Original Research Article

Recalibration in counting and accounting practices: Dealing with algorithmic output in public and private

Farzana Dudhwala [iD] and Lotta Björklund Larsen

Abstract

Algorithms are increasingly affecting us in our daily lives. They seem to be everywhere, yet they are seldom seen by the humans dealing with the consequences that result from them. Yet, in recent theorisations, there is a risk that the *algorithm* is being given too much prominence. This article addresses the interaction between algorithmic outputs and the humans engaging with them by drawing on studies of two distinct empirical fields – self-quantification and audit controls of taxpayers. We explore *recalibration* as a way to understand the practices and processes involved when, on the one hand, decisions are made based on results from algorithmic calculations in counting and accounting software, and on the other hand, when decisions are made based on human experience/knowledge. In particular, we are concerned with moments when an algorithmic output differs from expectations of 'normalcy' and 'normativity' in any given situation. This could be a 'normal' relation between sales and VAT deductions for a business, or a 'normal' number of steps one takes in a day, or 'normative' as it is according to the book, following guidelines and recommendations from other sources. In these moments, we argue that a process of recalibration occurs – an effortful moment where, rather than treat the algorithmic output as given, individuals' tacit knowledge, experiences and intuition are brought into play to address the deviation from the normal and normative.

Keywords

Algorithms, recalibration, normativity, accounting, algorithmic output, quantified self

This article is a part of special theme on Algorithmic Normativities. To see a full list of all articles in this special theme, please click here: https://journals.sagepub.com/page/bds/collections/algorithmic_normativities.

Introduction

Algorithms seem to be everywhere. Recent scholarship suggests an increasing proliferation of digitisation in decision-making across, and within, organisations, corporations and bureaucracies (Domingos, 2015; Lyon, 2014; McQuillan, 2015). For the end-user, this digitisation often appears as a maze where only the outcome is seen, and the enormous impact this is having in people's everyday life is already under examination (e.g. Isin and Ruppert, 2015; Striphas, 2015; Zarsky, 2016). In this article, we discuss how humans make decisions when engaging with systems and devices that produce algorithmic outputs. We aim to shed light on situations where algorithms propose a way, a solution, a move, to return a situation to a state of normalcy. What happens when an algorithmic output does not resonate with human expectations of what should be normal? Under what circumstances does the human feel able to reject the algorithmic output, and when does the algorithm win out?

Tax Administration Research Centre, University of Exeter Business School, Exeter, UK

Corresponding author:
Farzana Dudhwala.
Email: fdudhwala@gmail.com

Our two empirical fields are very distinct – self-quantification and the sampling of taxpayers for audit – but both derive from the ontological concern of the impact of algorithms on everyday decisions. We are concerned with moments when an algorithmic output differs from the tacit knowledge of 'normalcy' in any given situation – be this the 'normal' relation between sales and VAT deductions for a business, or the 'normal' amount of steps one ought to take in a day. By introducing three varieties of the concept *recalibration*, we draw attention to the many instances where algorithmic output is questioned and the implications such questioning have for further action.

In the first part of the article, recalibration is proposed as a way to account for the multiple other considerations that need to be made when examining phenomenon involving algorithms – including human and contextual decision-making processes. These are instances depicting the human work that takes results from algorithmic calculations in various counting and accounting software and compares these results with circumstantial expectations on what is normal. Such expectations are referred to as tacit knowledge; the particular set of embodied knowledge practices that can be revealed by ethnographic and critical engagements (see Elyachar, 2012; cf. Kockelman, 2007; Strathern, 2005). Although we do not delve into the tacit knowledge per se, it provides the necessary backdrop for understanding the circumstances in which recalibration of the algorithmic output occurs.

In the next section, we locate the human–algorithm relation that we find in moments of decision making in our cases within the current literature. We concur with many recent theorisations; that the *algorithm* is at risk of being given too much prominence in practices that form the world (e.g. Pink et al., 2017). Although algorithms might be everywhere, there are many environments, contexts, and instances where humans still tinker with algorithmically produced results not just following their recommendations and instructions. Instead they may reject, adapt, ignore or as we propose – recalibrate – the algorithmic output. We illustrate instances of recalibration with two empirical cases. Our first case is based on controls at the Swedish Tax Agency, where tax auditors check for unsolicited cost deductions in annual tax statements. The second case draws on the practices of self-quantification, where the self is counted using technologies and devices to gain an account of the self. We conclude our article teasing out instances when humans bow to the algorithmic solution.

At first glance, our cases may seem disparate in many ways, yet we find commonalities in how our informants relate to algorithmic outputs that propose an action and accordingly recalibrate their decisions. Both our cases build on long and extensive ethnographic fieldwork. Lotta Björklund Larsen (henceforth LBL) followed a risk assessment project at the Agency

for its entire duration of three years: from its initiation, through the research phase, intermingled with presentations of its conclusions within the Agency and to see the final project report finally buried (Björklund Larsen, 2017). The resulting material builds on over 100 hours of transcribed meetings and interviews at the Agency, all their internal e-mails, policy regulations and instructions for their work, as well as copies of work in progress. A substantial part of this project was a random audit control; the largest ever performed by the regional department specialising in such audits and is the focus of this article.

Farzana Dudhwala (henceforth FD) spent five years with members of the Quantified Self (QS), predominantly based in London. She attended 26 Meetups (events run by, and for, members of the QS to share learnings and experiences of self-tracking, watched 113 'show and tell' presentations (the predominant format to share said learnings and experiences), did 32 interviews, and participated in two QS 'unconferences' (unstructured conferences where sessions are formed from interest in the participants). There were extensive field notes taken throughout these ethnographic engagements, as well as transcriptions of all of the interviews.

The data from both our projects was input into NVivo software to then be coded iteratively to make sense of the themes coming out of it. The vignettes and quotes chosen for this paper are used to make broader points about the concepts that are derived from them, and specifically because they were deemed to be fairly typical of the kinds of behaviour seen throughout.

We have each participated in our respective groups' activities, be that performing audits or using self tracking devices; we have observed what our informants do; and we have scrutinised how they engage with the technologies given to them. In doing so, in each of these two cases, we have been able to examine moments where recalibration of varying degrees has taken place.

Recalibration to normalcy

According to the Cambridge dictionary, *recalibration* is defined as changing the way one does, or thinks, about something. This definition is distinct from *calibration*, which refers to checking the accuracy of a measuring instrument. The difference is small, yet important. Strictly speaking, a calibration is the act of using a set standard to compare a measure against the thing it is actually measuring; it does not include any subsequent adjustment of behaviour. Recalibration, on the other hand, describes how a decision or thought changes or is adapted given various types of input. It is, in other words, to carefully assess or adjust something based on information other than a normative set standard as the outcome of an algorithm often is. Such information may

be previous experience, instinct, or common sense, etc. – tacit knowledge. The aim of this paper is to specifically address human interaction between the output of algorithms and the subsequent decision(s) made as a result. This process is what we call *re*calibration and it occurs when an activity that involves counting or accounting through an algorithm does not fit with reasonable expectations of the outcome by the human decision maker.

Our use of recalibration also involves an understanding of what *normalcy* implies. In the case of quantifying the self, specific numbers are often given as a 'target' or 'goal' to which one is expected to aspire to – for example, taking an average of 10,000 steps a day was often cited as a worthy aim and this number therefore became normative for many from which variations were seen as deviations – whether good or bad. That derivation of normalcy was obtained in the absolute, where the figure in itself had achieved status of normal condition. But normalcy may also be gauged in relativity. In auditing practices, absolute numbers are often less important than a number relative to others: a corporation's registered VAT amount or stock value, for example, only makes sense when compared with the total business turnover. The algorithmic output is the result of a calibration, whereas recalibration occurs when human considerations of contextual, relative, and other factors are brought into the situation and thus the course of action changes, however slightly.

Making a decision about and changing the course of an action one takes is then not only about following the numbers given by the algorithm, but about taking into consideration knowledge or expectations about the normal condition – what the algorithmic output is reasonably expected to be, given other information. Based on these considerations, do people base decisions on the resulting numbers given by the algorithm, or by trusting their own knowledge of what the number is supposed to measure? Or, perhaps, is there something else going on entirely? We are therefore not looking at the recalibration of the algorithm itself but rather we focus on how the output from the algorithm is used in adjusting the expectations one has of normalcy and how subsequent decisions are made.

Based on our ethnographic engagements, we found three varieties of such recalibration. *Relative recalibration* denotes a comparison between various numbers suggested by the algorithmic systems. These can be algorithmic calculations, hard-coded recommendations set by the system or other instructions that the human follows. *Tacit recalibration* illustrates the process where the algorithmic output is set in relation to what is the human's own experience of such instances. In these situations, humans trust their own knowledge of the best way to create normal situations. *Anomalous recalibration* relates to abnormal, even dangerous, situations

where the algorithmic output is trusted. Here, the human seldom questions the algorithm's correctness; the human knowledge is deemed inferior and incomplete to what the algorithms can produce.

As Bateson (1972: xxiv) argued, '[A]n explorer can never know what he is exploring until it has been explored'. Through recalibration, we can explore what algorithms do as part of a larger whole, neither giving them a defined and specific role at the outset, nor setting them apart as a category of their own – as 'algorithmic configurations'. Algorithms are 'there' but the user seldom knows what they actually do. *Recalibration* looks at instances when the algorithmic outcome is given precedence despite tacit knowledge, as well as those instances when tacit human knowledge wins out in the face of the algorithmic output – how is this done? To answer these questions, we turn to examples from both counting and accounting to show how recalibration can figure into decisions made both at a larger scale involving decisions made about others, as well as the more private and intimate level about one's own self. Although these human actors recalibrate in different capacities, they all refer and compare their decisions to a perceived state of normalcy. Before turning to our fields of study, we situate our take on algorithms in the existing literature.

Locating the algorithm

Algorithms are seldom seen for humans dealing with their results as they hide deep down in software programmes, obscured by computer code languages and are thus inaccessible to most (Deville and van der Velden, 2016). As an algorithm is 'a process or set of rules to be followed in calculations or other problem-solving operations, especially by a computer' (*New Oxford American Dictionary*), one may be tempted to try and locate the specific 'process or set of rules' to interrogate it. Yet, as algorithms are constantly changing (see Kubler, 2017; Seaver, 2013), focussing too much on the algorithm itself can be akin to a dog chasing its own tail. The focus here is to disentangle the various aspects of decision-making in relation to algorithmic outputs, and this is where Paul Dourish's (2016) proposal is a useful starting point (cf. Wirth, 1977). First, he points to the need for separating how data is made available for algorithmic processing from how the result of the calculations is handled. Second, he highlights the puzzle of how to identify the algorithms in the maze of digital locations when they themselves are subject to constant change (cf. Gillespie, 2017). Third, he argues that we need to be attune to the temporalities of algorithms, where certain algorithms are selected, re-used, re-made and spread to other users and contexts. This fuels Dourish's fourth and final point, that the study of algorithms could be ethnographic – putting

the gaze on the communities constructing and working with algorithms. It is the latter work of algorithms that we turn our attention to in this article.

Rather than focus on the algorithm in isolation, we thus study how 'algorithms do work in the world' (Kitchin, 2017: 22; cf. Ziewitz, 2017). We put the gaze on communities using algorithmic output, *after* the algorithms have purportedly done their job. As we will show, there is not a universal acceptance of their outcome, and there is much more at play than *just* the algorithmic output. We need to take into account how decisions are made *with* algorithms where decisions draw on the tacit knowledge of the given situation. Whilst algorithms certainly do work in the world, they do so in interaction with users who are not necessarily knowledgeable about the actual content of the algorithms but only of the output they produce. Such outputs are seen on computer screens, digital displays and even on old fashioned paper printouts. Yet, as we will show the output of algorithms does not always determine further action. We investigate if something general can be said about what aspects are considered when a human decides to trust her tacit knowledge instead of the result from the algorithms.

From the outset, we consider algorithmic output to be part of a *larger network of actors* (cf. Latour, 2005). Michel Callon and Fabian Muniesa (2005), for example, suggest the term 'algorithmic configurations' in their study of how commodities are priced when computers were involved in price-setting. Price-setting is determined by algorithms, but never in isolation. 'Algorithmic configurations' describe how algorithms are part of a market practice that weaves together the goods, agents and exchanges and define prices on a variety of markets; from the Paris bourse, to the Marseille open-door fish market, right through to the decision of a customer doing their weekly shop at the supermarket. In particular, we note Callon and Muniesa's (2005: 1245) point that regardless of how calculations are made, there is 'not a single implacable logic, one that is becoming hegemonic: that of calculation as the only possibility for action'.

Malte Ziewitz's (2017) not-so-random walk through the streets and alleys of Oxford is also a useful inspiration to us where humans adapt to the outputs of an algorithm. In an experiment with an algorithm that gives him ad hoc instructions on where to turn at each junction, he engages with the instructions as 'practical reasoning' (Ziewitz, 2017: 2). Instead of analysing the logics of the algorithm itself, he deliberates with its suggestions while observing and engaging with the surroundings he meets on his stroll. He cannot always make the turns as suggested: the algorithm he brought with him was not specific about whether an alley should count as a public road; it did not take into account

where to run in a Y-junction; and he is not allowed to cross a private car-park. To our mind, his practical reasoning resonates with recalibrating engaging with what he experiences in the world.

In one sense, we could see the results of his algorithms as being 'broken', a concept that Sarah Pink et al. (2018) propose in a study of varieties of empirical fields. They emphasise that even all the dysfunctional, decaying, out-dated data need to be taken into account when studying digitalisation, to account for 'broken' algorithms. It is an important aspect to consider, although broken is perhaps a too strong a word for our case. We instead maintain that the algorithmic outputs that we study need continuous tinkering and maintenance to adapt to a changing world. In line with Ziewitz's (2017: 11) suggestion, therefore, we investigate what happens when we study 'algorithms not as objects to be known in theory but as figures to be used in practice'. As both Callon and Muniesa, as well as Ziewitz, show, algorithms do not lead blindly to future actions. These are the instances in which we argue that the algorithmic outputs need *recalibration*. We now turn to our empirical cases to show how this works in practice.

An audit control at the Swedish Tax Agency

Audit controls are one of the tools the Swedish Tax Agency use to make sure taxpayers report the right information and pay the right tax in on time. Such an audit is a systematic and independent verification of (financial) accounts and documents of an organisation or individual to ensure that they present a true and fair view of their activities and follow the law. An audit aims to 'draw conclusions from a limited inspection of documents, such as budgets and written representations, in addition to reliance on oral testimony and direct observation' (Power, 2000: 111). For a tax administration, the ultimate goal of an audit is to ensure that a taxpayer reports and pays the correct tax that is owed – the normal situation. A Swedish audit control is heavily regulated by law and there are two standard types of controls; the 'desk control' and the more intrusive 'audit control' which involves a visit to the taxpayer.

This Agency has an astonishing standing among Swedish governmental authorities and is, since the last 10 years, among the top-contenders of the citizens' most trusted governmental authorities. The Agency says it has obtained such a standing by being active, trustworthy and helpful (Skatteverket, 2013). They apply these values in their four public strategies: make sure all taxpayers do the right thing; earn the trust of citizens and corporations; ensure that all pay the right tax; and simplify the taxation process for all.

Audit controls address all the Agency strategies while paying careful attention to the values they are said to enact (Björklund Larsen, 2017). It is thus of particular interest to see how auditors handle intrusive audits based on calculations from tax returns while enacting the Agency values of being active, trustworthy and helpful.

This particular audit control was part of a large risk assessment project that the Agency undertook in 2010–2013. A random control, *slumpkontroll*, is a seeming contradiction in terms. To control, steer, and discipline the haphazard and irregular appears a daunting, if almost impossible task. It is making the irregular regular. A random audit control is a statistical project where human intervention is large; the challenge is to work in a consistent and uniform way, both in the sampling procedures and in the aim for equal treatment of taxpayers regardless of *who* conducts the audit. Tax collectors have tried out many different ways of identifying those most prone to make errors (Skatteverket, 2014); currently it means that the Agency has access to *really* Big Data and foresees revolutionary changes in its way of working with the advances, even disruptions, in technology. Yet, it is still up to individual tax auditor(s) to choose which *measure* should be applied in any given case and suggest what the tax ought to be for taxpayers that are seen to have reported questionable data. One of these measures could be to change the decision of taxation, another to require that additional information be given, a third measure could be no action at all (see Figure 1).[1]

The actual selection of taxpayers for audit is conducted according to strict statistical sampling methods. A third of Sweden's population of small businesses was found in the four selected regions, in total 100,000. When analysts do not know the extent of wrongdoing, 50% is automatically assumed. With a statistical confidence interval of 10%, 45–55% of these businesses were assumed to deduct incorrect costs, resulting in a minimum of 383 entities to audit. This number was adjusted upwards to 400, but as it was also assumed that there would be some taxpayers found without any business activity, as well as some unable to be reached, a final list of 450 was randomly selected.

Once a taxpayer is selected for audit, the question of *what* issues the human auditor considers when following or overriding the suggestions made by the algorithms arises. The audit control depicted below focuses on unreasonable costs for small businesses. Finding unreasonable costs implies an assumption of what reasonable costs are; a comparison against normal costs. An audit control is not just an inspection of documents supported by oral testimonies and direct observation as Power (2000: 111) suggested, but an algorithmic task unpacking the calculated output; the additions and subtractions, the depreciations, etc. that the taxpayer has calculated and entered into particular posts on the tax return statement.

Alice, an experienced tax auditor at the Agency, takes LBL through all the steps in the audit control procedure. It aims to ensure a consistent way of working for all 40 auditors taking part in this particular
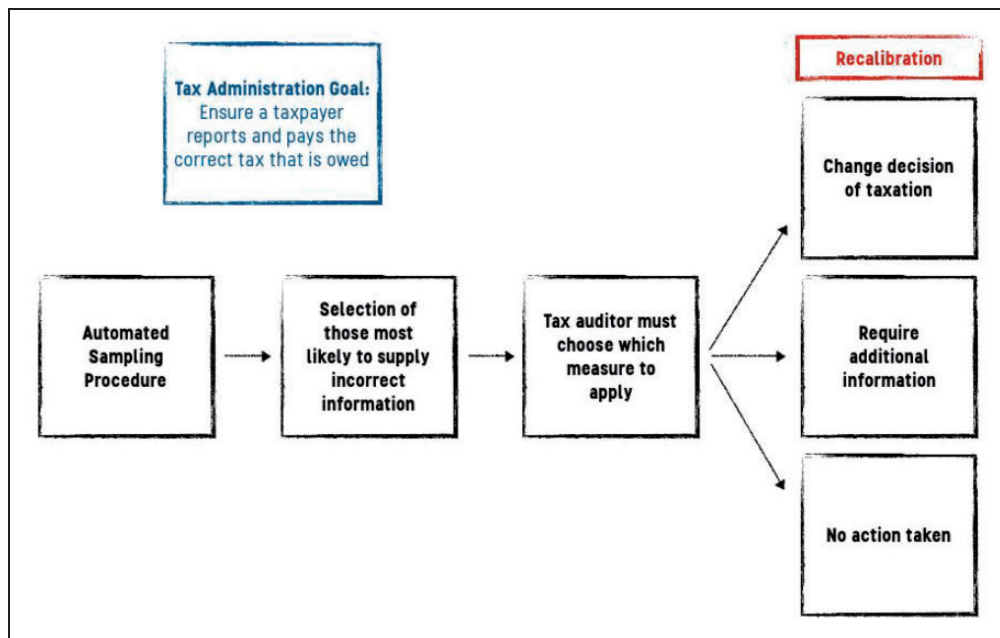

Figure 1. Audit control process.

audit control. Alice shows queries and memos made in different data systems, the binder with the auditees' physical documentation, the logbook recording each step of the audit, the diary registrations procedure she has to go through at each step of official contact with the audited entity, and the different letter templates for written communications. To ensure a uniform audit control there are 17 checklists in total: procedural instructions, information materials, and questionnaires to follow during phone contacts and visits. Ten working days are allocated to each audited taxpayer; two-thirds of the time involves administrative tasks and one-third concerns actual auditing.

Alice gathers information about the taxpayer from numerous registers and databases and notes them in the first checklist in order to make a plan to support a formal decision for the control audit of this particular taxpayer. A letter is sent to the taxpayer informing about the upcoming audit control including a visit. Alice follows each checklist but also works in her own way using her extensive experience. She says she increasingly uses her computer, happily acknowledging her two screens that allow her access to double the amount of information at the same time. For each audited taxpayer she has created a (physical) binder that fills up with material as the audit control proceeds. For example, after the visit she typically returns with the taxpayer's invoices and receipts, often accompanied by a data file with the ledger. Importantly, she also gathers knowledge about the audited subject during the telephone call(s) and by visual impressions at the visit.

Alice finds the audit visits very informative. It is really a gift to go on a visit, she says, as you then understand more about taxpayers' business activities. Seeing their 'reality' makes the audit control simpler. Most people are very patient, some are a bit unsure and afraid and then you have to be smooth. After she and her colleague pose their questions, which they are very careful not to make into an interrogation, they ask to bring supporting material back to the Agency office. Alice does not believe in the proposition that an orderly book-keeping reflects an orderly business. There are so many people who are very practical; they might be terrific performing their job which might not be mirrored in their book-keeping of these activities. There might also be sickness or other problems that make their book-keeping and other paperwork lag.

Alice tries to understand the everyday business reality of the people she audits and aims to be reasonable in her decision of which costs can be deducted. Her reasonableness commences with the first phone call and continues at the actual visit. Although they only want to see what concerns business activities, they often pass through the home. Alice can then see what can be considered assets. Are there paintings on the walls? How separate is the 'office' from the rest of the home? If any computers or cars are present, they can ask if these are also used for private use? If they see other things that are not needed for the business they audit, they can ask about them. Alice says:

> We should always pose the question "how do you use this thing?" This means that we listen to the taxpayer. We aim to create an understanding for the work we do, but also that we try to understand how they work. If we get off on the wrong foot, the audit might be unnecessarily long and cumbersome and we might irritate them. You know, we aim to interfere as little as possible so that they think we do our job well and with quality.

By understanding as much as possible of the taxpayers' business activities, the auditors can claim to be reasonable in their control. Glancing into the taxpayers' private life, Alice expects that the audited subject ought to behave in a cooperative and reasonable way and commensurate with his business role for – and especially with regard to what costs she/he deducts.

Back at the office, Alice starts the actual audit alternating between calculated information on her computer screens, the uploaded ledger from the auditee, a checklist of issues to look into. She considers and compares the relation between numbers that can indicate wrongdoings. Is there a large storage value but little turnover? Is a cash posting negative? Has a corrected VAT entry not been registered with the Agency? Are relations strange between different posts such as a lot of income yet little VAT? The actual audit is a continuous recalibration among and between various digital information.

Alice unpacks book-keeping posts, for example looking for invoices for purchases that seem irrelevant for the business, information that can also be found by browsing through some of the physical binder. She usually choose those for December or the end of the accounting year looking for purchases that seem irrelevant. It could be an invoice for a bed, bookshelves or other things that can be used outside business activities. Recalibration also includes her understanding of the taxpayer's business and the impression from the visit. In addition, Alice always has to pay attention to Agency values; she needs to be reasonable in her judgement. But there are limits to how understanding an auditor can be. Alice recalls an audit control of a consultant who had a big new screen prominently placed in his private living room, but the invoice was part of the book-keeping, a cost he could not deduct. A large TV screen could be very important for a business; Alice herself is very happy about her two screens. But here it is the location of the screen – in the taxpayers living

room – that makes it an invalid cost for the business. Correcting the tax return brings the taxpayer back to normalcy for the work of a tax auditor.

Recalibration in the quantification of self

Self-quantification is the practice through which people use (most often digital) technologies to track, measure, and record certain aspects of their behaviour and self. Digital self-tracking technologies exist in many different forms. Wearable technologies such as the Fitbit, Nike Fuelband, and Jawbone Up (devices that the user wears on the wrist to record steps taken using an in-built accelerometer), blood glucose monitors that measure blood sugar levels, wifi-enabled scales to track weight, mobile applications used to measure and analyse sleep patterns, monitor menstruation cycles, track calories, runs, and bike rides, are all examples of the types of self-quantifying technologies used.

The practice of self-quantification has evolved beyond 'early adopters' and has become much more mainstream since it was formalised in late 2008 through the group known as the 'Quantified Self'. On the face of it, self-quantification involves using a technology of some sort to quantify the self in pursuit of increased self-knowledge – the motto of the 'Quantified Self' is 'Self knowledge through numbers'. Often, people use devices which have algorithms built into them that turn the data into information, and this information will – either implicitly or explicitly – have particular courses of action associated with them. For example,

a blood glucose device may tell the user that they have 4 mmol/L glucose in their system and this, with the knowledge of 'normal' blood glucose ranges, might imply that a dose of insulin is required. Someone using a sleep app may find that their device advises or recommends that they go to bed earlier so that they have their full 8 hours of sleep. In the former example, the information was not given as a direct instruction for action, but knowledge about that particular measurement implies a certain action to be taken (if hypoglycemia is to be avoided). In the latter example, the instruction is much more explicit – it is telling the user what they ought to do (see Figure 2).

In many instances of self-quantification, these algorithmic outputs inform or advise users about particular behaviours or states of health. Much like the algorithmic recommendations at the Tax Agency, the algorithmic outputs here are not stand-alone readings devoid of context – as discussed above, these outputs may only *imply* certain realities or potentialities by virtue of the user's own knowledge surrounding the greater context about what they are tracking and what expectations of normalcy might be.

Emma, for example, is a member of the QS group interviewed by FD. She had been using the 'My Days' app on her phone to track her menstruation and her bowel movements. *My Days* is used to predict fertility levels, ovulation, basal metabolic temperature, cervical mucus, and other related data. Emma had been having irregular and infrequent bowel movements for around two years, and seemed to think that this was not an
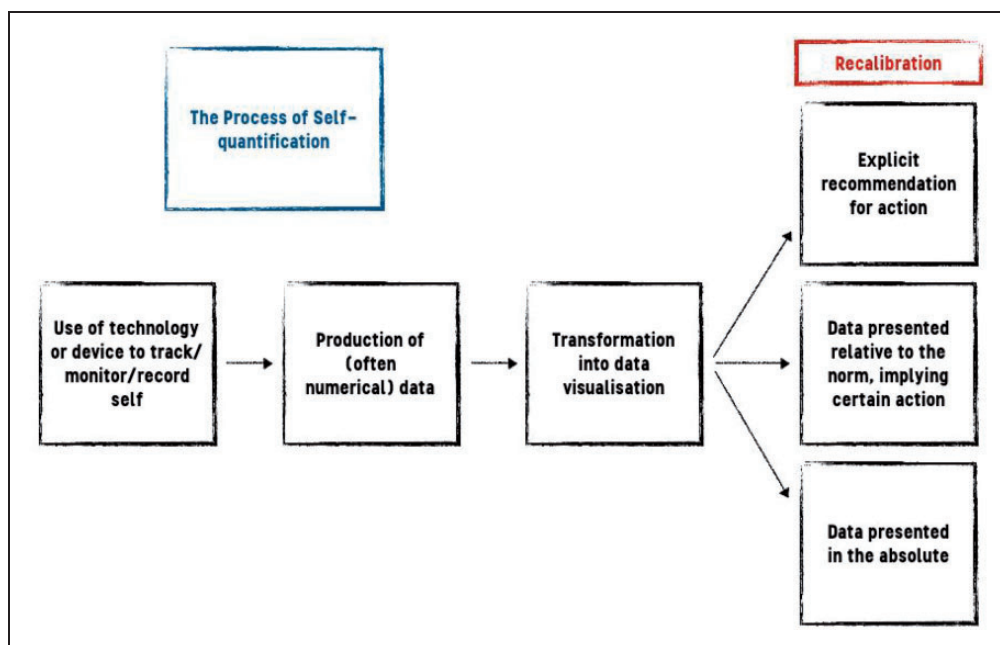

Figure 2. Process of self-quantification.

issue, in fact, she indicated that she thought it might even be convenient, at least until she started to use the *My Days* app:

> I used to have a problem, well I still have a problem, with constipation. And, that's another reason why I started using the MyDays things, and it's probably been like two years so I kind of sometimes think that OK I don't have a problem, but ... sometimes your perception is totally different from the facts when they're actually in front of your face.
>
> When I started putting stuff on MyDays app and started tracking when I go – when I use the toilet – and actually I've realised that I have a massive problem because sometimes there are weeks when I go to the toilet only twice a week! And if I kind of go that's great, but one thing I'm facing is Bowel Cancer, right?

Interestingly, Emma makes the instant distinction between her 'perceptions' and the 'facts' *in front of her face*. The facts for her are the algorithmically derived results gained through using her app. There is a very clear sense here that she puts these 'facts' above her own perceptions of what is going on. She uses the results from the app and considers them alongside assumptions about what is a normal number of times one ought to be opening their bowels, and what might be the consequence if one falls outside of these normal parameters. The recalibration of her constipation issues from being a scenario which may have been seen as somewhat convenient, to one in which she worries about the potential of bowel cancer shows how a discrepancy in algorithmic outputs and expectations of normalcy can result in changes in intuitions and the narratives given to felt experiences. In this case, her own feelings about the consequences of not opening her bowels more frequently and consistently (convenience) are recalibrated by the possible consequences implied by the algorithmic 'facts' laid bare through the charts on her app (bowel cancer).

Of course, the recalibration does not always have to side with the algorithmic output. There are many instances where the user might push back against what the device or technology indicates to them. Stephen, another member of the QS who was using a sleep-band and an associated app called Zio, described to FD how he grappled with tapping into his own intuition rather than automatically following the algorithmic reading on the device:

> FD: Has there ever been a time where you thought your sleep was one thing but then Zio would tell you something different?
> Stephen: Yes. Yeah that has happened.

> FD: And what sort of stuff do you do to ... what do you tend to agree with?
> Stephen: So there is a tendency to trust the technology a bit too much: so thinking that if Zio tells me I have had a bad night then it means I must have a bad night – and I think it's easy to give up our responsibility for that kind of that thing, to let the technology lead us, when really we should be first asking "Well how do I feel? Do I feel okay? And comparing that with "What does Zio say?" It's very easy to not do that I think so I don't know the answer to your question really as to what I go with.

Stephen questioned whether or not to take what the algorithm and/or technology said as being the 'truth'. He reflected that of course you should be asking about how you feel before you blindly accept what the algorithm and/or technology is telling you, but that it is difficult to do that part and much easier to just take the device for granted. The negotiation that Stephen has about whether to trust the algorithmic outputs, or to take himself seriously as a reliable judge of his own sleep is a great moment in which to show how recalibration involves more than *just* an algorithm.

Algorithmically determined sleep quality is, it turns out, a good example of when algorithmic output is often questioned using intuition and experience. Another QS informant, Tamar, showed how the reading on his device was rejected altogether as he compared his algorithmically derived sleep analysis with his own intuitions and recollection of how his night went:

> There have been numerous occasions when I've looked at that readout in the morning and it says "Congratulations you've got 98% efficiency." And I've sat here and questioned it because I think that I've not slept all that well. There are numerous occasions, I mean very frequently, where it says "You took 7 minutes to fall asleep." When I know damn well I was laying there for 20 minutes or more.

Tamar's reflection highlights an instance where the algorithm seems to be too far removed from the felt intuitions, and therefore the algorithm is not enough to convince him that he had a good sleep. There are also some instances where the algorithmic output is used as a way to initially calibrate in a situation, where there are little expectations in the first place. Take the example of Beth, a woman in her late twenties working at a 'tech-hub' in Berkeley, California. In this excerpt of an interview with FD, Beth talked about using a Fitbit, a device which records the number of steps taken per day by the user. It measures the distance travelled through these steps and the equivalent

number of flights of stairs climbed based on elevation. The device, which is roughly the size of the length of a human thumb and can be attached to a bra-strap or put in a pocket, connects via Bluetooth to a smartphone device on which the user's activity data is wirelessly uploaded and visualised.

> Beth: I was given this [Fitbit] from my dad … what's interesting is that, like, the main thing that I've been using this for is, at the end of the day, figuring out how many steps I took, and then kind of being able to connect that to how, like, I'm feeling, how my body is feeling. It's really interesting because I feel like the way that I feel after walking, let's say, 6000 steps a day, compared to, like, let's say, 26,000 steps in a day, a lot of times I, like, kind of feel the same. Like, I don't feel more tired. I wouldn't really know that I had walked a lot more on, like, day one versus day two. So that's really interesting. But I mean, the big thing that I really like as far as, like, the goal settings, so like, if I reach ten thousand steps in a day, it pauses and lets me know. So I'll definitely, like, make an effort to reach that every day. But you know, at the end of the day if I go home and I'm like, you know, 2–3000 steps short, I'm not going to take another walk around the neighbourhood to meet that.
>
> FD: Yeah, that's interesting. So, have you found that sometimes you haven't realised how many steps you've taken, when it said 26,000 as opposed to the 6000?
>
> Beth: Yeah, and I really would not know that's the steps. I mean, maybe I'm just really not self-aware, but it's been interesting for me to really just kind of connect, you know, number of steps to how my body's feeling at the end of the day … Because you would think that without this, I would somehow kind of … be able to know how much I've walked. That's really not the case.

Beth said that she was unable to tell the difference between days when she walks 6000 steps and when she walks 26,000 steps: she was not more tired, and she largely felt the same. She told FD that she felt as though she was not really self-aware and that she used the device to be able to connect the amount of steps she had taken in the day with how her body was feeling.

It is important to note that Beth does not refer specifically to the algorithm here, but only the device that she uses. However, within the device, there is of course an algorithm determining metrics like stride length, what counts as a step, categorising certain movements as being walking rather than cycling for instance, elevation gained, and so on. This algorithm is then converted into numbers and data visualisations that are shown on the smartphone app so that the user can be informed of all of these metrics. These are the very

metrics that allow Beth to 'connect' the amount of steps she is taking in a day with how her body feels. The last part of the conversation above really clarifies the point that the device and the data are really crucial for her to be able to know/feel how many steps she has taken in the day. She was almost surprised at herself that she did not seem to have any intuitive indicator of her daily activity without the app. There is a calibration taking place here: the device helps her to calibrate her intuitive bodily feelings with the number of steps she has taken as measured by her device.

Beth's case is at odds with other accounts of the effects of algorithms and data on the self. Other theories point to the potentially alienating features of such devices: Gardner and Jenkins (2016: 4), for example, argue that 'there are few technological experiences with more potential for creating a sense of disembodied alienation than seeing one's physical self portrayed two-dimensionally as data via algorithmic code.'. This Heideggerian view of technological alienation, however, seems to be refuted by our example here which shows the opposite to be the case. Rather than being alienating, Beth is in fact enabled by the device and data to be more in touch with her bodily feelings, supposedly giving her the ability to connect her daily activity with how she feels at the end of each day.

Further, there is another aspect to this device and the algorithm within it. And that is the part where it normalises certain behaviour and has the effect of changing the way Beth behaves to a certain extent, without being too overbearing. Beth says that the device will let her know when she has reached the goal of 10,000 steps and therefore she makes a concerted effort to try and reach that every day. However, the pull of this goal is not so overbearing that it compels her to get out of her house and go for an additional walk just for the sake of meeting that goal.

In this case, we see how calibration might occur with someone who does not have any strong sense of how they are feeling and so are happy to have the technology help them to connect their activity with bodily feelings. It is almost like an exercise in learning how to feel using the device and the algorithms. Recalibration is thus made possible in future uses of the Fitbit as her bodily feelings in conjunction with the device start to become more normalised and intuitive. We can imagine that deviations from these expectations may kickstart a whole new process of recalibration.

Discussion: Recalibration

These examples of counting and accounting share a very similar root: they are both practices as old as numerical digits themselves, and yet, the *digitisation* of these practices are changing the way in which

people engage with them. They both, therefore, provide wonderfully complex ground from which to explore and understand how algorithmic outputs are adjusted with tacit knowledge, intuition, and ideas about normalcy and normativity – in other words, how they are *recalibrated*. Throughout our respective ethnographic engagements, we have shown three ways that recalibration is predominantly done.

Relative recalibration takes place to identify the relations between numbers suggested by the algorithms. It follows an initial inspection as any given number can seem very big or small on its' own. This is where the recalibration starts. Alice, for example, considers different posts found in the taxpayer's yearly tax return statement and makes a comparison between them. If all numbers are present and the relations between them seem normal, she goes on to the next point on her procedural list, otherwise she pauses to probe further into the respective numbers. She unpacks a calculated number and moves on to a tacit recalibration. Elsewhere, Beth recalibrates how she feels about the number of steps she has taken when comparing that figure with the number of steps that her Fitbit tells her that she ought to have done (10,000). She may think that 4000 is a good amount, but in relation to the number that the algorithm has determined is a healthy number of average steps that a person ought to take in a given day, it seems much lower and she must decide whether or not she wants to make the effort to achieve a number closer to the recommended number.

Tacit recalibration illustrates the process where the algorithmic output is set in relation to what has been the human's own experience. Drawing on her own experience of past audit control visits, Alice 'knows' the set-up of a normal business office and the work being performed. Yet, the world changes and so do business opportunities which, according to Agency strategies in dealing with taxpayers, means being empathetic to the taxpayer's explanation of what is necessary in order to carry out the business. The surroundings and daily activities of a forest farmer are probably quite different from that of a PR consultant, yet they both have to account for their business activities and Alice has to understand each of 'their' realities and relate it to the calculated numbers present in each of their tax return statements. Engaging with the reality of a taxpayer includes browsing through the binders and comparing invoices with calculated amounts in the ledgers while comparing it to her impression of the visit as well as her understanding and knowledge of what 'normal' business activities entails. In the best of worlds, it is a normative stand, but Alice also needs to recalibrate her decision in terms of being reasonable.

Similarly, when Stephen and Tamar's respective sleep tracking technologies tell them that they had a sleep of certain quality, they start to ask questions about their experience and how they felt. They are unconvinced by the data from the device and start to engage with their intuitions and feelings to inform them of how their sleep was. Do they feel tired? Do they remember having many interruptions during the night? Do they remember lying awake for a long time before they eventually drifted off? Here, they take into consideration their lived experience and intuitions much more than taking the reading on their device for granted, and in both these cases they challenge the numbers shown to them.

Anomalous recalibration relates to abnormal, even dangerous, situations where the algorithmic output is trusted. The human decision-maker does not dare to question the algorithm's correctness. In the case of Emma and her bowels, she has a vague idea that she has *unusual* bowel movements, and at some point, far from seeing it as indicative of a potentially fatal condition, she even sees it as being convenient. After her app records and stores the information of her bowel movements over a period of time, she sees a pattern and becomes much more worried that they are not just unusual but statistically abnormal, and thus starts to worry about the potential of bowel cancer. She now sees what was once, perhaps, a convenient pattern is now a 'toxic' thing that is something to worry about. Whilst the algorithm here has meant that her intuitive feelings are given less weight than the app, which indicates to her that something may not be right, she is of course drawing upon knowledge of what she has learnt about bowel cancer, what the algorithm implies about 'normalcy', and what is an appropriate number of times to empty her bowels. There is recalibration going on here, but it has gone in favour of the algorithm.

When Alice realises that there are more serious irregularities between the taxpayer's book-keeping, his business activities and what has been reported to the Tax Agency, she needs to correct the taxpayer to bring the taxpayer's position back to normal – one without errors. Although the taxpayer is wrong in the given tax return, it is a delicate situation. She acknowledges the different capacities of these small businesses when she points out that the owners do not run a business due to their terrific book-keeping skills, but because they are good at the work they provide. She says she gives the taxpayers the benefit of doubt as to their record-keeping. But if the calculated numbers are too deviant from what a normal situation entails, she is compelled to be professional and act on the algorithmic output to correct errors.

Recalibration is thus seen as a way to understand the relationships between algorithmic outputs and ideas

about normalcy in both the case of the quantified selves and of the Swedish tax auditors. In these cases, normalcy and reasonableness can be seen as proxies for navigating the overarching normativities implied in each situation – whether this be the achievement of 10,000 steps and 8 hours sleep, or spending a reasonable amount of money on business expenses. Recalibration as a concept thus allows us to articulate the different normative notions that people consider when making decisions based on information gleaned from algorithmic output.

Conclusion

Through our ethnographic encounters with counting and accounting we have examined how decisions are adjusted between the output of algorithmic calculations and the experiences, and the tacit knowledge, of humans dealing with issues that stem from those calculations. We call this act of comparative adjustment *recalibration* as it involves making a decision about something – the algorithmic output – and then changing that decision given some other input, e.g. other numbers given by the algorithms, own experience, or contextual knowledge from elsewhere.

Often, the digital devices and systems that we focussed on made implicit or explicit recommendations for action which are, ultimately, powered by algorithms that take into account all of the data that is being collected and stored. Despite this, however, it was not the case that these outputs were unquestioningly thought of as gospel, and therefore that any action commanded by the technology was then taken. We found many instances where there was a 'technological dissonance' between what the user expected and what the technology implied (Dudhwala, 2017). In the cases of those quantifying themselves, if the user had one idea about the state of their body or some aspect of their behaviour, but the algorithm stated otherwise, a process occurred to align the two of them. Likewise, in the tax examples, the auditor often had an idea of the state of the audited taxpayer and their behaviour which was then compared to the algorithmic output.

We found that when the algorithmic output indicates something different from normal situations, be that the reasonably compliant taxpayer or an abnormal account of the self, there will be a recalibration, whether relative, tacit, or anomalous, to return the situation back to perceived normalcy. This involves deliberate work for the individual reacting to these abnormal situations, accounting and recalibrating far more than just the algorithmic output. Algorithmic outputs are thus never considered in isolation, but incorporate the individuals' tacit knowledge, experiences, and intuition to allow for moments where recalibration can occur.

Authors' note

Lotta Björklund Larsen is now affiliated with Tax Administration Research Centre, University of Exeter Business School, Exeter, UK.

Note

1. A large number of the decisions tend towards nothing further being done, and as such, it has been suggested that both the sampling and the following handling of audit controls could be made more efficient (Riksrevisionen, 2012). This official criticism of inefficient usage of algorithmic calculations implies a trust in the algorithmic output.

ORCID iD

Farzana Dudhwala (iD) https://orcid.org/0000-0002-4847-4542

References

Bateson G (1972) *Steps to an Ecology of Mind: Collected Essays in Anthropology Psychiatry Evolution and Epistemology*. Chicago, IL: Chicago University Press.

Björklund Larsen L (2017) *Shaping Taxpayers. Values in Action at the Swedish Tax Agency*. Oxford: Berghahn Books.

Callon M and Muniesa F (2005) Economic markets as calculative collective devices. *Organization Studies* 26(8): 1229–1250.

Deville J and van der Velden L (2015) Seeing the invisible algorithm. The practical politics of tracking the credit trackers. In: Amoore L and Piotukh V (eds) *Algorithmic Life. Calculative Devices in the Age of Big Data*. New York, NY: Routledge, pp. 87–106.

Domingos P (2015) *The Master Algorithm: How the Quest for Ultimate Machine Learning will Remake Our World*. New York, NY: Basic Books.

Dourish P (2016) Algorithms and their others: Algorithmic culture in context. *Big Data & Society* July–December: 1–11.

Dudhwala F (2017) Redrawing boundaries around the self and the body: The case of self-quantifying technologies. In: Lynch R and Farrington C (eds) *Understanding Personal Medical Devices: Critical Approaches*. London: Palgrave MacMillan.

Elyachar J (2012) BEFORE (AND AFTER) NEOLIBERALISM: Tacit Knowledge, Secrets of the Trade, and the Public Sector in Egypt. *Cultural Anthropology* 27(1): 76–96.

Gardner P and Jenkins B (2016) Bodily intra-actions with biometric devices. *Body & Society* 22(1): 3–30.

Gillespie T (2017) The relevance of algorithms. In: Gillespie T, Boczkowski P and Foot K (eds) *Media Technologies. Essays on Communication Materiality and Society*. Cambridge, MA: MIT Press, pp. 167–194.

Isin E and Ruppert E (2015) Doing things with words and saying words with things. In: Isin E and Ruppert E (eds) *Being Digital Citizens*. London: Rowman & Littlefield International, pp. 1–17.

Kitchin R (2017) Thinking critically about and researching algorithms. *Information Communication & Society* 20(1): 1429.

Kockelman P (2007) Enclosure and Disclosure. *Public Culture* 19(2): 303–305.

Kubler K (2017) State of urgency: Surveillance power and algorithms in France s state of emergency. *Big Data & Society* 4(2). https://doi.org/10.1177/2053951717736338.

Latour B (2005) *Reassembling the Social: An Introduction to Actor-Network-Theory*. Oxford: Oxford University Press.

Lyon D (2014) Surveillance Snowden and Big Data: Capacities consequences critique. *Big Data & Society* 1(2): 1–13.

McQuillan D (2015) Algorithmic states of exception. *European Journal of Cultural Studies* 18(4–5): 564–576.

Pink S, Ruckenstein M, Willim R, et al. (2018) Broken data: Conceptualising data in an emerging world. *Big Data & Society* 5(1): 1–13.

Pink S, Sumartojo S, Lupton D, et al. (2017) Mundane data: The routines, contingencies and accomplishments of digital living. *Big Data & Society* 4(1): 1–12.

Power M (2000) The Audit society – Second thoughts. *International Journal of Auditing* 4(1): 111–119.

Riksrevisionen (2012) *Skattekontroll av företag. Urval, volymutveckling och uppföljning*. Stockholm.

Seaver N (2013) Knowing algorithms. *Media in Transition 8*, Cambridge, MA, USA. Available at: http://nickseaver.net/papers/seaverMiT8.pdf (accessed 21 August 2014).

Skatteverket (2013) The Swedish Tax Agency's Strategic Direction. Solna.

Strathern M (2005) *Kinship, Law and the Unexpected: Relatives Are Always a Surprise*. New York: Cambridge University Press.

Stevenson A and Lindberg C (2010) *New Oxford American Dictionary*, Third Edition. USA: Oxford University Press.

Skatteverket (2014) The development of the tax gap in Sweden 2007–12. Solna.

Striphas T (2015) Algorithmic culture. *European Journal of Cultural Studies* 18(4–5): 395–412.

Wirth N (1977) Design and implementation of modula. *Software: Practice and Experience* 7(1): 67–84.

Zarsky T (2016) The trouble with algorithmic decisions: An analytic road map to examine efficiency and fairness in automated and opaque decision making. *Science Technology and Human Values* 41(1).

Ziewitz M (2017) A not quite random walk: Experimenting with the ethnomethods of the algorithm. *Big Data & Society* 4(2): 1–13.