the U.S. Judiciary administrative processes but did not disclose the complete structure of the algorithms. These algorithms are protected by property rights. To address this, we focused on studying how the structuring and aggregation of data by the algorithms impacted the administrative process.

Second, we are aware that the version of the algorithm we examined is relatively basic compared to more recent ones, such as those utilizing Machine Learning or Generative AI. However, we believe that the generalizability of our study is not compromised by the specific type of algorithm examined. The investigation into how algorithms reshape administrative processes and transform workflows within public administrations offers relevant insights regardless of the algorithm's nature. The external validity of our research is strengthened by anchoring the case study to a well-established theory, formalization. Reflecting theoretical propositions is crucial to correct imprecisions in the case narrative and enhance the explanation's solidity (Yin, 2018). Combining thorough data analysis with meaningful theoretical constructs supports the study's generalizability because the theoretical findings can be challenged, replicated, or tested in other contexts (Bacharach, 1989).

We encourage further research on the algorithmic impact on administrative processes. First, scholars can utilize the concept of algorithmic formalization and assess it in other organizational settings. The increasing adoption of algorithms to mediate workflows and decisions offers opportunities to discuss algorithmic formalization effects in other sectors of public administration. Second, a nuanced analysis of the specific processes transformed by algorithm adoption can provide insights into how human actors react or adapt to reshaped administrative processes, and how this affects the design and delivery of key public services. Lastly, research can examine whether algorithmic formalization, beyond its impact on administrative processes, poses challenges to the regulatory and normative underpinnings of public administration.

ACKNOWLEDGMENTS

The authors thank the Associate Editor, Bruce McDonald III, and the three anonymous reviewers for their valuable guidance during the revision process. The authors received helpful feedback and comments from participants at seminars at the London School of Economics, at the Center for Technology in Government (SUNY at Albany), and at the Hawaii International Conference on System Sciences 2022, the pre-International Conference on Information Systems e-Government Workshop 2022, and the Academy of Management Annual Meeting 2023.

CONFLICT OF INTEREST STATEMENT

The authors certify that they have no affiliations with or involvement in any organization or entity with any financial interest (such as honoraria; educational grants; participation in speakers' bureaus; membership, employment, consultancies, stock ownership, or other equity interest; and expert testimony or patent-licensing arrangements), or non-financial interest (such as personal or professional relationships, affiliations, knowledge or beliefs) in the subject matter or materials discussed in this manuscript.

DATA AVAILABILITY STATEMENT

To properly manage, store, and access the data, the authors have followed the Self-Archiving policy jointly agreed upon by Wiley and the London School of Economics. Accordingly, a Data Management Plan named after the title of the paper, "Algorithmic formalization: impacts on administrative processes", has been created with the ID154522. It can be shared upon request by the authors. The secondary data supporting the findings of this study are stored in "PA Algorithmic formalization" repository at: https://github.com/franzgualdi/PA-Algorithmic-formalization.git. These data were derived from the following resources available in the public domain: Equivant (2019, 2022) and Northpointe (2009, 2015).

PERMISSION TO REPRODUCE MATERIAL FROM OTHER SOURCES

All the figures and tables in the paper are the output of authors' own elaboration. No copyrighted material has been reproduced. When figures and tables have been created using publicly available data from copyrighted material, we have credited the source and added "adapting from". In any case, no figure or table exceeds the 75%–25% rule as prescribed by Wiley rules.

ORCID

Francesco Gualdi https://orcid.org/0000-0003-4975-9133

ENDNOTES

[1] Algorithmic formalization partially resonates with the concept of formalization in Computer Science discipline, which involves translating procedures into a formal, rigid, executable language suitable for software and computer models (Giannakopoulou et al., 2021). However, algorithmic formalization, as presented in this paper, is the outcome of the interdependent processes of data structuring, aggregation and algorithmic computation rather than solely of the process by which procedures and operations ae translated into a formal, precise, and unambiguous language that can be executed by software and computer models.

[2] On the 9th of January 2017, Northpointe Inc. merged with Courtview Justice Solutions Inc. and Constellation Justice Systems Inc. to create Equivant, a new company. For clarity purpose, in references and bibliography we use Northpointe before 2017 and Equivant after that date.

REFERENCES

Administrative Office of the U.S. Courts. (2013) *Pretrial Services Risk Assessment (PTRA) user guide*. Available from: https://cjdata.tooltrack.org/sites/default/files/2018-10/PTRA%202.3%20User%20Guide%20May%202013_0.pdf

Agarwal, P.K. (2018) Public administration challenges in the world of AI and bots. *Public Administration Review*, *78*(6), 917–921. Available from: https://doi.org/10.1111/puar.12979

Alaimo, C. & Kallinikos, J. (2017) Computing the everyday: social media as data platforms. *The Information Society*, *33*(4), 175–191. Available from: https://doi.org/10.1080/01972243.2017.1318327

Ammitzbøll Flügge, A., Hildebrandt, T. & Møller, N.H. (2021) Street-level algorithms and AI in bureaucratic decision-making: a caseworker perspective. In: *Proceedings of the ACM on human-computer interaction, 5 (CSCW1)*. Available from: https://doi.org/10.1145/3449114

Au, W. (2022) *Unequal by design: high-stakes testing and the standardization of inequality*, 2nd edition. New York: Routledge. Available from: https://doi.org/10.4324/9781003005179

Ávila, F., Hannah-Moffat, K. & Maurutto, P. (2020) The seductiveness of fairness: is machine learning the answer? – algorithmic fairness in criminal justice systems. In: Schuilenburg, M. & Peeters, R. (Eds.) *The algorithmic society: technology, power, and knowledge* (1st ed., pp. 87–103). London: Routledge. Available from: https://doi.org/10.4324/9780429261404

Bacharach, S.B. (1989) Organizational theories: some criteria for evaluation. *Academy of Management Review*, *14*(4), 496–515. Available from: https://doi.org/10.5465/amr.1989.4308374

Baxter, P. & Jack, S. (2012) Qualitative case study methodology: study design and implementation for novice researchers. *The Qualitative Report*, *13*(4), 544–559. Available from: https://doi.org/10.46743/2160-3715/2008.1573

Berente, N., Gu, B., Recker, J. & Santhanam, R. (2021) Managing Artificial Intelligence. *MIS Quarterly*, *45*(3), 1433–1450. Available from: https://doi.org/10.25300/MISQ/2021/16274

Bouckaert, G., Peters, B.G. & Verhoest, K. (2016) *The coordination of public sector organizations: shifting patterns of public management*. Basingstoke: Palgrave MacMillan.

Brunsson, N., Rasche, A. & Seidl, D. (2012) The dynamics of standardization: three perspectives on standards in organization studies. *Organization Studies*, *33*(5–6), 613–632. Available from: https://doi.org/10.1177/0170840612450120

Brynjolfsson, E. & McAfee, A. (2014) *The second machine age: work, Progress, and prosperity in a time of brilliant technologies*. New York: W. W. Norton & Company.

Bullock, J.B., Young, M.M. & Wang, Y.-F. (2020) Artificial intelligence, bureaucratic form, and discretion in public service. *Information Polity*, *25*(4), 491–506. Available from: https://doi.org/10.3233/IP-200223

Busuioc, M. (2021) Accountable artificial intelligence: holding algorithms to account. *Public Administration Review*, *81*(5), 825–836. Available from: https://doi.org/10.1111/puar.13293

Chander, A. (2017) The racist algorithm? *Michigan Law Review*, *115*(6), 1023–1045.

Chouldechova, A. (2017) Fair prediction with disparate impact: a study of bias in recidivism prediction instruments. *Big Data*, *5*(2), 153–163. Available from: https://doi.org/10.1089/big.2016.0047

Christensen, T. & Lægreid, P. (2002) New public management: puzzles of democracy and the influence of citizens. *Journal of Political Philosophy*, *10*(3), 267–295. Available from: https://doi.org/10.1111/1467-9760.00153

Conte, S.D. & De Boor, C. (2018) *Elementary numerical analysis: an algorithmic approach*, 3rd edition. Philadelphia, PA: Society for Industrial and Applied Mathematics.

Cordella, A. (2007) E-government: towards the e-bureaucratic form? *Journal of Information Technology*, *22*(3), 265–274. Available from: https://doi.org/10.1057/palgrave.jit.2000105

Cormen, T., Leiserson, C., Rivest, R. & Stein, C. (2009) *Introduction to algorithm*, 3rd edition. Cambdrige, MA: The MIT Press.

Criado, J.I., Valero, J. & Villodre, J. (2020) Algorithmic transparency and bureaucratic discretion: the case of SALER early warning system. *Information Polity*, *25*(4), 449–470. Available from: https://doi.org/10.3233/IP-200260

Dalton, D.R., Todor, W.D., Spendolini, M.J., Fielding, G.J. & Porter, L.W. (1980) Organization structure and performance: a critical review. *Academy of Management Review*, *5*(1), 49–64. Available from: https://doi.org/10.2307/257804

de Sousa, W.G., Fidelis, R.A., de Souza Bermejo, P.H., da Silva Gonçalo, A.G. & de Souza Melo, B. (2022) Artificial intelligence and speedy trial in the judiciary: myth, reality or need? A case study in the Brazilian Supreme Court (STF). *Government Information Quarterly*, *39*(1), 101660. Available from: https://doi.org/10.1016/j.giq.2021.101660

Denhardt, R.B. & Denhardt, J.V. (2000) The new public service: serving rather than steering. *Public Administration Review*, *60*(6), 549–559. Available from: https://doi.org/10.1111/0033-3352.00117

Equivant. (2019) *Practitioner's guide to COMPAS Core*. Available from: https://www.equivant.com/wp-content/uploads/Practitioners-Guide-to-COMPAS-Core-040419.pdf

Equivant. (2022) *COMPAS-R core: The revised version of the standard COMPAS core*.

Fayol, H. (1949) *General and industrial management*. London: Pitman Publishing.

Fishenden, J. & Thompson, M. (2012) Digital government, open architecture, and innovation: why public sector IT will never be the same again. *Journal of Public Administration Research and Theory*, *23*(4), 977–1004. Available from: https://doi.org/10.1093/jopart/mus022

Giannakopoulou, D., Pressburger, T., Mavridou, A. & Schumann, J. (2021) Automated formalization of structured natural language requirements. *Information and Software Technology*, *137*, 106590. Available from: https://doi.org/10.1016/j.infsof.2021.106590

Gillespie, T., Boczkowski, P.J. & Foot, K.A. (2014) *Media technologies: essays on communication, materiality, and society*. Cambdrige, MA: MIT Press.

Grimmelikhuijsen, S. (2022) Explaining why the computer says no: algorithmic transparency affects the perceived trustworthiness of automated decision-making. *Public Administration Review*, *83*(2), 241–262. Available from: https://doi.org/10.1111/puar.13483

Gualdi, F. & Cordella, A. (2024) Artificial intelligence to support public sector decision-making: the emergence of entangled accountability. In: Constantiou, I.D., Joshi, M.P. & Stelmaszak, M. (Eds.) *Research handbook on artificial intelligence and decision making in organizations*. Cheltenham: Edward Elgar Publishing, pp. 266–281. Available from: https://doi.org/10.4337/9781803926216.00024

Gulick, L. (1937) Notes on the theory of organization. In: Gulick, L. & Urwick, L. (Eds.) *Papers on the science of administration* (Vol. 3, pp. 1–45). New York: Institute of Public Administration, Columbia University.

Guy, M.E. & McCandless, S.A. (2012) Social equity: its legacy, its promise. *Public Administration Review*, *72*(S1), S5–S13. Available from: https://doi.org/10.1111/j.1540-6210.2012.02635.x

Hamilton, M. (2019) The sexist algorithm. *Behavioral Sciences & the Law*, *37*(2), 145–157. Available from: https://doi.org/10.1002/bsl.2406

Hansen, J.A. (2022) Can details depoliticize? An examination of the formalization strategy. *Public Administration*, *102*(1), 249–263. Available from: https://doi.org/10.1111/padm.12902

Hood, C. (1991) A public management for all seasons? *Public Administration*, *69*(1), 3–19. Available from: https://doi.org/10.1111/j.1467-9299.1991.tb00779.x

Hughes, O.E. (2017) *Public management and administration*. London: Bloomsbury Publishing.

Humerick, J.D. (2019) Reprogramming fairness: affirmative action in algorithmic criminal sentencing. *HRLR Online*, *4*, 213.

Ingrams, A. (2020) A machine learning approach to open public comments for policymaking. *Information Polity*, *25*(4), 433–448. Available from: https://doi.org/10.3233/IP-200256

Kehl, D.L., Guo, P. & Kessler, S. (2017) *Algorithms in the criminal justice system: assessing the use of risk assessments in sentencing*. Available from: http://nrs.harvard.edu/urn-3:HUL.InstRepos:33746041

Kirkpatrick, K. (2017) It's not the algorithm, it's the data. *Communications of the ACM*, *60*(2), 21–23. Available from: https://doi.org/10.1145/3022181

Kitchin, R. (2019) Thinking critically about and researching algorithms. In: Beer, D. (Ed.) *The social power of algorithms*. New York: Routledge, pp. 14–29.

König, P.D. & Wenzelburger, G. (2020) Opportunity for renewal or disruptive force? How artificial intelligence alters democratic politics. *Government Information Quarterly*, *37*(3), 101489. Available from: https://doi.org/10.1016/j.giq.2020.101489

Kwon, S.-W. (2008) Does the standardization process matter? A study of cost effectiveness in hospital drug formularies. *Management Science*, *54*(6), 1065–1079. Available from: https://doi.org/10.1287/mnsc.1070.0839

Larson, J., Mattu, S., Kirchner, L. & Angwin, J. (2016) *How we analyzed the COMPAS recidivism algorithm*. ProPublica. Available from: https://www.propublica.org/article/how-we-analyzed-the-compas-recidivism-algorithm

Lee, J. (2024) Behind and beyond the standard(ization) trap: diversifying power sources by differentiating centralization and standardization. *International Review of Public Administration*, *29*(1), 1–19. Available from: https://doi.org/10.1080/12294659.2024.2310899

Lim, J.H. & Tang, S.-Y. (2008) Urban E-government initiatives and environmental decision performance in Korea. *Journal of Public Administration Research and Theory*, *18*(1), 109–138. Available from: https://doi.org/10.1093/jopart/mum005

Margetts, H. & John, P. (2023) How rediscovering nodality can improve democratic governance in a digital world. *Public Administration*, *1–15*, 12960. Available from: https://doi.org/10.1111/padm.12960

Martin, K. (2019) Ethical implications and accountability of algorithms. *Journal of Business Ethics*, *160*(4), 835–850. Available from: https://doi.org/10.1007/s10551-018-3921-3

McCorduck, P. (2004) *Machines who think: a personal inquiry into the history and prospects of artificial intelligence*. New York: CRC Press.

McDonald, B.D., Hall, J.L., O'Flynn, J. & van Thiel, S. (2022) The future of public administration research: an editor's perspective. *Public Administration*, *100*(1), 59–71. Available from: https://doi.org/10.1111/padm.12829

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

Meijer, A., Lorenz, L. & Wessels, M. (2021) Algorithmization of bureaucratic organizations: using a practice lens to study how context shapes predictive policing systems. *Public Administration Review*, *81*(5), 837–846. Available from: https://doi.org/10.1111/puar.13391

Northpointe. (2009) *Measurement & treatment implications of COMPAS core scals*. Available from: https://www.michigan.gov/documents/corrections/Timothy_Brenne_Ph.D.__Meaning_and_Treatment_Implications_of_COMPA_Core_Scales_297495_7.pdf

Northpointe. (2015) *Practitioner's guide to COMPAS core*. Available from: http://www.northpointeinc.com/downloads/compas/Practitioners-Guide-COMPAS-Core-_031915.pdf

O'Neil, C. (2017) *Weapons of math destruction: how big data increases inequality and threatens democracy*. New York: Crown Publishing Group.

Ospina, S.M., Esteve, M. & Lee, S. (2018) Assessing qualitative studies in public administration research. *Public Administration Review*, *78*(4), 593–605. Available from: https://doi.org/10.1111/puar.12837

Pencheva, I., Esteve, M. & Mikhaylov, S.J. (2020) Big data and AI–A transformational shift for government: so, what next for research? *Public Policy and Administration*, *35*(1), 24–44. Available from: https://doi.org/10.1177/0952076718780537

Pollitt, C. & Bouckaert, G. (2017) *Public management reform: a comparative analysis - into the age of austerity*. Oxford: Oxford University Press.

Rainey, H.G. & Bozeman, B. (2000) Comparing public and private organizations: empirical research and the power of the a priori. *Journal of Public Administration Research and Theory*, *10*(2), 447–470. Available from: https://doi.org/10.1093/oxfordjournals.jpart.a024276

Schiff, D.S., Schiff, K.J. & Pierson, P. (2021) Assessing public value failure in government adoption of artificial intelligence. *Public Administration*, *100*(3), 653–673. Available from: https://doi.org/10.1111/padm.12742

Scholta, H., Balta, D., Räckers, M., Becker, J. & Krcmar, H. (2020) Standardization of forms in governments. *Business & Information Systems Engineering*, *62*(6), 535–560. Available from: https://doi.org/10.1007/s12599-019-00623-1

Stevenson, M. & Doleac, J.L. (2022) *Algorithmic risk assessment in the hands of humans*. Available from: https://doi.org/10.2139/ssrn.3489440

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. Available from: https://doi.org/10.1016/j.giq.2018.09.008

Turner, S., Hess, J., Bradstreet, C., Chapman, S. & Murphy, A. (2013) *Development of the California Static Risk Assessment (CSRA): recidivism risk prediction in the California Department of Corrections and Rehabilitation*. Available from: https://ucicorrections.seweb.uci.edu/files/2013/12/Development-of-the-CSRA-Recidivism-Risk-Prediction-in-the-CDCR.pdf

United Nations. (2003) *Human rights in the administration of justice: a manual on human rights for judges, Presecutors and lawyers*. New York and Geneva: United Nations.

Vaccaro, M.A. (2019) *Algorithms in human decision-making: a case study with the COMPAS risk assessment software*. Cambridge, MA: Harvard University. Available from: https://nrs.harvard.edu/URN-3:HUL.INSTREPOS:37364659

Van der Voort, H.G., Klievink, B., Arnaboldi, M. & Meijer, A.J. (2019) Rationality and politics of algorithms. Will the promise of big data survive the dynamics of public decision making? *Government Information Quarterly*, *36*(1), 27–38. Available from: https://doi.org/10.1016/j.giq.2018.10.011

van Leijen, H. (2005) The role of Contextuality in process standardization. In: Baets, W. (Ed.) *Knowledge management and management learning: extending the horizons of knowledge-based management*. New York: Springer US, pp. 251–286. Available from: https://doi.org/10.1007/0-387-25846-9_15

von Krogh, G. (2018) Artificial intelligence in organizations: new opportunities for phenomenon-based theorizing. *Academy of Management Discoveries*, *4*(4), 404–409. Available from: https://doi.org/10.5465/amd.2018.0084

Walsh, J.P. & Dewar, R.D. (1987) Formalization and the organizational life cycle. *Journal of Management Studies*, *24*(3), 215–231. Available from: https://doi.org/10.1111/j.1467-6486.1987.tb00700.x

Washington, A.L. (2018) How to argue with an algorithm: lessons from the COMPAS-ProPublica debate. *Colorado Technology Law Journal*, *17*(1), 131–160.

Waugh, W.L., Jr. & Streib, G. (2006) Collaboration and leadership for effective emergency management. *Public Administration Review*, *66*(s1), 131–140. Available from: https://doi.org/10.1111/j.1540-6210.2006.00673.x

Weber, M. (1922) *Economy and society: an outline of interpretive sociology*. Berkeley, CA: University of California Press.

Wenzelburger, G., König, P.D., Felfeli, J. & Achtziger, A. (2024) Algorithms in the public sector. Why context matters. *Public Administration*, *102*(1), 40–60. Available from: https://doi.org/10.1111/padm.12901

Wirtz, B.W., Weyerer, J.C. & Sturm, B.J. (2020) The dark sides of artificial intelligence: an integrated AI governance framework for public administration. *International Journal of Public Administration*, *43*(9), 818–829. Available from: https://doi.org/10.1080/01900692.2020.1749851

Witten, I.H., Frank, E. & Hall, M.A. (2011) Data mining. In: *Practical machine learning tools and techniques*, 3rd edition. San Francisco, CA: Morgan Kaufmann.

Yin, R.K. (2018) *Case study research and applications: design and methods*, 6th edition. Thousand Oaks, CA: SAGE Publications.

Young, M.M., Bullock, J.B. & 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*, *2*(4), 301–313. Available from: https://doi.org/10.1093/ppmgov/gvz014

How to cite this article: Cordella, A., & Gualdi, F. (2024). Algorithmic formalization: Impacts on administrative processes. *Public Administration*, 1–26. https://doi.org/10.1111/padm.13030