



Uncertainty, risk and the use of algorithms in policy decisions: a case study on criminal justice in the USA

Kathrin Hartmann¹ · Georg Wenzelburger¹ 

Accepted: 4 December 2020 / Published online: 29 January 2021
© The Author(s) 2021

Abstract

Algorithms are increasingly used in different domains of public policy. They help humans to profile unemployed, support administrations to detect tax fraud and give recidivism risk scores that judges or criminal justice managers take into account when they make bail decisions. In recent years, critics have increasingly pointed to ethical challenges of these tools and emphasized problems of discrimination, opaqueness or accountability, and computer scientists have proposed technical solutions to these issues. In contrast to these important debates, the literature on how these tools are implemented in the actual everyday decision-making process has remained cursory. This is problematic because the consequences of ADM systems are at least as dependent on the implementation in an actual decision-making context as on their technical features. In this study, we show how the introduction of risk assessment tools in the criminal justice sector on the local level in the USA has deeply transformed the decision-making process. We argue that this is mainly due to the fact that the evidence generated by the algorithm introduces a notion of statistical prediction to a situation which was dominated by fundamental uncertainty about the outcome before. While this expectation is supported by the case study evidence, the possibility to shift blame to the algorithm does seem much less important to the criminal justice actors.

Keywords Artificial intelligence · Big data · Public policy · Public administration · Criminal justice

Introduction

The increased use of algorithmic decision-making (ADM) systems in many domains of public life has spurred a debate about the chances and risks involved. Optimists emphasize that algorithms are capable of recognizing patterns in enormous amounts of data very rapidly—tasks humans would never be able to fulfill at similar speed. They also hold that evidence-based decision-making is enhanced by the use of artificial intelligence, not least because ADM systems do not suffer from the well-known psychological biases (Kahneman

✉ Georg Wenzelburger
georg.wenzelburger@sowi.uni-kl.de

¹ Fachbereich Sozialwissenschaften, Policy-Analyse und Politische Ökonomie, Technische Universität Kaiserslautern, Postfach 3049, 67653 Kaiserslautern, Germany

2012) that plague human decision-making (for a review, see Einhorn and Hogarth 1981). In sum, this view suggests that the use of ADM will increase the efficiency of public services, provide well-informed predictions or boost the speed of decision-making by bureaucracies (on “digital era governance,” see Dunleavy et al. (2005)). In contrast, critical voices ring the alarm bells emphasizing the opaqueness of algorithms (Ananny and Crawford 2018; Pasquale 2016) and potentially in-built biases (Angwin et al. 2016). Moreover, legal challenges related to the lack of accountability, due process and equal protection arise when ADM systems are used (Yeung 2018; Kehl et al. 2017).

While all these debates on the risks and chances of ADM use are important, we know surprisingly little about the real-life implementation and the effects of ADM systems on decision-making processes (on the use in the Criminal Justice (CJ) system, see Stevenson 2018). In fact, there are only very few studies that analyze how political and bureaucratic actors use ADM and what this changes to their behavior (see for instance Berk 2017; Stevenson 2018). Van der Voort et al. argue in a recent paper, most analyses “neglect the institutions that shape the process from data generation to the decisions taken” (van der Voort et al. 2019, 27) and Stevenson (Stevenson 2018, 341) deplores a “sore lack of research on the impacts of risk assessment in practice.” This is problematic because the consequences of ADM systems are at least as dependent on the implementation in an actual decision-making context as on their technical features (Zweig et al. 2018, 189–191). Zavarsnik holds, for instance, that even if algorithms tools are used in semi-automated contexts, “the process of arriving at a decision changes. The perception of accountability for the final decision changes too. The decision-makers will be inclined to tweak their own estimates of risk to match the model’s” (Završnik 2019, 13). Therefore, only if we know how ADM systems work on the ground can we assess chances and risks of ADM use (see also Veale et al. 2018).

To address this shortcoming of the current literature, our article provides a case study on how risk assessment software based on machine learning (Correctional Offender Management Profiling for Alternative Sanctions, COMPAS) has been implemented in the CJ system in Eau Claire County in Wisconsin. Our goal is explorative in nature. We investigate (1) how ADM systems are used in day-to-day public administration decision-making and (2) how practitioners experience the introduction of the algorithmic tool. Eau Claire County has been chosen as it is one of the front-runner regions of “evidence-based decision-making” in CJ and has introduced the software COMPAS to provide risk assessment for pre-trial and post-trial decisions. The case therefore provides an illuminating illustration on how ADM systems affect decision-making on the ground. Based on a close reading of primary source material and qualitative expert interviews, we show how COMPAS has been introduced to provide risk assessments of offenders.

Our findings indicate that the main appeal of using the ADM system comes from two sources. *First*, decisions in the CJ system are often taken under high uncertainty concerning the outcomes while at the same time carrying the potential of far-reaching outcomes. If a judge decides to grant early release from jail, for instance, she is not able to give exact probabilities on the risk of reoffending, which places her decision under high uncertainty. Similarly, the decision entails the possibility of far-reaching outcomes, because the outcome may prove strongly harmful for society if the person on parole indeed reoffends. In this situation, gaining evidence about statistical correlations that help to predict probabilities for certain outcomes readily delivered by software and put into a score substantially reduces uncertainty. Applying the framework by Mousavi and Gigerenzer (2014, 1673), we argue that the availability of a algorithmically generated risk score changes the basic characteristics of the decision-making situation from fundamental uncertainty to statistical risk.

Therefore, even though decision-makers sometimes only possess incomplete knowledge about the inner workings of an ADM system, they are very open to using it.

Second, using software-based evidence provides a possibility for decision-makers to avoid blame for decisions that may have harmful consequences for society. This is why using the ADM has the big advantage to open up the possibility to deflect blame (to the software). While this may create problems of accountability and responsibility (Ananny and Crawford 2018; Veale et al. 2018), it is—from an actor’s perspective—an instrument of blame avoidance (Hood 2011). At the same time, using ADM also means that deciding against the “score” is probably rare, because doing so increases the stakes for being held accountable personally. The findings from our case study reveal that while decision-makers do not directly allude to blame avoidance, the possibility to rely on an algorithmically generated score (and to deflect blame) seems to affect their decision-making: risk-averse strategies, such as incarceration when in doubt, have been replaced by a strong reliance on the risk score as created by COMPAS. In sum, our results point out that it is at least equally important to think about the consequences the use of ADM systems has for the broader decision-making context as to evaluate the quality of the technical features of such systems.

The remainder of this paper is structured as follows. In the next section, we briefly discuss the state-of-the-art on the use of ADM systems in public administration and relate them to our theoretical framework, which is rooted in the literature on risk and uncertainty as well as on blame avoidance (Hood 2002, 2011; König and Wenzelburger 2014). The third section presents the case study and discusses the implications; the final section concludes.

ADM systems in public administration: putting actors center stage

Algorithms and public administration: from the laboratory to messy reality

In recent years, algorithms have made a fast career in public administration. Today, ADM systems are used in different contexts and support bureaucrats when detecting tax fraud (Botelho and Antunes 2011), assigning future students to universities (Grenet 2018; van Zanten and Legavre 2014), matching job seekers to training schemes (Desiere et al. 2019; Fröhlich and Spiecker 2019) or calculating the risk of reoffending for early release from a prison sentence (Berk 2017; Berk et al. 2017). While many public administration scholars emphasize the chances that big data and automated pattern detection entails and see the bureaucracy on the path toward “digital era governance” (Margetts and Dunleavy 2013), critical voices emphasize the lack of transparency and accountability of algorithms (e.g., Mittelstadt et al. 2016; Ananny and Crawford 2018; Zweig et al. 2018). These studies maintain that ADM systems may even produce biased decisions—not only because they incorporate certain values (which may be biased, e.g., Hildebrandt 2016; Yeung 2018), but also because they learn from input data and reproduce the biases found in this data, e.g., concerning ethnic or gender inequalities (Barocas and Selbst 2016; Lepri et al. 2018).

While the debate on these important issues is vivid (Singh et al. 2018), only more recently have scholars turned their interest toward non-technical issues that involve the question of how ADM systems are actually implemented in the messy real-life decision-making contexts of public administration and how the use of ADM transforms these systems. From a policy-science perspective, this is perhaps an even more relevant aspect than the technical features of the algorithm itself. Although it is certainly true that quality and

fairness features of ADM systems need to be discussed in terms of their lawfulness and ethical principles, the consequences of the introduction of ADM systems in society depend crucially on their implementation on the ground. If, for instance, an ADM system has been bought by a department to inform decision-making in a certain area but is hardly applied in everyday processes, consequences for citizen's lives are minimal—with biased or not biased, ethical or unethical, transparent or intransparent ADM systems.

To date, although the necessity to understand algorithms as part of social processes has been emphasized repeatedly (Beer 2017), this important aspect of how ADM systems are implemented in everyday bureaucratic decision-making has only been analyzed in a handful of recent studies. Veale et al. (2018), for instance, have interviewed public sector practitioners in five OECD countries on the use of machine learning systems in their everyday life. While their results suggest that some practitioners are indeed well aware of the ethical considerations that come with the introduction of algorithms, they also conclude that “those interested in transformative impact in the area of fair and accountable machine learning must move toward studying these processes in vivo, in the messy, socio-technical contexts in which they inevitably exist.” (Veale et al. 2018: 10). In a similar vein, van der Voort et al. (2019) have examined two concrete implementation examples and focused on how ADM systems have been implemented in a predictive policing case in a Dutch city as well as in a project on digital traces in Milan, Italy. Their main interest was to investigate how big data is used for public decisions and whether political actors or data analysts use the new decision-making contexts to further their own interests. They find that it is important to systematically analyze the institutional setting that shapes the decision-making process and point to increased possibilities for data analysts and decision-makers to pursue their own interests (van der Voort et al. 2019: 36). Such an interpretation runs against a functionalist view of public decision-making processes, which stresses the instrumental use of knowledge by competent actors for solving given problems.

In this paper, we take these most recent studies on how algorithms are used in public administration as our starting point. Building on two strands of literature on political decision-making—decisions under uncertainty and blame avoidance theory—we analyze whether the introduction of ADM systems can be explained by motives of uncertainty reduction and blame avoidance by political and bureaucratic actors. To this end, the next section will briefly summarize the theoretical framework and how it can be applied to our case, before we present our case study.

Algorithms and decision-making in a context of uncertainty and fatal consequences

In their conceptualization of how ADM systems affect administrative decision-making, van der Voort et al. argue that in order to get a full understanding of how algorithms affect outcomes, we need to provide a more adequate model the decision-making process that takes actors seriously (van der Voort et al. 2019: 29). Policy scholars have for many years pointed out that decision-making is hardly functionalistic and only geared toward problem-solving, but often follows a much more erratic process in which certain actors, such as policy entrepreneurs, have a key role in defining problems, setting them on the agenda and attaching policy solutions to them (Kingdon 2003; Herweg et al. 2017). The key question therefore is what the motives and preferences of bureaucrats and political actors are with regard to the use of ADM systems in administrative decision-making processes in general, and, more concretely on our case, in the area of criminal justice.

Decision-makers in public administration are—as every human—keen on reducing uncertainty and ambiguity when they take decisions (Gajduschek 2003, 715–717). However, doing so can be a difficult task, as information is sometimes incomplete and sometimes so overwhelming that it cannot be easily processed by human in a reasonable amount of time (Jones and Baumgartner 2012; Boin 2009). Confronted with such situations, decision-makers have therefore mostly relied on heuristics (Tversky and Kahneman 1974; Vis 2018) or on pragmatic abduction (Ansell and Boin 2019) in their standard practices while, at the same time, following established guidelines and rules that put such decisions on a solid legal foundation and assure accountability (Bovens et al. 2014). Such “procedural rationality” that is behavior which is the “outcome of appropriate deliberation” (Simon 1976, 67) is therefore key in uncertain contexts which are the rule in administrative decision-making.

With the introduction of ADM systems, this emphasis on procedural rationality, satisficing and reliance on heuristics may change, however, because the decision context changes from one of uncertainty into one of risk (see Table 1). Following the famous distinction by Knight (1921), uncertainty relates to situations in which an actor does neither know the outcome of her decision nor the odds of the outcome (“unknown unknowns”). This is when procedural rationality and heuristics are key. In contrast, situations of risk occur, when an actor does not know what happens, but is able to calculate the odds of the outcomes (“known unknowns”). This risk calculation can be done analytically, when probabilities are known (second line in Table 1), or through statistical analysis, when registered correlations in data yield regularities that are used to predict outcomes (third line in Table 1).

In the context of criminal justice, for instance, decision-makers—prosecutors, judges, state attorneys and bureaucrats in the CJ administration¹—cannot be sure about the probability with which a person may reoffend once released from jail or set free on bail. There simply is no way to calculate this in an appropriate way, which is why administrative rules and procedures are crucial. Hence, for decennia, judges and other actors in the CJ system have followed strict rules and standard procedures in these cases and relied mainly on expert knowledge (mainly delivered by reports from psychologists, social workers and others) to inform their decision. With algorithmic tools, this situation of fundamental uncertainty changes into one of statistical risk. If statistical evidence generated by algorithms provides aggregated risk scores that indicate a statistical risk of recidivism, the decision context in which the administrative actor² has to decide changes from one of uncertainty to one of statistical risk (line 3 in Table 1). This is a particular far-reaching transformation of the decision situation if the actor does not need to collect the data herself, but gets an output from a scoring algorithm or a classifier which is easy to interpret. The quantified empirical evidence can then be interpreted as “scientific objectivity”, and “provides an answer to a moral demand for impartiality and fairness” (Porter 1995, 8). In fact, while it is clear for a social scientist that such “risk assessments yield probabilities, not certainties, and that they measure correlations and

¹ All these actors take part in the decision-making process, at different points in time and to different degrees. We will detail below how this plays out in our case study.

² This is true on all levels of decision-making in the CJ system—for attorneys, lawyers and judges who decide about whether to hold a suspect in contempt awaiting trial as well as for bureaucrats that prepare a report for early release of a convict. We do not make any claims here about political decision-makers in the narrow sense, such as Members of State Congress, Mayors or Ministers in Charge of Criminal Justice as these are usually not concerned with the concrete decisions we are interested in (such as parole, bail, etc.).

not causations” (Završnik 2019, 10), this may be much less clear for a practitioner on the ground who is happy to receive additional information. Such information cannot be simply ignored but represents an anchor for any further interpretation by the human. This insight meshes well with findings about an “automation bias” (Dzindolet et al. 2003), which indicate that humans seem to consider decisions generated or supported by computers as overly trustworthy.

A second point has to be added to these considerations: The question of how far-reaching the consequences of decisions in a certain policy domain are and whether they can have negative consequences for the responsible decision-maker. In fact, scholars of public administration and political scientists have both found actors are particularly reluctant to take decisions that may have harmful consequences. In such situations, blame-avoidance strategies are used to delegate, blur or shift responsibility for the decision—just in case that it turns out to have negative consequences (Hinterleitner 2017; Weaver 1986; Vis and Van Kersbergen 2007; König and Wenzelburger 2014; Hood 2011). This is true for unpopular welfare state cutbacks (Jensen et al. 2018; Wenzelburger et al. 2019; Vis 2009) as well as for decisions in the area of crime (Hinterleitner 2018). In fact, blame-avoidance behavior is a much more common phenomenon and can be expected to structure the public decision-making process in many cases, and especially when risky decisions are involved (Hood 2011).

Algorithms can be seen as a welcome blame-avoidance instrument on which blame for decisions with unpopular outcomes can be shifted. In fact, Christopher Hood (2011) identifies “formulae, algorithms, computer programs” and several other instruments as “policy strategies” to avoid blame (Hood 2011, 93), because they limit the responsibility or even the formal liability for a decision, by limiting the decision-maker’s discretion. Whereas detailed protocols (“playing it by the books”) has always been an example for such a strategy, algorithms even take this idea one step further, because they provide a score or a classification result that indicates what decision to take. Clearly, if a bureaucrat can justify her decision with the recommendation of an ADM system, blame avoidance is almost perfect.

From this consideration follows that the extent to which political actors resort to such strategies very much depends on the policy at stake. Vis and van Kersbergen (2007) have nicely explained how blame avoidance can be linked to the psychological mechanism of negativity bias (Kahneman and Tversky 1979). Hence, the more risky and uncertain a decision seems to a decision-maker, the more important blame-avoidance techniques are. Applying this framework to the field of criminal justice yields a clear result: In this area, decisions taken by public administration are highly risky, because they concern the security of the citizens (if, for instance, an early released offender commits a crime, or the police decides not to control a terrorist). Moreover, crime and insecurity are part of a class of problems that are highly mediatized (Cere et al. 2014). Hence, the chance of a malign decision not to be discovered is low. Blame-avoidance strategies should therefore be of utmost importance for decision-makers that want to shield themselves from negative repercussions in case of a decision that turned out to be harmful (Welsh et al. 1990).

Taking these two arguments together, it seems straightforward to expect that decision-makers in the judicial system—both publicly elected officials such as attorneys or judges, and bureaucrats in the CJ administration—will report that the introduction of COMPAS has transformed the way how they make decisions. Moreover, from our theoretical considerations, we can derive two more specific questions that will guide our empirical exploration. First, we explore *how the involved persons experience the uncertainties and risks related to their decision*. And second, we investigate *whether and to what extent blame avoidance is mentioned in the actor’s description of the changed decision-making context*.

Table 1 Uncertainty and algorithmic decisions. *Source:* adapted from (Mousavi and Gigerenzer 2014: 1673)

Nature of unknown	Decision process	Result
Risk (a priori, stochastic) “known unknown”	Deductive (probabilities)	Objective odds
Risk (statistical)	Inductive (statistical inference, algorithmic predictions)	Correlational patterns, predictions of effects
Uncertainty (fundamental) “unknown unknown”	Heuristic	Intuitive satisficing

ADM systems in the criminal justice sector: the case of Eau Claire

In order to study how evidence policy-making via algorithms is seen by practitioners as an alteration of their decision-making context, we have studied the implementation of the COMPAS risk assessment software in the Criminal Justice system in Eau Claire County, Wisconsin. The case selection is driven by two considerations. First, the CJ sector in the USA is, arguably, the most advanced in terms of implementing ADM in real-life settings (see Kehl et al. 2017). Therefore, we can expect that discussions about how to implement ADMs and the repercussions on the broader decision-making context are most advanced. Second, Eau Claire County, Wisconsin, was chosen as it is one of the front-runner regions of “evidence-based decision-making” in criminal justice. Therefore, actors can be expected to have at least some longer experience with the working of ADMs and have integrated these systems in their everyday routines. This is important, if we want to study how ADM changes decision-making processes in the CJ system and whether blame avoidance actually plays a role.

Methodologically, our case study is based on a close reading of secondary sources and primary documents that we have obtained while visiting Eau Claire. Some of this primary material was internal so that we could only inspect the documents and take field notes. Document analysis was important throughout the whole process of this case study. The decision to choose Eau Claire County, for example, was made when we discovered through media research, the county’s special role as one of forerunners in implementing “evidence-based decision-making” in criminal justice. The analysis of documents that were given to us during the interviews was helpful to use some of the additional information in the next interviews. Moreover, these documents helped us to check some statements of the interviewees for the analysis of the case. This was especially done when interviewees were not certain about time sequences during the implementation process or when they referred to those documents in particular. Moreover, we have conducted six qualitative expert interviews. While the interviews also touched upon the policy-making process that led to the introduction of COMPAS in Eau Claire (e.g., the EBDM process), the main focus was on the concrete implementation of COMPAS in the everyday decision-making of the practitioners. However, the questionnaires were adjusted depending on the interview partner’s expertise. Moreover, we also discussed information obtained from the desk study of secondary sources with the actors in order to evaluate in how far issues raised in official documents were actually relevant for the practitioners on the ground.³

³ We did not directly ask questions about the sensitive issue of blame-avoidance behavior in the interviews. Instead we asked the interview partners to describe the positive and negative changes that the implementa-

The interviewees were chosen by their expert status in the local criminal justice system of Eau Claire County. Since the aim of the case study was to discover how practitioners in the field implement ADM systems into their daily working routines, the interviewees were (1) elected officials in court, (2) certified professionals who represent natural and juristic persons in legal matters and (3) professionals in the county administration.⁴ Moreover, on the state level, (4) an elected representative was interviewed on his impression concerning the implementation of COMPAS in local criminal justice systems in Wisconsin. In addition, the expert's role in the implementation process was an important aspect for the choice. Since the implementation of COMPAS was closely linked to the implementation of the national EBDM program, preliminary research of the Eau Claire case showed that most of the experts were also involved in the decision process concerning the question if and how to implement ADM systems into the local criminal justice sector. The experts were lawyers, bureaucrats in the CJ administration and high-level public officials. As access to involved actors proved to be rather difficult in several cases, all interviews are anonymous. Based on this empirical evidence, we are able to provide an in-depth insight into the dynamics during the implementation process of COMPAS and can describe how the involved actors reflect on their role in this process. In this section, we will therefore first provide an overview of the implementation of ADM in Eau Claire County. This will be followed by a brief discussion of the risk assessment tools used in the county. In a third step, we will discuss to what extent our theoretical expectations are supported by the case.

The implementation of ADM in Eau Claire

The first attempt to implement evidence-based decision-making in the CJ system of Eau Claire County has been the use of the so-called “Wisconsin Risk Need Assessment Instrument” (Henderson and Miller 2013). The tool's severe limitations in terms of predicting risks of recidivism brought up the necessity to find a new tool with higher predictability rates⁵ (Henderson and Miller 2013). About the same time, the National Institute of Corrections launched an initiative to “build a systemwide framework that [should] result in more collaborative, evidence-based decision-making (EBDM) and practices in local criminal justice systems” (National Institute of Corrections 2017: 9). Based upon the desire to measurably reduce recidivism rates in crime, the strategy was created to standardize information procedures at arrests and in court proceedings and equip professionals in the county administration and the CJ system with ADM tools in order to support their decisions. The strategy, referred to as “The Framework”, consisted of six phases overall.

Phase 1 started in May 2008 and ended in March 2010. During that period, counties in the USA were asked to hand in their applications to participate in the initiative's work. Participating counties were asked to create so-called “EBDM action groups”. In Eau Claire County, the action group was formed as a subcommittee of the Eau Claire Criminal Justice Collaborating Council (CJCC)—a group that had already existed since 2006 to enhance public safety through system and community collaboration, to maintain and

Footnote 3 (continued)

tion of COMPAS brought. Moreover, every interviewee was asked how their individual working routines have changed since the implementation of COMPAS.

⁴ As we have guaranteed anonymity to our interviewees, we cannot go into more detail concerning their positions.

⁵ Expert interview, June 25, 2019.

establish effective rehabilitation programs, and to foster innovative correctional programs. The council consisted of 10 members from the criminal justice system such as judges, district attorneys, public defenders, and members of the local police department and of three additional citizen members—and the fact that it already existed since several years seems to have been a major reason for the selection of Eau Claire as one of seven counties to participate in the EBDM process and to implement the strategy under technical assistance by the National Institute of Corrections (National Institute of Corrections 2017, 13).

This newly formed EBDM action group brought together all actors who were working in the local criminal justice systems—such as police officers, state defenders, district attorneys, and lawyers. The members were asked to review their current working routines as well as to analyze the current prison population in terms of their risk to reoffend. Having that knowledge at hand, actors then defined the guiding principles to implement evidence-based decision-making practices and structured the steps that needed to be taken in order to optimize their current working routines following the guidelines.

Guided by the NIC, Eau Claire County decided to implement COMPAS as their new risk assessment tool in 2009. Initially testing the functions of COMPAS, it was first applied to the Eau Claire jail population:

I needed to get this process moving, so we entered into the contract with COMPAS in 2009. I believe it was and said Let's try it out on our jail population. So my program director at the time [...], she interviewed everyone, she did a, she went to all the trainings and did an interview on every person that was in the jail, so we could get a baseline as to what the top criminogenic needs were. And then, we were able to take that data and then start targeting the types of programs we should be offering in the jail.⁶

The actual implementation of COMPAS into the decision-making practices within the local criminal justice system then happened during Phase 3, which started in August 2011. The NIC mainly provided assistance by educating the local EBDM teams through workshops and individual coaching. During these trainings, actors were given information “[...] about research-based policies and practices (“evidence-based practices”) and their application to decision points spanning the entire justice system” (National Institute of Corrections 2017: 2). However, it is unclear to what extent these measures were standardized and continued after the initial implementation phase so that new judges, district attorneys or defenders coming into the system also receive the relevant information. In addition, an agent from the NIC came to Eau Claire and helped the actors with interpreting statistics and choosing risk assessment tools to implement the evidence-based strategy in the local criminal justice system.

Besides COMPAS, which was used as a more encompassing risk assessment tool, a second tool—PROXY—was also introduced as a quick three-question tool used at the point of law-enforcement contact (e.g., interpellation by a policeman) for risk screening routines (see below). Phase 4 started 2 years later, in September 2013. While the seven counties were asked to review the steps they have been taking since the initial implementation, the NIC expanded the EBDM program to the state level. Phases 5 and 6 were then used to make further adjustments at both levels. Phase 6, which is still running at the moment, has terminated by the end of the year 2019.

⁶ Expert interview June 25, 2019.

The ADM tools PROXY and COMPAS

One of the main goals of the EBDM strategy was to implement an empirically based three level-classification within the decision-making system in criminal justice with respect to reoffending: (1) low risk (2) medium risk and (3) high risk of reoffending. This classification is used in pre-trial decision-making (e.g., in plea negotiations) and post-trial case management. The three risk groups are constructed on the basis of a score given by the assessment tool.⁷ Based upon the individual reoffending risk as calculated by the tool, the first decision in pre-trial would therefore be, which treatment a person should receive: diversion program, probation (with different treatment programs), or straight jail. While offenders with a low risk of reoffending normally do not get sentenced to jail, individuals with a medium or high risk might do so.⁸

The separation of the three risk groups involved two different ADM systems in Eau Claire County—(1) PROXY and (2) COMPAS. These tools differ both in terms of content—the questions on which the risk assessment is based and their associated goals—and in terms of their utilization in the local CJ system. As a pre-trial assessment tool (1) PROXY consists of three questions and aims at obtaining a first impression of a person's crime record. It basically serves to assess a person's risk to reoffend at its first arrest. In contrast, (2) COMPAS consists of 137 questions, is used pre- and post-trial and includes a risk assessment as well as a needs assessment, which is based upon eight criminogenic factors.⁹ Moreover, while PROXY is mostly used for misdemeanors or deviant behavior (drug abuse), COMPAS is often used in felony cases. According to a professional in the administration of Eau Claire County, a COMPAS assessment is often done as an additional assessment to get a comprehensive impression on a person's deficiencies, which might have caused the criminal behavior (needs assessment).¹⁰ This assessment will then be used to determine the treatment program. Those COMPAS assessments will only be done, in cases where the PROXY identified the risk level to be at medium or high risk. Finally, while PROXY is always used when a person gets arrested and does not need any approval by the suspect, the COMPAS assessment will only be done after the consultation with a lawyer and the following approval by the suspect.¹¹ When COMPAS was implemented in Eau Claire as part of Phase 3 of the EBDM program, it was therefore used as a complementary system to PROXY, mainly for those individuals that were classified as medium and high risk by PROXY.

⁷ In the case of PROXY, only three indicators are taken into account: age, age of first arrest, number of prior arrests.

⁸ This separation of groups of offenders is based on empirical research, showing that people with a low risk to reoffend tend to self-correct in their natural living environment and with the help of friends and family. Sentencing low-risk individuals to prison would, in fact, have an opposite effect as the low-risk individuals will be in contact with offenders with a medium or high risk (see: Lowenkamp et al. 2006).

⁹ The needs assessment includes the following criminogenic factors: (1) cognitive behavioral, (2) criminal associates/peers, (3) criminal involvement, (4) criminal opportunity, (5) criminal personality, (6) criminal thinking self report, (7) current violence, (8) family criminality (Northpointe 2019).

¹⁰ Expert interview, June 24, 2019.

¹¹ In the very beginning of the introduction process of COMPAS, the public defenders could decide whether to give the COMPAS results to the other actors involved in the respective process.

Algorithms as instruments to reduce uncertainty and to avoid blame

Building on the differentiation between situations of risk and uncertainty as well as on the concept of blame avoidance, we have theorized that the introduction of ADM systems in the CJ sector may be experienced by the involved actors—public officials and bureaucrats—as a transformation of their decision-making situation because of two reasons. First, the availability of an algorithmically created risk score transforms a situation of fundamental uncertainty in one of statistical risk; and second, it enables decision-makers to relate their choices on evidence delivered by an algorithm, which creates a possibility to shift the responsibility to software. The following analysis is organized along the two guiding questions that we have derived theoretically above. We will first discuss how the involved persons experience the introduction of COMPAS in relation to the uncertainty of their decisions; in a second step, we explore to what extent actors mention blame-avoidance opportunities.

ADM systems and uncertainty reduction

Our first and most general expectation has been that actors in the CJ system can be expected to see the use of COMPAS as reducing uncertainty in decision-making therefore leading to decisions of higher quality. The interviews with the decision-makers indeed point to a generally rather positive view of ADM systems. In two interviews with persons from different levels of hierarchy in the CJ system, the respondents described that before the use of COMPAS, they simply had no clue about the persons incarcerated in the county jail.¹² What was emphasized most strongly was that the information included in COMPAS is used as an additional input when making decisions thereby reducing uncertainty. One of the actors explained the situation as follows:

And that's to me, what in the end what this is all about: Be as informed as possible to make as informed decision as possible. So at least you can say: With a clear conscience, you know I utilized the information that is available to me versus: Well I, dart board, I just guess and hope I get it right. And it doesn't mean that you always get it right but I mean you're gonna get it more right and this is what the research tells everybody. That's one of the core concepts of evidence-based decision making. If you combine everything that you always do, hard work, analysis, the facts of the case, all of these types of things, the evidence, combine all of that with the use of validated risk assessment tools, your outcomes will consistently outperform just using the things you've always used without the tool or vice versa and so that's it.¹³

The quote shows that actors seem to be well aware of the important repercussions that their decisions have on a person's future and that for such a decision every additional source of information is welcome. The practitioner interviewed clearly believes that the evidence generated by using COMPAS along other available information will produce better decisions—and a clear conscience of the decision-maker.

¹² Expert interviews, June 24, 2019 and June 25, 2019.

¹³ Expert interview, June 24, 2019.

Moreover, the actors not only welcome the increased availability of information via COMPAS, but link it directly to their decision—and the risks involved in these decisions. As one actor put it:

And if they are aware of all of these concepts that we've been talking about, and they use those tools to help make decisions regarding individuals, to me that's a really good thing. Versus a judge who sits upon the bench and takes the [...] dart and throws it at the dart board and says: 'Well here is, here is what we are going to do for you.' I don't know what that means. It's a guess. I don't like guesses. You know, you like certainty. And right, wrong or different. These validated risk assessment tools bring a level of certainty to the analysis of an individual that we've never had before.¹⁴

In this quote, it is evident how the practitioner feels about using COMPAS. It is seen as a tool to reduce uncertainty—or even more than that: as a tool, that brings a “level of certainty”. Therefore, this statement strongly underscores that the statistical evidence provided by an algorithm is interpreted not only as a statistical probability (what it actually is), but as more than that.

Given that the COMPAS as well as the PROXY score is used to generate three levels of risk (high, medium, low), it seems crystal clear from the quote that these three classes of outcomes are seen as more than just indications of statistical probability. The availability of such information is seen as a relief by the actors in their highly delicate and uncertain decision contexts.

It allows the judges and the DA to craft a better plea negotiation, meaning: Are they better suited and settled straight jail time or probation. [...]. So what we try to do is: Work it through the EBDM process and talk with the judges, the DAs and public defenders and the Department of Corrections and say: Are we comfortable with structuring sanctions that does not include probation for those lower risk people and diverting or deflecting those people that are coming up low risk on the PROXY. [...]. So if you are low risk on the PROXY and you are in contact with law enforcement, that risk tool will then trigger that diversion of that person from the system, so that person is never charged.¹⁵

In a similar vein, another expert shows how far the categorization of people in three risk classes as put forward by the system has affected the entire decision-making context. It is considered to produce “better” decisions in the sense that the new information by COMPAS has led the actors to step back from an over-cautionary approach that they seemed to have used before. For example, a clear consequence of the introduction of COMPAS in the post-trial phase has been a reduced prison time for persons scoring at “low risk” according to the ADM system:

So the whole idea EBDM and the COMPAS and some of these assessment tools was to try and limit the number of people coming in, keeping those low-risk people out, keeping them in the community, giving them the treatment and services they needed in community and not in custody We didn't have a good measure as to who is in our jail and what their risks were. [...] We decided to go with the COMPAS after doing

¹⁴ Expert interview, June 24, 2019.

¹⁵ Expert interview, June 25, 2019.

some research because it was automated, it would produce a chart as to the top criminogenic factors and then it would give a risk on that person. So we could focus the risk, focus the service on those that are the medium and high risk and not focus our services on the low risk because they are gonna self-correct naturally, that's what the research says.¹⁶

Actually, interviewees from all levels of administration emphasized the fact that the risk scores delivered by the tool were “research-based”. This also indicates that the users see the tool as providing reliable evidence of high quality that guides their decision-making.

Interestingly, all actors still maintained that the COMPAS scores did not involve an “in and out”-decision and that the decisions by actors in the judicial system still “started and ended with facts”. Some of them still emphasized the discretion of human actors and the non-binding role the risk scoring by the ADM system played. At the same time, however, given that the evidence generated by the software was seen by all of the actors as positive because it reduces uncertainty and helps the actors in their efforts to get it right, it seems obvious that the introduction of COMPAS into the decision-making process massively transformed the context in which judicial decisions were taken. This change in the decision-making process in criminal justice is clearly acknowledged by the actors, as shown in the following statement by one of the practitioners interviewed:

After we implemented some of these evidence-based practices, same two parties would go in front of the judge. The judge says: ‘What’s the agreement?’—‘Judge, the agreement is to put this person on probation for 2 years with all these conditions. We’re asking you to follow the agreement.’

The first question, the judge is gonna ask us is: ‘Was a COMPAS done?’ That never happened before. Now, you may say: ‘Yeah, it was done and here it is.’—‘Okay, great.’ Then the judge can look at that. Or if you said ‘No’, you better have a reason to tell the judge why that step wasn’t taken¹⁷

The evidence on this first guiding question therefore indicates that the use of algorithmic tools in the CJ system is mainly seen as a way to reduce uncertainty in the decision-making process. COMPAS is praised by all interviewed practitioners to deliver new information and most of them seem to follow the categorization of low, medium, and high risk as put forward by the software. While it is true that most interviewed experts also emphasized that other sources of information were still consulted, and only the joint interpretation of all available information would lead to a decision, the emphasis of the research-based and validated character of COMPAS as well as the importance of the three-category risk classification which was mentioned in all interviews points to the central role played by the risk assessment tool in the entire decision-making process. This important impact is palpable in the following exchange with a practitioner on the question of how COMPAS changed the everyday routines of the actors in the Criminal Justice system:

Interviewee: [...] We were using it [COMPAS, the authors] for pre-trial and one way it really helped was on the first offense drug deliveries especially Marihuana. We were able to get those low-risk people on like diversion agreements and deferred agreements as opposed to convicted as felonies; And so that was based on their low-risk COMPAS and that was something we weren’t able to do before. So I think that

¹⁶ Expert interview, June 24, 2019.

¹⁷ Expert interview, June 24, 2019.

was one of the main positives I saw with the COMPAS.

Interviewer: Okay, you weren't able to do that before?

Interviewee: It was hard to get the DAs to agree to like deferred on only felony delivery charges because they didn't know who were the real deliverers and who were not. So what the COMPAS..., if we could show that they were low-risk individuals, they were willing to reduce some of those like deferred agreements where they would defer at or make them misdemeanor or something less serious.

Interviewer: So you give data to the DA?

Interviewee: Yes, the DA gets the COMPAS results, yes. And then they would look at it and see the person was low risk and then they would give you a better plea agreement.¹⁸

ADM systems and blame avoidance

Our second exploratory question concerned the issue whether users of COMPAS in Eau Claire mention that relying on an algorithm protects them from blame. We theorized that this may be an important consideration because decisions in the CJ system involve questions of recidivism and may have important consequences for society (Welsh et al. 1990). The use of such “anticipatory blame avoidance” (Hinterleitner and Sager 2017) strategies would therefore shield the decision-makers from blame for malign decisions.

Empirically, however, the case study and the interviews with the practitioners provide no clear-cut evidence for the relevance of blame avoidance. In fact, most actors we talked to emphasized the reduced uncertainty and the additional information provided by the ADM as a tool helping them in their effort to get it right (see above), but did not mention the possible blame involved with malign decisions. This comes as a surprise, because the literature on blame avoidance in the CJ system (Eckhouse et al. 2019, 203) argues that some actors, such as judges, should be very concerned about blame avoidance, because they are at a very exposed position and blame can be attributed rather easily. However, even judges asked about how COMPAS transformed their decision-making, did not mention blame avoidance as part of their rationale, but explained the changes as follows:

I think that judges want to do the right thing. We may not always be right but we want to do the right thing. And so the big picture evidence-based decision making is smarter sentencing. You know things like using risk assessment tools to figure out whether someone should go on probation or not. We don't want to put low-risk people on probation because they will then be over supervised. They don't need all of that. They interfere with their natural ability to get over this if we overload them with programming they don't need, we negatively impact their lives, their families, their jobs and so forth. And of course: Mixing low-risk people with high risk people winds up influencing low-risk people in a negative way. They learn to be better criminals perhaps, if you want to use that phrase. So I think that generally speaking, evidence-based decision making is just doing what the research tells us is effective. And once again: It's becoming the norm. It pretty much is the norm [...] now.¹⁹

¹⁸ Expert interview, June 25, 2019.

¹⁹ Expert interview, June 24, 2019.

In this quote, the interviewed judge describes the role of COMPAS as helping them achieve smarter sentencing. Again, research and quality of the evidence is alluded to—whereas blame does not seem to play a role.

In addition, the availability of the risk score seems to have helped to come to more time efficient decisions. Asked about how COMPAS affected plea bargains, one interviewee pointed out that the evidence generated by the algorithm actually helps reaching an agreement more quickly because it narrows down the key points of negotiation:

What the risk assessments and COMPAS have done is taking discussions [...] closer to that window because we get objective information about a person, and so we don't have to start discussions way out of our spectrum [...]. Because, for example, if a COMPAS shows someone is low risk and that they don't have any criminogenic needs... Absence of some really compelling factors... That person is not going on probation, okay... In the past without that COMPAS assessment and showing that objective analysis of that person, the prosecutor and defense attorney may have argued for months just to get over the "probation or not"-question. Whereas with that COMPAS showing that, so definitely, the parties may be like: Okay, we know probation is off the table. Now, what else do we talk about in terms of a disposition? A street sentence? Some community service? A fine? Whatever it is, it helps narrow the issues for discussion. Just like if someone is high risk on everything and scores high on every criminogenic need, well, ... you know that there is a lot going on with that person."²⁰

Hence, in sum, the evidence that judges are increasingly open to refrain from probation or jail sentences (e.g., in pre-trial bail decisions) *because* they can base their judgment on the score provided by COMPAS or PROXY (indicating that a person belongs to the low-risk category) is rather weak. While several interviewees acknowledges that key actors—judges as well as district attorneys—seemed now more open to use alternative sentences since the introduction of COMPAS, they did not relate this change to blame-avoidance behavior, but rather to the higher probability of the actors to follow the risk score, and not to incarcerate a person which belongs to the low-risk class according to the algorithm.²¹

Conclusion

In this article, we have shown that the increased use of algorithms to inform decisions in the CJ system has transformed the decision-making process. Based on evidence from a case in Wisconsin, our article has illustrated that actors in the CJ system from different levels of hierarchy—public officials as well as bureaucrats in the administration—are rather open to use ADM systems in the decision-making process. While the practitioners on the ground emphasize that the software is just an additional tool in a toolbox and does not solely drive in-and-out decisions, it is palpable how deeply the risk scoring using three categories has transformed the decision-making process. In fact, using the software is “pretty much the norm”²²

²⁰ Expert interview, June 24, 2019.

²¹ Evidently, it could also be the case that the actors simply did not want to tell that blame avoidance played a role in their considerations.

²² Expert interview, June 25, 2019.

now and all interviewed persons have used the categorization of three risk levels as put forward by the software when discussing how recidivism risk is assessed.²³

The reason for this positive stance toward using algorithms clearly is that practitioners welcome additional information that helps them to come to a decision in a context characterized by high uncertainty.²⁴ The actors interpret the scores delivered by the ADM system as being research-led and evidence-based, which makes them a valuable reference point for their decision. The interviews suggest that, indeed, the context of the situation changes from one of fundamental uncertainty to one of statistical probability in which decision-makers use the empirical evidence provided by the ADM to calculate risks. This is exactly what we theorized. In contrast, blame avoidance has not been mentioned explicitly by the practitioners as a reason to follow the advice from algorithms. Hence, it seems that the main impetus for the use of algorithmic evidence indeed is the perceived reduction in uncertainty.

With these findings, our article contributes to the emerging but still sketchy literature that investigates how ADM systems are implemented in real-life decision-making processes and what consequences they have. They speak at least to three bodies of literature. First, in the realm of criminal justice, our findings, according to which the use of risk assessment tools might have reduced the share of probation decisions for low-risk offenders, speaks to several case studies that report similar tendencies (see the discussion by Stevenson 2018, 337–340). It would be definitely important to study more systematically what the downstream consequences of the introduction of algorithmic tools were—be it by means of statistical large-N studies or through experiments. Second, the positive view of the interviewed actors here speaks to the more general PA literature on algorithm aversion and algorithm appreciation which report rather inconclusive findings on how open decision-makers are to use ADM tools in public administration (Burton et al. 2019). Whether the openness of the actors in our case has to do with the fact that the County has been successful in a nation-wide application process or with the long experience of some actors, can only be guessed from the case. More research on the conditions that lead to acceptance or rejection of ADM in public administration more generally is clearly needed.

On a more general level, we believe that the results from the case study can be used to inform further research on ADM in public policy beyond the narrow field of criminal justice inspected here. While criminal justice is certainly particular because most decisions have far-reaching consequences, the general context of risk and uncertainty does indeed characterize many decisions in public administration. If a frontline worker in an employment agency has to decide whether she will grant a certain retraining program to a job-seeker, this decision is less far-reaching than imprisonment but is equally uncertain—which is why we should

²³ The only critical aspects mentioned in the interviews related (1) to the matter of discrimination that may built-in into the ADM system due to the use of variables that may identify certain sensitive groups and (2) to the question whether the data used to train the dataset was actually a good proxy for the criminal context in Eau Claire, where the predictive algorithm was implemented. However, these concerns did not refer to the general advantage seen by all actors to increase the certainty of decisions.

²⁴ An additional aspect mentioned by some of the interviewees was the need to reduce the high costs of imprisonment (see also NIC 2017, 12). In several interviews, the experts related to the fact that when a new county jail was planned, a reflection about the composition of prison inmates was started and how the number, and hence, the financial burden of imprisonment could be reduced. However, this aspect of resources does not seem to have been the dominant consideration in the day-to-day decision-making processes we are interested in.

observe similar behavioral patterns. Following from this, at least two lessons from our case can contribute to the emerging broader public administration and public policy literature on algorithms. First of all, it seems critical to observe the real-life implementation of ADM systems in public administration in order to assess the consequences for decision-making. Second, as the changed process may have palpable downstream consequences on society, it is of prime importance to investigate not only whether the rules that are implemented *within* the algorithm are ethically defensible, but also to assess whether the *system as a whole* and the transformation of the decision-making process it involves is politically legitimate. In contexts where human discretion is involved because “soft” factors play a role, such as the credibility of an offender to change its habits, algorithms may on the one hand provide important information that makes a human decision more evidence-based; on the other hand, an algorithmically generated score may also be an important anchor point from which a human decision-maker will only rarely deviate. However, in order to answer such intricate question, we first need more evidence on how algorithms play out in real-life contexts. The present study can be seen as one step on this way.

Funding Open Access funding enabled and organized by Projekt DEAL.

Open Access This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons licence, and indicate if changes were made. The images or other third party material in this article are included in the article’s Creative Commons licence, unless indicated otherwise in a credit line to the material. If material is not included in the article’s Creative Commons licence and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this licence, visit <http://creativecommons.org/licenses/by/4.0/>.

References

- Ananny, M., & Crawford, K. (2018). Seeing without knowing: Limitations of the transparency ideal and its application to algorithmic accountability. *New Media & Society*, 20, 973–989.
- Angwin, J., Larson, J., Mattu, S., & Kirchner, L. (2016). Machine bias—There’s software used across the country to predict future criminals. And it’s biased against blacks. *ProPublica*. <https://www.propublica.org/article/machine-bias-risk-assessments-in-criminal-sentencing>.
- Ansell, C., & Boin, A. (2019). Taming deep uncertainty: The potential of pragmatist principles for understanding and improving strategic crisis management. *Administration & Society*, 51(7), 1079–1112. <https://doi.org/10.1177/0095399717747655>.
- Barocas, S., & Selbst, A. D. (2016). Big data’s disparate impact. *California Law Review*, 104, 671. <https://doi.org/10.15779/Z38BG31>.
- Beer, D. (2017). The social power of algorithms. *Information, Communication & Society*, 20(1), 1–13. <https://doi.org/10.1080/1369118X.2016.1216147>.
- Berk, R. (2017). An impact assessment of machine learning risk forecasts on parole board decisions and recidivism. *Journal of Experimental Criminology*, 13(2), 193–216. <https://doi.org/10.1007/s11292-017-9286-2>.
- Berk, R., Heidari, H., Jabbari, S., Kearns, M., & Roth, A. (2017). *Fairness in criminal justice risk assessments: The state of the art*. Philadelphia: University of Pennsylvania.
- Boin, A. (2009). The new world of crises and crisis management: Implications for policymaking and research. *Review of Policy Research*, 26(4), 367–377. <https://doi.org/10.1111/j.1541-1338.2009.00389.x>.
- Botelho, J., & Antunes, C. (2011). Combining social network analysis with semi-supervised clustering: A case study on fraud detection. In *Proceeding of mining data semantics (MDS’2011) in conjunction with SIGKDD*, pp. 1–7.

- Bovens, M., Schillemans, T., & Goodin, R. E. (2014). Public accountability. In M. Bovens, R. E. Goodin, & T. Schillemans (Eds.), *The Oxford handbook of public accountability*. Oxford: Oxford University Press. <https://doi.org/10.1093/oxfordhb/9780199641253.013.0012>.
- Burton, J. W., Stein, M.-K., & Jensen, T. B. (2019). A systematic review of algorithm aversion in augmented decision making. *Journal of Behavioral Decision Making*. <https://doi.org/10.1002/bdm.2155>.
- Cere, R., Jewkes, Y., & Ugelvik, T. (2014). Media and crime. In S. Body-Gendrot, M. Hough, K. Kerezi, R. Lévy, & S. Snacken (Eds.), *The Routledge handbook of European criminology* (pp. 266–279). London/ New York: Routledge.
- Desiere, S., Langenbucher, K., & Struyven, L. (2019). *Statistical profiling in public employment services (OECD social, employment and migration working papers, no. 224)*. Paris: OECD Publishing.
- Dunleavy, P., Margetts, H., Bastow, S., & Tinkler, J. (2005). New public management is dead—Long live digital-Era governance. *Journal of Public Administration Research and Theory*, 16(3), 467–494. <https://doi.org/10.1093/jopart/mui057>.
- Dzindolet, M. T., Peterson, S. A., Pomranky, R. A., Pierce, L. G., & Beck, H. P. (2003). The role of trust in automation reliance. *International Journal of Human–Computer Studies*, 58(6), 697–718.
- Eckhouse, L., Lum, K., Conti-Cook, C., & Ciccolini, J. (2019). Layers of bias: A unified approach for understanding problems with risk assessment. *Criminal Justice and Behavior*, 46(2), 185–209. <https://doi.org/10.1177/0093854818811379>.
- Einhorn, H. J., & Hogarth, R. M. (1981). Behavioral decision theory: Processes of judgement and choice. *Annual Review of Psychology*, 32(1), 53–88.
- Fröhlich, W., & Spiecker, I. (2019). Können Algorithmen diskriminieren? *djbZ Zeitschrift des Deutschen Juristinnenbundes*, 22(2), 91–93. <https://doi.org/10.5771/1866-377x-2019-2-91>.
- Gajduscheck, G. (2003). Bureaucracy: Is it efficient? Is it not? Is that the question? Uncertainty reduction: An ignored element of bureaucratic rationality. *Administration & Society*, 34(6), 700–723. <https://doi.org/10.1177/0095399702239171>.
- Grenet, J. (2018). Orientation postbac: Une question technique ou politique? *Administration & Éducation*, 159(3), 123–127. <https://doi.org/10.3917/admed.159.0123>.
- Henderson, H., & Miller, H. (2013). The (twice) failure of the Wisconsin risk need assessment in a sample of probationers. *Criminal Justice Policy Review*, 24, 199–221. <https://doi.org/10.1177/0887403411422410>.
- Herweg, N., Zahariadis, N., & Zohlnhöfer, R. (2017). The multiple streams framework: Foundations, refinements and empirical applications. In M. Christopher, S. Weible, & A. Paul (Eds.), *Theories of the policy process* (pp. 17–54). Boulder: Westview.
- Hildebrandt, M. (2016). Law as information in the Era of data driven agency. *The Modern Law Review Limited*, 79(1), 1–30. <https://doi.org/10.1111/1468-2230.12165>.
- Hinterleitner, M. (2017). Reconciling perspectives on blame avoidance behaviour. *Political Studies Review*, 15(2), 243–254. <https://doi.org/10.1111/1478-9302.12099>.
- Hinterleitner, M. (2018). *Policy controversies and political blame games*. Bern: University of Bern.
- Hinterleitner, M., & Sager, F. (2017). Anticipatory and reactive forms of blame avoidance: Of foxes and lions. *European Political Science Review*, 9(4), 587–606. <https://doi.org/10.1017/S1755773916000126>.
- Hood, C. (2002). The risk game and the blame game. *Government and Opposition*, 37(1), 15–37.
- Hood, C. (2011). *The blame game: Spin, bureaucracy, and self-preservation in government*. Princeton: Princeton University Press.
- Jensen, C., Arndt, C., Lee, S., & Wenzelburger, G. (2018). Policy instruments and welfare state reform. *Journal of European Social Policy*, 28(2), 161–176.
- Jones, B. D., & Baumgartner, F. R. (2012). From there to here: Punctuated equilibrium to the general punctuation thesis to a theory of government information processing. *Policy Studies Journal*, 40(1), 1–20. <https://doi.org/10.1111/j.1541-0072.2011.00431.x>.
- Kahneman, D. (2012). *Thinking, fast and slow*. London: Penguin Books.
- Kahneman, D., & Tversky, A. (1979). Prospect theory: An analysis of decision under risk. *Econometrica*, 47(2), 263–292.
- Kehl, D., Guo, P., & Kessler, S. (2017). *Algorithms in the criminal justice system: Assessing the use of risk assessments in sentencing*. Cambridge: Responsive Communities Initiative, Berkman Klein Center for Internet & Society, Harvard Law School.
- Kingdon, J. W. (2003). *Agendas, alternatives, and public policies, longman classics in political science* (2nd ed.). New York: Longman.
- Knight, F. H. (1921). Risk, uncertainty and profit. In *Hart, Schaffner and Marx prize essays*. Boston: Houghton Mifflin.
- König, P., & Wenzelburger, G. (2014). Towards a theory of political strategy in policy analysis. *Politics & Policy*, 42(3), 397–427.

- Lepri, B., Oliver, N., Letouzé, E., Pentland, A., & Vinck, P. (2018). Fair, transparent, and accountable algorithmic decision-making processes. *Philosophy & Technology*, 31(4), 611–627. <https://doi.org/10.1007/s13347-017-0279-x>.
- Lowenkamp, C. T., Latessa, E. J., & Holsinger, A. M. (2006). The risk principle in action: What have we learned from 13,676 offenders and 97 correctional programs? *Crime & Delinquency*, 52(1), 77–93. <https://doi.org/10.1177/0011128705281747>.
- Margetts, H., & Dunleavy, P. (2013). The second wave of digital-era governance: A quasi-paradigm for government on the Web. *Phil Trans R Soc A*, 371, 20120382. <https://doi.org/10.1098/rsta.2012.0382>.
- Mittelstadt, B. D., Allo, P., Taddeo, M., Wachter, S., & Floridi, L. (2016). The ethics of algorithms: Mapping the debate. *Big Data & Society*, 3(2). <https://doi.org/10.1177/2053951716679679>.
- Mousavi, S., & Gigerenzer, G. (2014). Risk, uncertainty, and heuristics. *Journal of Business Research*, 67(8), 1671–1678. <https://doi.org/10.1016/j.jbusres.2014.02.013>.
- National Institute of Corrections. (2017). *EBDM case studies: Highlights from the original seven pilot studies*. Washington: National Institute of Corrections (US Department of Justice).
- Northpointe. (2019). Practitioner's guide to COMPAS Core. Traverse City, MI. Retrieved from <http://www.equivant.com/wpcontent/uploads/Practitioners-Guide-to-COMPAS-Core-040419.pdf>.
- Pasquale, F. (2016). *The black box society. The secret algorithms that control money and information*. Cambridge: Harvard University Press.
- Porter, T. M. (1995). *Trust in numbers: The pursuit of objectivity in science and public life*. Princeton: Princeton University Press.
- Simon, H. A. (1976). From substantive to procedural rationality. In T. J. Kastelein, S. K. Kuipers, W. A. Nijenhuis, & G. R. Wagenaar (Eds.), *25 Years of economic theory: Retrospect and prospect* (pp. 65–86). Boston, MA: Springer.
- Singh, J. P., Kroner, D. G., Wormith, J. S., Desmarais, S. L., & Hamilton, Z. K. (2018). *Handbook of recidivism risk/needs assessment tools*. Hoboken: Wiley-Blackwell.
- Stevenson, M. (2018). Assessing risk assessment in action. *Minnesota Law Review*, 103, 303–384.
- Tversky, A., & Kahneman, D. (1974). Judgment under uncertainty: Heuristics and biases. *Science*, 185(4157), 1124–1131. <https://doi.org/10.1126/science.185.4157.1124>.
- van der Voort, H. G., Klievink, A. J., 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. <https://doi.org/10.1016/j.giq.2018.10.011>.
- van Zanten, A., & Legavre, A. (2014). Engineering access to higher education through higher education fairs. In G. Goastellec & F. Picard (Eds.), *Higher education in societies: A multi scale perspective* (pp. 183–203). Rotterdam: SensePublishers.
- Veale, M., van Kleek, M., & Binns, R. (2018). *Fairness and accountability design needs for algorithmic support in high-stakes public sector decision-making*. Paper presented at the CHI 2018, Montréal, Canada.
- Vis, B. (2009). Governments and unpopular social policy reform: Biting the bullet or steering clear? *European Journal of Political Research*, 48(1), 31–57. <https://doi.org/10.1111/j.1475-6765.2008.00783.x>.
- Vis, B. (2018). Heuristics and political elites' judgment and decision making. *Political Studies Review* (forthcoming).
- Vis, B., & Van Kersbergen, K. (2007). Why and how do political actors pursue risky reforms. *Journal of Theoretical Politics*, 19(2), 153–172.
- Weaver, K. R. (1986). The politics of blame avoidance. *Journal of Public Policy*, 6(4), 371–391.
- Welsh, W. N., Ponetll, H. N., Leone, M. C., & Kinkade, P. (1990). Jail overcrowding: An analysis of policy makers' perceptions. *Justice Quarterly*, 7(2), 341–370.
- Wenzelburger, G., Jensen, C., Lee, S., & Arndt, C. (2019). How governments strategically time welfare state reform legislation: empirical evidence from five European countries. *West European Politics* online first.
- Yeung, K. (2018). Algorithmic regulation: A critical interrogation. *Regulation & Governance*, 12(4), 505–523. <https://doi.org/10.1111/rego.12158>.
- Završnik, A. (2019). Algorithmic justice: Algorithms and big data in criminal justice settings. *European Journal of Criminology*. <https://doi.org/10.1177/1477370819876762>.
- Zweig, K. A., Wenzelburger, G., & Krafft, T. D. (2018). On chances and risks of security related algorithmic decision making systems. *European Journal for Security Research*. <https://doi.org/10.1007/s41125-018-0031-2>.