8 Algorithms in Government

This chapter is about the use of algorithmic decision tools by governments, specifically government agencies. Of course, *every* chapter in the book so far has (to some extent) touched upon government uses of AI and algorithmic decision tools. Consequently, you may wonder whether a chapter dedicated solely to the use of such tools by government agencies has anything new to add to the discussion. We think it does. The issues raised in previous chapters are relevant to how we assess the use of algorithmic decision tools by government agencies, but there are distinctive issues that arise from this use too, and it's worth giving them a fair hearing. It will help if we start with a clearer characterization of what those issues might be.

In *On the Wealth of Nations,* Adam Smith argues that specialization (or what he calls the "division of labor") is a key engine of economic growth. He illustrates this with the example of a pin factory, which makes pins consisting of a pointed metallic shaft and a flat metallic cap. He observes that "a workman not educated to this business" would struggle to make more than one of these pins per day "and certainly could not make twenty." But if the process of making the pins is broken down into a number of distinct processes, and if individual workers are trained and specialized in performing those distinct processes, a team of ten pin-makers can make "upwards of forty eight thousand pins in a day." And so, with specialization, you can dramatically increase the productivity of labor and, by extension, the "wealth of nations."[1] Specialization of this Smithian kind has long been embraced by the private sector. It is at the heart of most businesses and firms in existence today. This is also reflected in how the private sector makes use of technology. Where once it was teams of human workers who trained in specialized tasks, now it is teams of workers, machines, and algorithms that specialize in task performance. (We already caught wind of this shift in chapters 4 and 5.)

But the value of specialization is not just limited to productivity in the private sector. It's a general principle of social organization. As Thomas Malone notes in his fascinating book *Superminds*, the success of human civilization is largely attributable to our "collective intelligence," that is, how well we work together in groups to solve problems.[2] As the problems we confront get more complex, a good rule of thumb is to try to break them down into more manageable sub-problems and get specific individuals or organizations to specialize in the resolution of those sub-problems. They can bring their unique knowledge to bear on those problems.

This rule of thumb has been applied with vigor to the management of modern nation-states. The business of government has probably never been simple, but today it is a massively complex task. Governments create specialized agencies to solve all manner of societal problems. These include agencies dedicated to managing and regulating healthcare, the payment of welfare, finance, public and private transport, communications technologies, data gathering technologies, energy usage, environmental protection, food safety, drug safety, and so on. The swelling in both the number and size of these agencies is one of the hallmarks of the "administrative state" that has become the norm in liberal democratic regimes since the mid-twentieth century. These agencies are often created in response to particular crises and at moments of political convenience. Nevertheless, the advantages of specialization are often very real and tangible. No single politician or elected official could possibly govern an advanced, multicultural, industrialized nation without the help of specialized problem-solving agencies.

But this specialization creates a problem from a democratic perspective. The creation of specialized agencies changes the relationship between citizens and institutions of power. In particular, it risks *attenuating the legitimacy relationship* between them. What does this mean? In broad outline, the problem is this: specialized government agencies usually have the power to affect individual citizens in significant ways. They can deny rights and privileges, impose penalties and fines, and otherwise interfere with their capacity to live a flourishing life. To give an obvious example, a government agency that denies welfare payments to a person with no other source of income is doing something that could profoundly impact that person's capacity to live a good life. In liberal democratic states, the fundamental moral assumption is that power of this sort cannot be exercised unless it is *legitimate*.[3] What that means depends to some extent on who you talk to. Philosophers and political

theorists have identified many different "legitimacy conditions" that may need to be satisfied in order for power to be exercised legitimately. The most obvious way to ensure legitimacy is to get the consent of the person who might be affected by the exercise of power. This is the route favored by private enterprises when getting customers to sign contracts (though whether they live up to that ideal is a matter that was touched upon in the last chapter). But government agencies can rarely rely on such consent to legitimize their power. Governments don't behave like merchants selling wares to passers-by. They can't feasibly transact with every citizen individually (not as a rule anyway). Instead, governments conduct "business" for their populations as collective wholes. But this obviously runs the risk of acting against the express wishes of *at least some* members of these populations. So they rely instead on giving citizens a meaningful say in shaping the institutions of power.

The most obvious way to ensure that citizens get a meaningful say in shaping the institutions of power is to either consult directly with citizens on the construction of those institutions (e.g., through constitutional ratification and amendment) or to get citizens to elect *representatives* who make laws and shape institutions of power on their behalf. The most pared-down version of parliamentary democracy is the clearest illustration of this method of legitimation. An individual citizen votes for a representative and then the representative votes on specific legislative proposals in the parliament. If the citizen does not like the way the representative votes, they can hold them to account at public meetings and clinics or use the power of the ballot box to vote them out at the next election.

In practice, however, the creation of specialized government agencies often compromises these lofty ideals. As these agencies are typically created by parliaments and not directly by citizens through plebiscite they come into existence at one step removed from the most powerful source of legitimation: the direct will of the people. This might be acceptable if we assume that (a) the directly elected representatives in the legislative assembly are exercising their power to create such agencies legitimately and (b) they retain control over how those administrative agencies function. But this may not always be the case. Often specialized agencies are created in such a way that they are directly insulated from the whims and vagaries of electoral politics. This is often intentional. They are designed so that they have some independence from government and so are not swayed by the same short-term, interest-group-guided concerns of elected officials.

A classic example of this is the setup for the typical central bank. Though the form varies, a central bank will usually have ultimate control over a country's financial system, acting as both a lender of last resort to private banks and the "printing press" that controls the money supply. From long and unfortunate experience, we have learned that these powers can be abused if exercised by elected governments, and so we now favor an institutional design that insulates central banks from direct government interference. But this, of course, means that central banks have significant and wide-ranging "unelected power."[4] This leads to many people questioning their legitimacy, particularly in times of financial crisis. People rail against the "technocrats" running these institutions and the power they have over our lives.

This example is, of course, just the tip of the iceberg. As specialized agencies proliferate, the problem of unelected power becomes more pervasive, and it can be exacerbated when the agencies are given considerable discretionary authority to design and enforce policies, and are allowed to outsource or subcontract their powers to others. The net result is a system of governance in which those who exercise power are further and further removed from the legitimizing source of that power: the citizens who are affected by it. The gains in problem-solving efficiency (assuming they are real) come at the expense of legitimacy. This is was what we meant when we earlier referred "attenuating the legitimacy relationship."

Of course, this is not a new problem. Governments have been dealing with it since the birth of the administrative state. They try to safeguard against it through various policies and legal doctrines. The names of these policies and doctrines vary from country to country, but broadly speaking we can identify three conditions that need to be satisfied if the creation of and exercise of power by specialized agencies is to be legally and politically justifiable.

1. **Sound policy rationale:** There must be some compelling public interest that is served by the agency and its use of power, and its policies and practices must be directed at that public interest.
2. **Appropriate delegation of power:** The agency must have its powers conferred upon it in a legitimate way by some legally legitimate authority. Usually this will be through a piece of legislation that specifies what the agency is supposed to do and how it is supposed to do it. The legislation may give discretion to the agency, including the power to subdelegate to third parties (including private companies), but usually this is tolerable only if there are clear limits to the discretion.

3. **Compliance with principles of natural justice/fair procedures:** The agency must comply with widely recognized principles of natural justice, due process, or fair procedures (different terms are used in different jurisdictions) when exercising its powers against citizens. This typically means that the agency should have some plausible reason or explanation for why it exercised its powers in a particular way, that it not exercise its power in a discriminatory and unfair way, and that the person affected has the right to be heard and the right to appeal (or review) a decision to an impartial tribunal. (See box 2.1 in chapter 2 on the difference between "appeal" and "review.")

The practical meaning of each of these conditions can be quite technical and complex in particular legal traditions. We'll encounter some of this complexity below, but for now, all that matters is that the conditions are understood at a general level.

This gives us everything we need to understand the problem that might arise from the use of algorithmic decision tools by government agencies. The *general* concern is that the proliferation of specialized agencies attenuates the legitimacy of power; the *specific* concern is that the use of algorithmic tools by those agencies might further exacerbate this attenuation of legitimacy. Although there are safeguards in place to deal with the attenuation of legitimacy at a general level, we might wonder whether algorithmic tools pose novel or unexpected problems when it comes to the attenuation of legitimacy that cannot adequately be met with those traditional safeguards. In other words—and to continue the theme from earlier chapters—we might ask, is there something different when it comes to the use of this technology?

Case Study 1: The Go Safe Automatic Speed Cameras

Before we get too mired in the legal and philosophical debate, let's consider a case study. This case study illustrates some of the ways in which the use of power by specialized agencies can go wrong, particularly when a technological aid is used to assist in the exercise of that power. The case study comes from Ireland, and it concerns the use of automatic speed cameras to prosecute speeding offenses.[5]

Ireland, like many countries, treats speeding as a minor offense. If you are ever caught speeding on an Irish road, you will more than likely receive a fixed penalty (or charge) notice in the mail. The penalty will consist of a

fine, the application of "penalty points" to your driver's license, or both. If you clock up too many penalty points, you might be banned from driving for a period of time. In most cases, when people receive these fixed penalty notices, they pay the fine immediately and think no more about it. If they fail to pay the fine by a due date, they might be summonsed to court and face a slightly more severe criminal prosecution. Under the relevant legislation in Ireland (the Road Traffic Act 2010, as amended) the Irish police force is given the authority to detect and prosecute speeding offenses. There is, however, a section of that act, section 81(7), that allows for this power to be delegated to a third party (e.g., a private company) as long as this is done via a written agreement entered into with the Irish Minister for Justice.

In 2009, the Irish police force, with the approval of the Irish government, decided to outsource the detection and prosecution of speeding offenses to a private company called Go Safe. They did so under a contract with the Minister for Justice that gave Go Safe the right to do this until 2015, with the option for renewal then. The company operated a fleet of vans with automatic speed cameras. These vans were placed in strategic locations for a number of hours. These locations had to be approved by high-ranking members of the Irish police force. While *in situ* the cameras in the vans would automatically detect violations of the local speed limit and take photographs of the offending vehicles. Subsequently, fixed penalty notices would be issued to the registered owners of those vehicles. Much of this process would take place automatically, with some minimal oversight by human operators. If the registered owner refused to pay the fixed penalty, they could be summonsed to court, and Go Safe workers would give evidence in court to support their prosecution.

On the face of it, this would appear to be a textbook example of how to effectively use technology to administer the business of the state. There is a clear and obvious public interest at stake—speeding is a contributing factor to road deaths and reducing it protects public safety.[6] Automatic speed cameras can accurately detect speeding violations without the need for constant human supervision and input. The mere presence of those cameras, or the suspicion of their presence, has a deterrent effect and prompts drivers to change their behavior. Using them saves precious policing resources and enables the efficient administration of justice. What's more, the way in which the power to detect and prosecute was given to the Go Safe company does not appear to raise any obvious red flags when it comes to the

attenuation of legitimacy. There was a piece of legislation that authorized the outsourcing of this power to a third party; there were statutorily prescribed limits to how much power they could have; and there had to be a written agreement spelling out the terms and conditions of their service.

But things didn't work out so smoothly in practice. Road traffic offenses are notoriously complex. There are a number of technical protocols that need to be followed to legally prove their occurrence. It is not uncommon for defense lawyers to use this technical complexity to the advantage of their clients, pointing out how certain protocols were not followed thus rendering a prosecution void. Once defense lawyers got to work on figuring out the flaws in the Go Safe system, things quickly turned sour. According to one report,[7] over 1,400 attempted prosecutions of speeding offenses using the Go Safe system were thrown out of court. A variety of grounds were given for these dismissals. One of the biggest problems was that Go Safe could not prove whether fixed penalty notices had been received by potential offenders or that registered owners were in fact driving their vehicles at the relevant time. A whistleblower from the company complained that the company demanded that its workers record potential speeding violations even when they couldn't set up the camera equipment appropriately and so would be running the risk of false positive identifications.[8] Judges complained that when officials from the company gave evidence in court they could not explain how the camera system worked, how far above the speed limit someone had to drive before being issued a fixed penalty notice, or how the fixed penalty notices were generated and sent to potential offenders. They also lamented the failure to provide an adequate chain of evidence and to prove that they had the authority to give evidence in court. In short, the practical implementation of the Go Safe system was a shambles, so much so that one Irish judge referred to it as an "abject failure."[9]

Some of the problems with the Go Safe system have since been addressed,[10] and the company's contract was renewed in 2015.[11] Nevertheless, the Irish experience with the Go Safe system provides us with a cautionary tale. Automatic speed cameras are not an advanced technology. They use simple radar reflection to calculate the speed of oncoming vehicles, determine whether this exceeds the local speed limit, and then take a photograph. They are nowhere near as sophisticated as some of the algorithmic tools discussed in this book. They don't make complex predictions or judgments. They don't depend on esoteric programming techniques or advanced artificial

intelligence. Even still, their use in the administration of justice caused untold practical headaches. A new technological system had to integrate itself into an old legal governance system—like a software upgrade on an old computer. The upgrade did not go well. Its problems were compounded by the fact that it involved a contractual relationship between a government agency and a private corporation. Simple errors and omissions were made when implementing the system. People representing the company could not adequately explain how their systems worked. Sometimes incentives were misaligned and corners were cut. The result was a failure to meet a key public policy objective and a failure to legitimately delegate power.

If all these things can happen with a relatively straightforward technology, we would be well advised to be on our guard with a more complex one.

Does the Use of Algorithmic Decision Tools Pose a Threat to Legitimacy?

Granted that we need to be careful about the use of algorithmic decision tools by government agencies, is there any reason to think that special caution is required? Or, to go even further, could it be that the use of algorithmic decision tools poses such serious risks to the legitimate use of power that there should be a general presumption against their use?

Several commentators have expressed skepticism about the need for such hyper-caution.[12] They point out that the administrative state has long battled with criticisms concerning the legitimate use of power, and it has survived these criticisms despite increased specialization and increased use of public-private partnerships in how it carries out its key functions. Even the fiasco of the Go Safe system in Ireland did not lead to any major doubts about the use of technology by government agencies or the legitimacy of public-private partnerships. It just led to changes in policy and practice. It seems likely that a similar pattern will be followed when it comes to the use of algorithmic tools by government agencies. Nevertheless, it is worth entertaining a hyper-cautious stance, if only to see what might be wrong with it.

To do this it helps if we clarify exactly how algorithmic tools can be and might be used by government agencies. Government agencies perform two main functions: (1) they create policies and rules (if they have been given the discretionary authority to do so) and (2) they implement and enforce policies and rules (either the ones they themselves have created or those that have been stipulated for them by other elected authorities).

Algorithmic tools can help with both tasks. Examples have been given in previous chapters. Algorithmic tools are already being used by government agencies to efficiently manage the implementation and enforcement of rules, for example, in predictive policing software that provides guidance on how to deploy policing resources, and to identify and repair flaws in existing policies or rule-making frameworks, for example, in smart energy grid systems or traffic signaling systems. Sometimes algorithmic tools will be used as mere aids or supplements to human decision-making within the relevant government agencies. At other times algorithmic tools will operate autonomously with minimal human supervision and interference. In their discussion of the issue, Cary Coglianese and David Lehr refer to the possible autonomous uses of algorithmic decision tools as "adjudication by algorithm" and "regulating by robot," respectively.[13] These names are chosen for their alliterative appeal, not for their descriptive accuracy. For example, as the authors point out, what they call "regulation by robot" may not involve a robot; it may just involve an algorithmic system making rules without direct input from a human controller.

The merely assistive and supplementary use of algorithmic decision tools seems to raise relatively few red flags when it comes to the legitimate use of power. If the algorithmic tools are being used *as* tools, then human decision makers retain the actual power, and any issues we might have with how they exercise that power are ones that we are ultimately familiar with. The existing legal and regulatory framework is designed to work with human decision makers. By contrast, genuinely "autonomous" uses of algorithmic decision tools have the potential to raise more red flags.* Depending on how much autonomy is given to a tool, there is a risk that humans will no longer be the ones in charge (i.e., in meaningful/effective control). Questions might then arise as to whether the existing legal and regulatory framework is fit for purpose. It is supposed to be government by *humans* for humans, not government by *machines* for humans.

There is, however, a danger that we get carried away when it comes to assessing the potential threat that autonomous algorithms pose to the legitimacy of government. In this respect, we are easily seduced by fictional

*In terms of chapter 5, assistive and supplementary tools don't pose the control problem. On the other hand, fully automated decisions (and subdecisions) do pose this problem.

motifs of machines taking over the reins of power. We are a long way from that with current technologies. A more sober analysis is needed, taking into consideration the three conditions for the legitimate use of power by government agencies that were outlined above. When assessed in light of these three conditions, does the use of autonomous algorithmic decision tools really pose a significant threat?

Well, we can quickly set aside the first condition. Whether there is a sound public policy rationale for the use of a decision tool will vary from case to case, but we can easily imagine that there is often going to be one. For example, we could create algorithmic adjudicators that scrape through data on financial transactions and automatically impose sanctions on entities that breach financial rules. These algorithmic adjudicators could very credibly do a better job than human adjudicators. Financial markets are already suffused with trading bots and algorithms, executing thousands of trades in the blink of an eye. No human adjudicator can keep on top of this. An algorithmic adjudicator might be just what we need to ensure the smooth functioning of the regulatory system.[14]

That leaves us with the other two conditions: (1) whether power has been legitimately delegated to the algorithmic tool and (2) whether the use of the algorithmic tool complies with principles of natural justice/fair procedures. The issue of legitimate delegation is tricky. As Coglianese and Lehr point out, there are a number of different ways in which power could be delegated to an autonomous algorithm.[15] The first and most straightforward would be if a piece of legislation explicitly provides for the use of the tool. Imagine, for example, a road traffic act that stated explicitly that the police force is entitled to use "algorithmic adjudicators" to determine whether someone has breached a speed limit and to automatically impose penalties if they do. As long as the legislative provision is clear about this—and it is supported by some intelligible public policy rationale—the delegation of power to the algorithm would be legally uncontroversial. That's not to say that it would be wise from a political or public policy perspective (a point to which we return); it is just to say that it doesn't create special problems from a legal perspective. This is the normal and most legally appropriate way to delegate power. In Australia, section 495A of the Migration Act authorizes the minister of immigration to use computer programs when making certain decisions, and decisions by the computer program are then taken to be decisions by the minister.

The second way in which power could be delegated to the algorithmic tool is through discretionary subdelegation by someone within a government agency. This is not an uncommon practice. Elected governments create specialized agencies to manage and regulate key sectors of society because they realize that they themselves lack the requisite knowledge and expertise. Consequently, they have to give the staff within those agencies some discretion as to how best to create and implement rules and policies. So, for example, there might be a road traffic act that gives the police force the power to use "whatever means they see fit" to enforce the speeding laws. An official working within the police force might read about a new "algorithmic adjudication" system that would allow them to enforce the speeding laws more efficiently. Following appropriate consultation and procurement, they might decide to use this tool to exercise the discretionary power that has been delegated to them by the road traffic act.

But this discretionary subdelegation to an algorithm might be legally problematic. Questions would necessarily arise as to how to construe the precise scope of the discretionary power of subdelegation. A provision that allows an agency to use "whatever means they see fit" might be deemed too vague and open-ended to legitimize that form of delegation. Surely the police couldn't use "any means"? Imagine if they started using smart AI-based landmines that detonated underneath cars that exceeded the speed limit by 0.1 of a km/h? That would surely be shut down, largely because it would not comply with other principles underlying the legitimate use of power (e.g., principles of fair and proportionate punishment). This is a silly example, but it illustrates an important point. Even where the discretionary power given to an agency under a statute seems quite broad, there must be some limits to it. Discretionary powers that are phrased in terms of using "appropriate means" or "proportionate means" will raise similar issues of scope uncertainty.

Writing from the US perspective, Coglianese and Lehr argue that, although the subdelegation of power to an algorithmic tool will raise questions as to how best to interpret the power of subdelegation under the relevant statute, there is "in principle" no reason why power could not be subdelegated to an algorithm.[16] They make two points in support of this argument. First, they argue that government agencies already unproblematically make use of measurement devices that perform the kinds of functions that could be subdelegated to an algorithm. No one questions their right to do so on grounds of illegitimate delegation of power. Second, they argue that given

the way in which algorithmic tools currently operate (they focus specifically on machine learning algorithms) humans will always retain some control over them, either by determining the objectives/goals they are supposed to meet or determining how and when they are to be used. There will never be a complete subdelegation of power to an algorithm.

But we might question Coglianese's and Lehr's sanguine outlook. In chapter 5, we considered the issue of control in relation to the use of algorithmic decision tools and spoke about the risk of "automation bias." As we argued, there is a very real risk that when people rely on algorithmic systems they might defer to them excessively and lose meaningful control over their outputs. This could happen in government agencies too. Government agencies that rely on autonomous algorithmic tools may adopt an uncritically deferential attitude toward them. They may, in principle, retain some ultimate control, but in practice it is the algorithm that exercises the power. The British sketch comedy show *Little Britain* satirized this problem in one of its recurring sketches involving a receptionist who could never answer a customer's query because "the computer says 'no.'" A similar attitude might creep into government agencies. The human officials may not be inclined to wrestle control back from algorithmic tools, not because the tools are smarter or more powerful than they are, but because habit and convenience make them unwilling to do so. Looking at this topic from the perspective of UK administrative law, Marion Oswald argues that this kind of deference to the machine would be legally problematic no matter whether the algorithm was developed by the agency themselves or by some third party.

> A public body whose staff come to rely unthinkingly upon an algorithmic result in the exercise of discretionary power could be illegally "fettering its discretion" to an internal "home-grown" algorithm, or be regarded as delegating decision-making illegally to an externally developed or externally run algorithm, or having predetermined its decision by surrendering its judgment.[17]

Before this happens, we might like to take some corrective action and ensure that we don't allow government agencies to fetter their discretionary power to such an extent. We might like to create a new norm whereby any delegation of power to an autonomous algorithm has to be done on the basis of an explicit legislative provision, not through discretionary subdelegation.

Finally, what about natural justice and fair procedures? Is there any risk that the use of algorithmic tools poses special problems in this regard? The concept of natural justice or fair procedures overlaps significantly with

topics discussed in previous chapters. A fair procedure, broadly construed, is one that is relatively impartial,* takes into consideration the interests of the parties affected, gives them a right to be heard or consulted, provides reasons for the decisions made, and gives the affected person a right of appeal in the event that they still feel aggrieved. It is certainly possible that the use of an algorithmic tool would violate these requirements. The problems of bias and transparency, discussed in earlier chapters, would be relevant in reaching such a conclusion. If an algorithmic adjudicator makes decisions in a systematically biased and opaque way, then it may very well fall foul of the fair procedure requirement. But, as was pointed out in those earlier chapters, there are some difficult tradeoffs when it comes to eliminating bias, and there are ways to ensure that algorithmic decisions are explainable. So even though an algorithmic adjudicator could fall foul of the fair procedure requirement, it doesn't have to. There is no categorical reason to oppose the use of algorithmic tools by government agencies on these grounds.

There is another point worth making here too. Although there is an idealized conception of what a fair procedure should look like, many legal systems do not insist that every decision made by a government agency meet this ideal. Some corners can be cut in the interests of efficiency and cost-effectiveness. This makes sense. If an agency had to hold an impartial tribunal and provide detailed explanations for every decision they made, the actual day-to-day business of that agency would grind to a halt. Courts usually accept this and adopt "balancing tests" when figuring out how close to the ideal of a fair procedure any particular decision-making process must get. For example, US courts focus on three factors: how the decision impacts the affected party, the potential cost to that individual if a wrongful decision was made, and the net gain (if any) from introducing additional procedural safeguards to protect against that potential cost. They then weigh these three factors and decide whether a particular decision-making process is acceptable or needs to be reformed to pass muster.[18] Given this sensitivity to cost-effectiveness, it's quite likely that many uses of algorithmic tools will be deemed legally legitimate even if their use prevents a decision-making procedure from living up to the ideals of natural justice. After all, one of the reasons why government agencies might be tempted to use this

*We say "relatively" because absolute impartiality is impossible.

technology is to help them manage complex systems in a cost-effective manner. Think back to the earlier example of an algorithmic adjudicator enforcing financial regulation against algorithmic trading bots.

None of this is to say that the use of algorithmic decision tools by government agencies is always going to be a good idea. A technological aid might be a good idea in principle but a bad idea in practice. Furthermore, even a tool that does enable the more effective management of social systems and is technically legally legitimate might not be *perceived* as legitimate. Public administration is as much about good public relations as it is about legal technicalities and economic efficiencies. Any government considering the use of such a system would be well advised to consult widely on its introduction, listen to the concerns of key stakeholders, and constantly review the practical workings of the system once it is up and running. These are standard practices in government agencies the world over anyway, but it is important that these practices are maintained and emphasized.

It is also not to say that the use of algorithmic decision tools throws up no new challenges for the legitimacy of public administration. They do. Algorithmic tools work on the basis of a precise quantified logic. Human decision makers often work more by qualitative reasoning and intuition. This means that if some authority is delegated to such tools, there is going to be a need to translate what was once a qualitative decision-making process into an explicitly quantitative one. This translation process might throw up some new issues. It is impossible to create a flawless decision-making process. There is always some risk of error. An automated speed camera system, for example, might sometimes (even if only rarely) fail to record a car that is traveling over the speed limit or mistakenly record a car that was not. The former would be a false negative error; the latter would be a false positive. We live with the risk of both errors all the time, but we often don't think about them in explicitly quantified terms. In other words, we don't explicitly decide that we are okay with a system that makes false positive (or false negative) errors 5 percent or 10 percent of the time. We often live with the illusion that we are aiming for perfection. The use of algorithmic decision tools will force us to discard this illusion. Although some government agencies are already comfortable with making explicit choices about error rate, others may not be. The fact that they might *have* to may also create a crisis of perceived legitimacy, as the public has to confront the quantitative realities of risk. If a government department is going to use

an autonomous decision tool, it should be able to support this choice by making publicly available evaluations showing that the tool's decisions are equivalent to, or better than, the decisions of the relevant human staff.* But this requires delicate handling.

There is another, more philosophical, problem with the quantitative logic of algorithmic decision tools. It has to do with how these tools treat individual human beings (see the previous chapter). As noted earlier, in liberal democratic states, the legitimacy of public power is founded on respect for individual citizens. They are autonomous, dignified beings whose lives must be taken into consideration. To paraphrase the philosopher Immanuel Kant, they must be treated as whole, integrated persons—as ends in themselves—rather than as means. A long-standing concern with the specialization and bureaucratization of public administration is that it fails to live up to this Kantian ideal. The dignity of the individual is undermined by complex, mechanized management. Individuals are cogs in the machinery of the state. They are confronted by labyrinthine processes and nameless officials. They are bundles of statistics, not fully-rounded characters. They are "cases" to be managed, not persons to be respected. Although this dehumanization concern is a long-standing one, it is a concern that could be exacerbated by the use of algorithmic decision tools. These tools necessarily reduce persons to bundles of data. They have to quantify and disaggregate people's lives into mathematically analyzable datasets. They don't "see" people; they see numbers. Give this necessity, special safeguards may need to be put in place to maintain the human touch and ensure that the dignity of the citizen is respected.

Case Study 2: The Allegheny County Family Screening Tool

Let's now consider another case study in the use of technology by government agencies. Unlike the Irish case study, this one involves a more complex algorithmic decision tool. The case study concerns the Allegheny County Family Screening Tool (AFST for short).[19] We mentioned it briefly in chapter 5 already. We return to it here to see whether it holds any lessons for the legitimate use of power by government agencies.

---

*We discussed evaluation methods in chapter 1.

The AFST is, as the name suggests, a screening tool used to identify children who may be at risk of abuse and neglect. It was created by a group of academics, led by Rhema Vaithianathan and Emily Putnam-Hornstein.[20] The group was originally commissioned by the New Zealand Ministry of Social Development to create a predictive risk modeling tool that could sort through information about how families interact with public services and criminal justice systems to predict which children were at the most risk of abuse or neglect. The tool is supposed to use this information to generate a risk score for each child that can then be used by child protection workers to investigate and prevent cases of abuse and neglect. In this respect, the tool is not dissimilar to predictive policing tools that generate heat maps to assist police departments with the efficient distribution of policing resources. Indeed, the visual display used by the family screening tool is quite similar to that used in predictive policing heat maps. It adopts a traffic light warning system that identifies high-risk cases with a red light and low-risk cases with a green light.

The team led by Vaithianathan and Putnam-Hornstein eventually came up with the predictive model that relied on 132 different variables—including information about the mother's age, whether the child was born to a single parent family, mental health history, criminal record, and so on—to generate its risk scores. They claimed that the system was reasonably accurate at predicting abuse and neglect. But they ran into problems when trying to implement it in New Zealand.

The role of the government in child protection has always been politically contentious. At the extreme, child protection law entitles government agents to take children away from their lawful parents. This is often an upsetting experience for both the parents and the children. It is usually a last step. In most countries, child protection caseworkers only escalate to this step when other interventions have failed, but historical experience has made some people very suspicious of the process. There is often the belief that government agents unfairly target poor parents from minority backgrounds or those with unconventional lifestyles, and there is a sordid history of children from indigenous minority ethnic families (e.g., New Zealand, Australia, and North America) being taken from their families at disproportionate rates.[21] It is unsurprising, then, that the New Zealand government halted the experimental use of Vaithianathan and Putnam-Hornstein's tool in 2015.[22] (Strictly speaking, the experiment was aborted

because the minister didn't want to use children as guinea pigs rather than because she acknowledged that the tool would be discriminatory, but the issues are often difficult to disentangle.[23])

The Allegheny County Department of Human Services (DHS)—based in Pennsylvania, however, did award a contract to the group to create the AFST. They have used the system since August 2016.[24] It is clear that, in designing and implementing the AFST, both the academic team and the Allegheny County DHS were very sensitive to concerns that local citizens might have about the tool. They held a number of meetings with key stakeholders in the county to determine how best to create and implement the system. They also hired an outside, independent team to conduct an ethical analysis of the AFST. This team concluded that use of the tool was ethically appropriate because it was more accurate than preexisting systems for screening cases.[25] Furthermore, the group behind the AFST have tried to ensure that the system operates in a highly transparent way, releasing information about the variables used and providing detailed and updated FAQs to the public about how it works.[26] This has led to some public praise. For example, writing in the *New York Times*, the journalist Dan Hurley was largely positive about the impact of the AFST and considered it an important development in the fight against child abuse and neglect.

But others have been more critical. The political scientist Virginia Eubanks, in her book, *Automating Inequality*, argues that the AFST still unfairly targets children from poor and minority families, and unfairly correlates poverty and use of public services with an increased risk of abuse.[27] She has also argued that the creators of the system have not been transparent enough in releasing information about how the AFST works. They have released details about the variables they use but not the weights they attach to these variables. She also argues that the system may have an unacceptably high error rate, noting that officials at the Allegheny County DHS state that 30 percent of cases flagged by the AFST are ultimately thrown out as baseless,[28] but further noting that this cannot be properly assessed until information is released about the error rate.

The debate about the AFST perfectly illustrates the themes of this chapter. The AFST was created to enable a government agency to more accurately and more efficiently go about its business. But the way in which it operates creates concerns about the legitimate use of power. This is true even though many of the steps to ensuring legitimacy were followed. Critics still worry

about the opacity and complexity of the system: they worry that it may be unfair or biased; they want to be more involved with its implementation; and they demand greater transparency and openness when they are. This is a cycle that never really ends. The legitimacy of power can always be called into question. It is incumbent on both citizens and those working in government agencies to be willing to scrutinize how power is wielded on an ongoing basis to ensure its legitimacy. The use of algorithmic decision tools adds a new technological flavor to this perpetual dynamic.

There is another lesson to be learned from the AFST case study. One of the major criticisms that Virginia Eubanks launches against the AFST and other systems like it is that they are examples of technological solutionism in practice.[29] We now have sophisticated, algorithmic risk-prediction tools. We know they are more accurate than human decision makers in certain cases, so we look around for problems that they could be used to solve. But we don't think it all the way through. We are so eager to find some application that we don't think about the side effects or indirect consequences of that application. Nor do we consider that there may be *other* problems—perhaps not so easily solved by risk-prediction software—that are more deserving of our attention. For instance, Eubanks worries that the AFST intervenes too late in the game, that is, when a child has called looking for help. At this stage in the game, the systematic biases and structural inequalities in society have taken their toll on the affected families. This is something that the independent ethical reviewers of the AFST themselves noted, observing that predictive risk modeling can be founded upon, and consequently help to reinforce, existing racial biases.[30] This is a point we made in chapter 3.

It is unrealistic to expect a tool like the AFST to completely correct for such structural inequalities, given its intended use. It has to work off a social reality in which members of poorer groups and racial minorities do suffer from disadvantages that may very well place them at a higher risk of some negative outcome. But in that case, the use of the tool may widen the perceived legitimacy gap, at least when it comes to members of those populations. Perhaps we should use this realization as an opportunity to reflect on how we want government agencies to use algorithmic decision tools. Perhaps there are other problems they could help us solve that might then reduce the legitimacy gap for those groups. This will obviously require more creative, outside-the-box thinking about how best to use algorithmic decision tools.

A Necessity for Algorithmic Decision Tools?

We started this chapter by noting that the proliferation of specialized government agencies has advantages when it comes managing social problems, but that this advantage comes with the risk of attenuated legitimacy. The use of algorithmic decision tools by those agencies must be interpreted and understood in light of that long-standing dynamic. Fortunately, we have developed political and legal safeguards to protect against the attenuation of legitimacy, and these can be used to evaluate and constrain the public use of algorithmic decision tools. This means that although there is no insuperable legal or political obstacle to the use of algorithmic decision tools, they should be deployed carefully and thoughtfully, and their use should be subject to public scrutiny and oversight.

In making this argument we have suggested, at various points, that the use of algorithmic decision tools could be a boon to government agencies. As the world becomes more complex and as algorithmic tools are used by private individuals and corporations, the use of similar tools by government agencies may graduate from being practically desirable to being practically necessary. It may be the only way to keep up. On the face of it, though, it seems like this is an argument that can only be made on case-by-case basis. Is there anything more general that can be said in favor of the use of algorithmic decision tools?

Perhaps. Consider the lessons to be learned from the history of societal collapse. In his influential book, *The Collapse of Complex Societies,* the archaeologist Joseph Tainter tries to provide a general explanation for the collapse of complex societies.[31] He looks at all the famous examples: the Egyptian Old Kingdom, the Hittite Empire, the Western Roman Empire, Lowland Classic Maya, and more. Although some archaeologists and historians dispute whether these societies actually collapsed—some argue that they simply adapted and changed[32]—it seems relatively uncontroversial to say that they underwent some decline in their social complexity (i.e., their administrative centers were dissolved; they abandoned settlements and discarded cultural and institutional artifacts). Tainter wonders whether there is some common cause of these declines.

After reviewing and dismissing several of the most popular explanations, Tainter comes up with his own explanation for societal collapse. His explanation consists of four key propositions. The first proposition is that

complex societies are problem-solving machines. In other words, they sustain themselves by addressing the physical and psychological needs of their members. If they don't do this, they lose their legitimacy and ultimately collapse (or undergo significant decline). The second proposition is that in order to solve problems, societies have to capture and expend energy. Classically, societies did this through foraging and farming; nowadays, we do it through burning fossil fuels and exploiting other energy sources. The general rule of thumb is that, in order to solve more problems and sustain greater complexity, societies have to increase their energy expenditure (or at a minimum make more efficient use of the energy available to them). This leads to the third proposition—that the survival of complex societies hinges on a basic cost-benefit equation. If the benefits of societal problem-solving outweigh the costs, then the society will continue to survive; if they do not, then the society is in trouble. The fourth proposition states that an increased investment in social complexity (e.g., an increased investment in specialized administrative agencies) results in more benefits, but only up to a point. The increased complexity comes with an increased energy cost so that, eventually, the marginal cost of increased complexity will outweigh the marginal benefit. If this happens, and if the inequality is of a sufficient magnitude, the society will collapse. This problem of declining marginal returns is the core of Tainter's explanation. He argues that the complex societies that have collapsed throughout history have all confronted the basic problem of declining marginal benefits and increasing marginal costs.

Tainter gives several illustrations of how societies have had to deal with this problem. Some of the best documented examples have to do with bureaucratic power and control. Tainter shows how many government agencies undergo administrative bloat and mission creep. These agencies, often central to societal problem-solving, proliferate and grow in size to cope with new challenges. This is initially beneficial, as it allows the society to address more complex problems, but it ultimately results in declining returns as costly administrative staff have to be hired to manage the organizations themselves. There are ever more of these administrators compared to relatively fewer frontline caseworkers who actually deal with the problems those organizations are designed to address.

Tainter's theory prompts reflection on the role of technology in preventing societal collapse. Perhaps one way to solve problems efficiently without dramatically increasing the cost of administrative bloat is to double down

on the use of algorithmic decision tools and other forms of AI. According to some of its advocates, a significant breakthrough in AI could be the *deus ex machina* we need to solve our growing social problems. Instead of relying on imperfect, squabbling human intelligences to manage our social problems, we could rely on a more perfected artificial form. This is an argument that Miles Brundage makes explicit in his paper on the case for "conditional optimism" about AI in global governance.[33] This doubling down on algorithmic decision tools and AI might come at the cost of some perceived legitimacy, but maybe that is the price we as citizens have to pay to maintain our complex social order. It is a provocative and disturbing suggestion, and may mean that we should be more open to the widespread use of algorithmic decision tools. Their legitimacy might be greater than we'd like to think.

Summing Up

The use of specialized agencies and specialized tools is essential to the modern business of government. You cannot govern a complex, industrial society without relying on expert agencies with specialized problem-solving skills. These agencies make judgments and set policies that have a significant impact on the lives of citizens and must exercise that power legitimately. The problem is that the increasing diversity and organizational depth of such agencies attenuates their legitimacy. The use of autonomous algorithmic decision tools can exacerbate this problem. Nevertheless, we shouldn't *overstate* the problem. Many countries have developed legal and political doctrines that can determine when and whether the exercise of power by a government agency is legitimate. There is no strong reason to think that these doctrines are no longer fit for purpose. We may just need to modify them and err on the side of caution when it comes to ensuring that power is legitimately delegated to algorithmic decision tools and that their use does not undermine the need for fair procedures. One way or another, AI will be increasingly indispensable if our ever more complex societies are to be prevented from collapsing under their own weight.