Calculating Therapeutic Compliance: An Ethnographic Account of Numerical Inference and Interference in Mobile Health Care

Martina Klausner

Department of European Ethnology, Humboldt University Berlin, Germany/ martina.klausner@hu-berlin.de

Abstract

This article discusses calculation practices in the development of a monitoring device, aimed at improving therapeutic compliance of children and teenagers suffering from a deformation of the spine. In managing the complexities of physical parameters, therapeutic measures, and interventions in everyday life, numbers are central participants in inferring from and interfering with bodies and behaviours. Numbers constitute the input and output of such monitoring systems, translating, circulating, and visualizing physical conditions and therapeutic effects, as well as suggesting action. This generative process of capturing and interpreting data has at the core algorithms, which process data and provide seemingly unambiguous numerical outcomes, based on mathematical and technological means of processing information. Attending to the incremental process of "learning algorithms" as a central feature of the system's development allows me to describe the robustness of certain modes of inference. Over and above using a specific case as an example for computer-based numerical inference and interference, this article attempts to probe and complement two theoretical approaches to the numerical management of complexity: Helen Verran's (e.g., 2001, 2010, 2013) focus on numbers' performative properties and the potential tensions arising from divergent numerical orderings, and Paul Kockelman's (e.g., 2013a, 2013b,) sieving of inferential and indexical chains along the generation of meaning and ontological transformativities.

Keywords: monitoring systems, numerical inference, ontic tensions, ontological transformativities, algorithmic processing, (non-)compliance

Introduction

Therapeutic monitoring systems are an increasingly common way of capturing, translating and visualizing physical parameters and the effects of therapeutic efforts on patients. One of many e-Health-technologies, such monitoring systems usually comprise of a sensor system and some form of communication interface to provide patients with feedback, prompting them to evaluate and potentially change their behaviour. In managing the complexities of physical parameters, therapeutic measures, and interventions in everyday life, numbers are central participants in inferring from and interfering with bodies and behaviours. Numbers comprise the input and output of these monitoring systems, translating, circulating, and visualizing physical conditions and therapeutic effects, as well as suggesting action. It is this participation in inference and interference, which makes these numbers particularly interesting. In my article I attend to the management of therapeutic complexities through numbers in one exemplary case. Having worked with engineers and doctors and their numbering practices in developing a therapeutic monitoring system to improve compliance, as well as with patients (in this case children and teenagers), and the mundane calculation practices which make up part of their therapeutic effort, I was interested in the way the monitoring system calculations would relate to and possibly transform the embodied calculation of patients. What happens if digital processing becomes the central basis for the therapeutic management of complexities?

In the case at hand, the aim of the therapeutic monitoring system is to improve patients' therapeutic compliance by giving them precise feedback on their therapeutic performance. The feedback focuses on the actual time patients are pursuing therapy, in this case, the time during which children and teenagers wear their braces to correct for scoliosis, a deformation of the spine. As the daily duration of time for which the brace is worn is considered to have a central impact on therapeutic outcome, complying with the recommended duration (usually between 16 to 23 hours a day for at least two years) is regarded as crucial. At the same time, this therapeutic prescription proves to be an enormous challenge for these young patients as the brace itself disables them in various ways: the brace is of a rigid plastic causing pain in some cases, making the teenage wearers sweat especially during the summer, limiting their mobility and activity, while some see it as an aesthetic imposition. To prevent non-compliant behaviour, the monitoring system would capture the daily number of hours the brace was worn and would provide real-time feedback of the current number of hours to the patient via a smartphone app. Feedback in the form of an objective number is thought to potentially increase motivation and compliance to the advised hours. I am writing in the subjunctive as this article is not about the routine use of such a system. It is about the development of the therapeutic monitoring system and how numbers are generated by and generate assumptions about therapeutic compliance. Algorithms are at the core of this generative process of capturing and interpreting data, constituting a common basis for managing the complexities created by vast amounts of data and providing the basis for interference in people's lives. They process data and provide seemingly unambiguous numerical outcomes, based on mathematical and technological means of processing information. Attending to the incremental process of "learning algorithms" as a central part of the system's development allows me to describe the robustness of certain modes of inference. While the system in question is specific in many ways, several aspects of the development process and the device point towards more general discussions of monitoring systems in therapeutic contexts.

Over and above using the specific case as an example for computer-based numerical inference and interference, this article attempts to probe and complement two theoretical approaches to the numerical management of complexity. Focusing on the performative properties of numbers, Helen Verran's (e.g., 2001, 2010, 2013) work invites us to attend carefully to numbers in their ability to generate multiple relations and generalizations. Offering an entire toolbox for the investigation of the performative properties of numbers, Verran (2013) allows us to look at numbers as being generated in and at the same time being generative of collective action and order. This dissecting of numbers and their performative properties is the means to recognizing the "ontic and ontological tensions" in numbering practices and might bring us to recognize and engage with contesting political ontologies (Verran, 2014; for an alternative analysis of Verran's notions of ontic and ontology see Lippert, 2018). This interest in the tensions between different "enumerated entities" (Verran, 2010) leads us to focus on the perpetually 'becoming' nature of numbers and their capacity to change and flip from one generalizing mode to another, one semiotic manifestation to another.

Alongside the thrust of this special issue, numbers appear to be rather unstable entities. Yet my research on the development of the monitoring system and especially the mathematical and technological processing of data into an automated system made me aware of the robustness and durability of certain numbers. Central to this is the equipment of these automated numbers with a durable inferential profile that is the basis for interference. Understanding the configuration of inferential profiles is crucial for any intervention in the politics of algorithmic numbering. My focus on the robustness of certain numbers does not principally question the various dynamics of numbers' instabilities, as demonstrated by other papers in this special issue. Rather, there are several overlaps: for example with Neyland's (2018) paper on the development of a surveillance system and the attempt to automate deletion of unnecessary data by algorithms. His case study is a crucial reminder of the failures of such development processes and the undoing of calculation and qualculation (Neyland, 2018). Still, even though such projects fail or develop in other ways than planned, "successful", meaning solid working algorithmic processing is not an exception, on the contrary. To further delve into the creation of robust algorithmic processing and inferring, I propose a complement to Verran's performative properties. To do so, I draw from the work of anthropologist Paul Kockelman (2017) who is not explicitly concerned with numbers but with inferential processes especially in computer-based generation of meaning. Drawing from semiotics (as does Verran), among many other theoretical sources, Kockelman (2017: 128) invites us to attend to the "tangled, indexical and inferential chains, mediated by machines and algorithms as much as by humans". Processing information inevitably enchains "ontological transformativities" (Kockelman, 2013, 2017). Taking up his focus on interpretation or inference, I will discuss how the numbers produced by the monitoring system work as 'inferential devices': how they interpret and transform assumptions and action while being generated by ontological assumptions.

In my article I will employ both analytical offerings to discuss my material and will thereby hopefully offer a productive permutation of their ideas through my case study. Rather than comparing the two, my aim is to produce a recombinant, a combination of Verran's performative properties of numbers and Kockelman's inferential profiles.

Figuring out numbers in STS as/in relations of relations

Numbers' remarkable capacity to represent truth and objectivity (Porter, 1995; Hacking, 1990) has been scrutinized and amplified in the last years by various investigations in Science and Technology Studies, attending to the various work numbers do in different professional and mundane domains: from counting in classrooms (Verran, 2001) and supermarkets (Lave, 1988; Cochoy, 2008), to calculation practices and devices in economics (Callon and Muniesa, 2005) and financial markets (Zaloom, 2003), to governing practices in environmental politics (Asdal, 2008) and policy making (Ballestero, 2015) to epidemiological modelling based on enumerated entities (Bauer, 2008; Mackenzie, 2014), to mention some central fields of study. These studies have a particular interest in the ambiguity and performativity of numbers: how numbers are produced, how they circulate, legitimize authority, and constitute the realities they claim to represent. Rather than presenting numbers as belonging to one form of practice and contrasting it with another, those studies also highlight how numbers, as highly mobile devices, not only travel across divergent fields of action and styles of reasoning, but are also productive in creating new relations across different fields.

Yet the focus on numbers' capacity to circulate, to relate, to merge diverse systems and practices, also bears the risk of diffusing what we actually mean when we talk about numbering practices. As Ballestero emphasises, a calculation grammar, the arrangement of "people, technical instruments, and semiotic signs", is not only highly dependent on the concrete technical properties of its infrastructures but also on the "mathematical implications" it invokes (Ballestero, 2015: 266-267). What I summed up as studies on numbering practices appear to be rather diverse in the concrete procedures they invoke: quantifying, accounting, calculating, equating, valuing. Are these similar practices simply because they all invoke numbers? To what degree do we need to differentiate between the concrete procedures and mathematical inferences invoked by 'numbering'?

By taking up this question I heed the plea made by Helen Verran who has emphasised the need to differentiate the diverse work numbers do. Verran has written extensively on numbers as "lively semiotic-material actants" (Verran, 2012: 66) in various fields and has defined them in her approach as "materialized relations" (Verran, 2010: 171). In her book *Science and an African Logic*, which generated years of focusing on numbers, Verran (2001) attends to numbering practices in Nigerian classrooms and the educational efforts to teach Yoruba speaking and English speaking children how to calculate in a scientifically sound, Western way. Disconcerted both by her research and by the relativist stance towards seemingly disparate numbering logics in her first draft, she variegates her own argument on numbers, developing an (re) account of numbers as multiple relations. Drawing from Marilyn Strathern's discussion of "the Relation" in kinship studies, Verran shows how numbers *have* relations and *are* relations, they are generated in and generative of collective action and order (Verran, 2001: 100-101). Consequently, numbers are addressed "as particulars, in time and place, *in situ* we might say – materialized; realized in specific practical ways" (Verran, 2010: 172). In this way, numbers are considered to be participants in collective actions where their performative properties variegate according to the microworlds in which they are embedded and embodied, the imaginaries they evoke and are evoked by, and the specific orderings they are engaged with.

Over the years Verran (2013: 28) has elaborated and refined the "epistemo-cultural properties of numbers" and their performative effects. Two of those properties proved to be especially relevant for my analysis: the semiotic manifestation of numbers (functioning as icons, indexes, or symbols) and the modes of generalizations they perform (whole-part or one-many generalizations). For example, the differentiation of numbers as icons, symbols, and indexes highlights which specific linkage between sign (e.g. number of hours the brace was worn;), object (e.g. the child wearing the brace), and the interpretant[1] (e.g. the correlation between hours and (non-)compliance) can be made. Are numbers in the therapeutic monitoring system working as icons, co-constituting (non-) compliance as a new whole? Or does calculating hours remain in the indexical zone continually referring to the embodied and situated practices of patients and doctors?

As I will demonstrate by juxtaposing the constitution of hours worn as duration (in the monitoring system) and time rhythm (in patients' mundane calculations), a focus on the various modes of performativity of numbers allows for a nuanced analysis of what is at stake in numbering practices. In Verran's work numbers' manifestations are rarely stable, but rather, they have the capacity to shift from one manifestation to the other, "flipping imperceptibly from their one-many manifestation to their whole-parts form of working, shifting between signing as symbols and signing as icons" (Verran, 2010: 177). Because of the limitation of my ethnographic research to the making of the system and the not-yet routine use of the system at this point of time, I am not able to venture into the potential "imperceptible flipping" of numbers' manifestations in therapeutic routines using the system. What I try to do here is slightly different. I wonder about the details of changing the performative properties of numbers from one relation to the other, from one semiotic manifestation to the other in the incremental machine learning process of the system, and ultimately about how they become durable. Elaborating on this specific case of machine learning, I attempt to open up the black box of algorithmic processing and similar to Adrian Mackenzie's (2017) work *Machine Learners*, though less comprehensive and archaeological, analyze algorithmic processing as a specific form of knowledge production and meaning making. How do local relations feed into generalizations and how do machines learn to make one-many generalizations? And how do these generalizations manifest as durable and opaque in this process? I argue that the potential force of algorithmic processing is the way in which numbers become durable, equipped with seemingly unequivocal inferential profiles. Inserting inferential profiles into Verran's toolbox for analysing enumerated entities to understand their relative obduracy is the central recombinant I offer in this text by engaging with the work of anthropologist Paul Kockelman.[2]

Kockelman draws from a variety of theoretical sources, all centring around the idea of 'relations between relations', to systematize various processes of selection and significance as an on-going process of transformation (Kockelman, 2011, 2013a, 2013b). With his background in linguistic anthropology and his comprehensive systemization of a whole range of phenomena on various scales[3], he offers an at-times irritating mixture of stringent formalization and constant movement and shifting. There is no starting point, only continuous transformation; and yet he attempts to figure out the recurrent pattern in this complex "multiverse" of relations of relations. While he insists on the potential to use this analysis of patterns across various scales (Kockelman, 2011), much of his work is focussed on the very minutiae of information processing: the gesture, the utterance, the spam mail, the ticking of a clock (Kockelman, 2011, 2013a, 2013b; Kockelman and Bernstein, 2012), and therefore proved to be especially productive for attending to the details of computer-based inference and interference.

Like Verran, Kockelman draws on Peirce's (1974) theory of signs to set out his conceptual framework. In his conceptualizations he offers some helpful analytical categories and their linkages to elaborate on types of inference: the index, the kind, the individual, the interpreting agent, the ontology. Using the workings of spam filters as an example, Kockelman (2013a: 45-49) elaborates five forms of equations that form the basis for algorithmic processes (and assumption making more generally) and produce distinctive kinds of "ontological transformativity". With the notion of ontological transformativity Kockelman (2013a: 33) wants to foreground "the way ontologies are both embodied in and transformed by such algorithms". How are assumptions of the world, of an individual and accordingly indices incorporated into algorithms? And how does algorithmic processing loop back into ways of being and meaning making? Even though Kockelman is not explicitly dealing with numbers, his systematic analysis of relations of index, kind, individual, and agent and the transformation of ontologies complements the analysis of the workings of numbers in the concrete calculation practices I address. To discuss the 'inferential profiles' (Kockelman, 2013b, 2017) involved in the working of spam filters, Kockelman concentrates on the three modes of inference: inductive, deductive and abductive modes of inference. My central question concerns how and when numbers produced in such equations shift from one mode into another and become increasingly robust and durable. Kockelman's systematic semiotic vocabulary is productive in tracing the workings of mathematical and technological information processing through algorithms as continuous "inferential and indexical chains" (Kockelman, 2017: 128) which have the unique capacity to appear free of context given their technicality and high level of abstraction.[4]

While Verran and Kockelman are both interested in the way objects such as numbers or signs more generally come into being, gain and generate meaning, they engage different analytical devices to work through complexity management: Verran's (e.g., 2001, 2013) analytical device is a focus on the disconcertment arising from an encounter between different modes of numerical orderings. Juxtaposing the various performative properties of numbers, she emphasises the "ontic and ontological tensions" in these encounters, enabling the analyst to scrutinize the ontological politics at stake. Kockelman's (e.g., 2013a) conceptual device, on the other hand, is that of sieving endless connections and gradual shifts or, as I would call it, the incremental processing of ontologies and interpretations, altering both. Both of these approaches were productive for me long before this article, as they enabled me to look at different things or rather to look at things differently. My aim is not to symmetrically compare but rather to recombine their analytical approaches to generate an analytical device which captures what is at stake in the algorithmic processing and calculating of therapeutic compliance. This implies more than simply adding another concept, but rather invokes a recombination of the emphasis on *ontological multiplicity* and the arising frictions and *ontological transformativities* – the gradual transformation of ontological assumptions, especially in automated inference processes.[5] Part of this recombination will entail the integration of Kockelman's inferential profiles into the set of performative properties of numbers as proposed by Verran. I thereby propose to attend carefully to numbers' capacity to produce relatively stable assumptions about the world.

Yet my attempt to use the case study to effect engagement between the two by mapping Verran's conceptual framings of numbers and their 'epistemo-cultural properties' onto Kockelman's sieving process of 'ontological transformativities' created a strange effect of a gestalt-switch. This was productive for my analytical process, but rather difficult to capture in a linear text. In this article, I engage first with Verran's (2014) 'ontic tensions' to emphasize what is at stake in the development of the therapeutic monitoring system and the production of 'objective' numbers through the device. This allowed me to pause and scrutinize the performative properties of the device at the end of the development process by juxtaposing it with the mundane calculation practices of the young patients in their therapeutic routines. Subsequently, I will attend to algorithmic inference as incremental calculation. Here I retrace the development process and follow Kockelman's (2017) invitation to move along the chain of transformations to show how the generalizing mode of the numbers at hand is constituted through a cumulative process comprising sensors, a spreadsheet, human reasoning and tinkering, source code, among others. This retracing enables me to discuss how the coming into being of one-many relations occurs and how ontological politics are put into practice. The conclusion details the effects of this gestalt-switch. Before I proceed, I will provide more insights into my research and my case study.

Calculating Compliance: Case and Method

My insistence on the "gestalt-switch" is also a result of my research commitment. The aim of my ethnographic research was to accompany the technological developments of a large research cluster[6], including the monitoring system, in two ways: serving as a so-called ELSI (ethical, legal, social implications of technology development) project, the task was to address potential blind spots in the design and development of the technologies and introduce a broader critical reflection of the potential effects of such technologies. At the same time my aim was to provide ethnographic insights into potential users' expectations and experiences and to feed those findings back into the development process. One of the technologies developed in the research cluster was the monitoring system I focus on in this article.

This so-called "multi-sensor monitoring system" addresses the potential non-compliance of children and teenagers in scoliosis therapy. The aim of the system is to provide patients with feedback on their therapeutic performance via a smartphone application, assuming it will enhance their therapeutic compliance. In therapy and in the concept of the monitoring system compliance is defined by the adherence to the advised number of hours the brace is to be worn. Brace therapy is a common treatment for milder variants of scoliosis, which is a three-dimensional deformation of the spine, usually developing in the early teens and most responsive to corrective therapy during the growth phase. Hence, children and teenagers are the main patient group. Depending on the degree of spine deformity and the point in treatment, children are advised to wear the brace for 16 to 23 hours everyday for several years, whereby the rigid plastic "presses" children into the upright position. To no surprise, scoliosis therapy is demanding for children and teenagers in various ways, and adhering to therapeutic advice cannot be considered self-evident. Yet clinical studies and orthopaedic guidelines suggest a direct correlation between the number of hours it is worn, therapeutic outcome and potential long-term impairment.

To increase young patients' compliance to the advised number of hours a research team involving engineers, psychologists, orthopaedics, computer scientists, and usability designers developed the so-called multi-sensor monitoring system. The system comprises two main parts: a sensor system built into the brace to measure certain bodily values (temperature, moisture, acceleration, pressure) for the calculation of the hours the brace is actually worn; the outcome of this measurement is then provided to the patients via a smartphone application. In more or less real time monitoring, children and teenagers are provided with a visualization of the actual hours and are expected to adjust their wearing performance accordingly. While the developers of the system acknowledge various factors leading to non-compliance, the main risk factor is considered to be the incapability and lack of motivation on the part of young patients to realistically estimate their hours through the day. Providing them with an "objective number" is considered to increase their motivation to comply. Additionally, the app offers information on scoliosis therapy, is equipped with an exercise programme, and a pin board with user stories as well as tips for the daily use of the brace.

My commitment to a critical dis/engagement with the development project shaped my argument in various ways. Sharing the task of providing patients with better solutions and assistance in handling the impositions of brace therapy, I was cautious of the powerful effects of monitoring users' therapeutic performance with respect to their (non-)compliance. To critically engage with these potential effects I carefully attended to the different calculation practices in the project and their potential implications. Through participant observation in therapeutic settings, such as a children's orthopaedic hospital specializing in scoliosis, and through open-ended interviews with 44 young patients (some at home, some in an orthopaedic hospital, some at the brace manufacturer), I learned about the young patients' everyday therapeutic routines and their struggles to adhere to the therapeutic advice. At the same time I regularly attended working meetings of the development team, conducted a series of interviews with the engineers, contributed to an observational study of the monitoring system at the end of the project and provided feedback on my preliminary findings to the development team during the process. So all along the research process I was constantly juxtaposing the therapeutic routines of the patients and the development process of the monitoring system. Numbers were crucial in both: in patients' day-by-day efforts to attain the expected number of hours; and in the project team's efforts to produce objective numbers by measurement and algorithmic processing of the sensor data.

This constant juxtaposing had two effects: first, it made me aware of the different generalizations of wearing time: In the development project, wearing time referred to the overall duration of hours the brace was worn within a 24 hour window; it was a total; in patients' therapeutic practices, wearing time mainly referred to the concrete time in the course of a day, a passage of time. Yet the second effect was somehow the opposite: Attending to the struggles of patients in their daily lives and engineers as they went about their daily work made me aware of the similar messy grounding of algorithmic and embodied calculations. Both "need to wrestle with the (…) buzzing real", as Verran (2012: 120) has phrased it. Just as patients have to learn to calculate "correctly", the monitoring system – and its programmers, engineers, and algorithms – had to gradually learn to translate and perform numbers in a specific way. Both are ways of managing complexity. Yet with regard to the monitoring system, the learning had to come to a closure by the end of the development project. The calculation of wearing time through the system had become valid and robust. Before I elaborate on the development process, I first attend to the system as it was at the end and juxtapose it with a young patient's mundane calculation practices.

Juxtaposing hours worn as duration and time

Let me start with the monitoring system as it was working at the end (as a prototype). The central question concerning the system was how many hours per day the teenager wore the brace and whether this conformed to the therapeutic prescription. So we have a number somewhere between 0 and 24. I choose '16' as an example, which is a fairly good wearing performance, but not entirely perfect. The hours worn are visualized on the screen of the teenager's smartphone in the form of 16 cute little kittens (see Figure 1). For each additional hour the brace was worn according to the sensor system, one more kitten appeared during the day. In the weekly overview the actual hours it was worn each day are presented in the form of a bar chart where a green line indicates the advised hours (see Figure 2). The number of the advised hours headlines the chart and the actual hours, e.g. 16, are displayed above the bar for each day. But how does the system calculate the numbers of kittens and bars? Generally, the kittens (and bars) are the outcome of a classification system: no kitten stands for "not worn" and a present kitten for "worn". It is a simple binary system – a 'yes' or a 'no'. The system classifies the captured data into yeses (1) and noes (0) and adds up all the "yeses" for a total sum, e.g. 16. Sensors do not capture data continuously, but every five minutes.[7] So for example, the sensor system might capture data at 9:55 a.m. and then classify whether the user was wearing the brace or not at this point in time according to the underlying algorithm and presumes that this classification is correct over the five minutes until the next measurement. Once there are 12 points of measurement classified as "yes, worn", another kitten appears on the display (12x1x5=60). This might be at 10:50 a.m. or much later. There is no quarter or half kitten. The system first expands point-measurements to 5-minute intervals and then adds those times (duration) up for an overall duration within 24 hours. In my conversations with the engineer who was developing the measurement system and the algorithms for the classification, she made me aware of an important distinction: Even though in the project we generally spoke about "wearing time" monitored by the system, it is actually "wearing duration". It is not the actual time of day, e.g. 9:55 a.m., that is important but the summation of discrete time units to quantify the duration the brace was worn. Duration in the system's construction therefore does not refer to an interval between two points in time, which might correspond to a more intuitive understanding of duration, but a cumulative length, consisting of discrete units. Before I further delve into this difference between time worn and duration worn, let me pause and explicate the 'performative properties' of the numbers produced in this calculation practice. As stated above, Verran (2013: 28) has elaborated and refined the "epistemo-cultural properties of numbers": their modes of generalizations, their ontological manifestation as well their semiotic manifestation

Figure 1: Screenshot of current hours worn

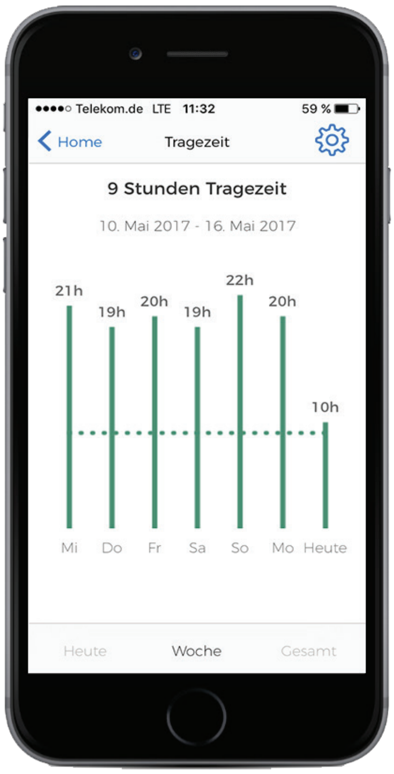

Figure 2: Screenshot of summary of weekly hours worn

Both figures were provided by the project partner at Berlin's University of Arts and the temporalities by which these modes are modified. Two properties proved to be especially helpful in my analysis of the numbering practices in the case at hand: the semiotic manifestation of numbers and their mode of generalizing unity-multiplicity relations.

Drawing from Peirce's theory of signs (yet twisting it in her own way, see for example Verran, 2010: 172), Verran invites ethnographers to attend to the different workings of signs as symbols, icons or indexes, and the specific co-constitution of signs, collective action, and the objects generated within these workings. While indexicality strongly implies the here and now and the existential co-constitution of object and referent, symbols and their objects perform a relation of supervenience, whereby "objects are accepted as affecting and effecting their signs but not vice-versa", as Verran (2010: 172) explains. Numbers represent objects, phenomena, 'reality'. Iconicity, the third semiotic mode of numbers, in contrast, highlights a collapse of any distinction between number and category and their capacity to generate order. Here sign and object are treated as one and alike (Verran, 2012: 116). This first distinction of numbers' performative properties allows me to analyse numbers' workings as different manifestations of the co-constitution of signs, the objects, and the collective actions in which they are embedded: literally pointing towards what is being counted (indexes), representing it (symbols) or constituting order (icon). These semiotic manifestations of numbers are intertwined with the way generalizations are performed through them. Starting with indexical numbers which "dwell in the mess of the real (…) generalizing can proceed simultaneously as whole-parts and one-many" (Verran, 2012: 120). Something is being (re)counted in the here and now, and from here generalizing can proceed in two ways. Performing one-many relations is a common generalization technique in many scientific practices and beyond; starting with discrete units, which are collected to a coherent cumulus of many; the resulting numeral (e.g. 16) abstracts and represents the plurality of many. Quite differently, whole-part generalizations refer to multiple emergent parts from a vague whole. While in the former, numerals are representations/symbols, in the latter they become iconic, constituting the world.[8]

Let me further discuss this in relation to the calculation made by the system described above. The calculation starts by capturing data every five minutes. The data, e.g. the temperature of the brace at this specific point in time, is processed by an algorithm for classification under yes (1) or no (0). In algorithmic processing, the existential relation between sign and its object (e.g. the temperature in the brace expressed as 36,9°) is transformed into the conventional binary system of 1 or 0. So while 1 still refers to the data of bodily parameters, its binary reworkings manifest it as symbolic. In the logic and processing mechanism of the monitoring system, calculation proceeds by adding discrete units to a total, first to a full hour and then to the overall duration the brace was worn during a day. In summing up discrete units of a defined measurement, the monitoring system and the produced number perform 'realistic', objective' representations where physical processes, e.g. temperature, are transformed into signs and visualized as numerals. As Verran (2010) has stated, in this semiotic manifestation as symbols, objects effect signs but not the other way around. This calculation appears to be a solid technical calculation process with clearly defined units and an unambiguous outcome. The reference point or frame is the 24 hour day, yet not as a course of time but duration as the cumulation of hours which are the result of sums of smaller time units classified as "yes, worn." It is a metric version of time, consisting of quantified units. In the generalization of a one-many relation, the number 16 manifests as a symbol. This is how many hours the teenager "really" wore the brace. Overall, this is the aim of the monitoring system: to provide children and teenagers with objective numbers to correct their often unrealistic calculation practices. In this sense, the 16 kittens are a truth claim, based on algorithmic processing of physical parameters, a seemingly neutral mathematical and technical procedure beyond subjective bias and human errors. In a combination of the connection between the medical correlation of numbers of hours worn to therapeutic outcome and the technical processing of physical parameters to an objective representation of therapeutic effort, numbers come to demarcate therapeutic compliance in an unambiguous way. It is either yes or no, either 1 or 0, and further, based on the unambiguous measurement of the total duration per day compared to the advised hours, it results in another binary interpretation of either compliant or non-compliant. There is no room for alternative interpretations or excuses, (non-) compliance is a fact.

The monitoring system is designed as an intervention in the wearing routines of the patients.[9] To inquire into the potential implications of the system, part of my research attended to the actual wearing and calculation routines of the patients. As I will show in the following section, children and teenagers (mostly) agreed on the difficulty of calculating wearing hours correctly (and were enthusiastic about using such a system); yet most of them nevertheless felt competent to handle the calculation of wearing time "most of the time". Their calculation, however, differed from the time measurement of the monitoring system. To elaborate on this difference, the generalizations made and the relations numbers are/have in these practices, I will juxtapose the performativity of numbers in the monitoring system with the embodied calculations of the young patients. What is the potential 'ontic tension' which might arise with the implementation of the monitoring system in everyday therapeutic routines? How might it interfere in the patients' everyday calculation practices?

Calculation practices in the 'indexical zone'

At first glance what children and teenagers were interested in was also duration worn. Their primary concern was: "Did I wear it enough?" For most of them, the actual number of hours they wore the brace during the day turned out to be hard to grasp. When asked whether they thought they could realistically estimate their hours, a common answer the young patients gave was: "Sort of, yes" or "most of the time". Inquiring more into their everyday practices of wearing the brace, this "sort of" and "most of the time" proved to be quite a challenge: "You think you just took it off for a couple of minutes and then it turns out it was more than an hour," a teenager explained. Wearing the brace, time seems to run slower, it makes everything more difficult: "I thought I wore it for ages but then recounting the time with my mom, it turned out it was just for an hour." Also those who were convinced they "sort of" knew most of the time, admitted that sometimes they got the hours totally wrong. Overall, the counting and calculation of hours worn very much depended on the daily routines and the regularity of their daily activities. In a way, the focus on the question "is it enough?" points towards an understanding of wearing-time as gradual and relative to a value-schema and highlights the situated judgments of these young patients. Yet the judgment around 'enough' is always made vis-à-vis clearly defined (by doctors, parents, therapists) quantities, which serve as reference.

This struggle to grasp the hours worn and to count "correctly" became obvious when several teenagers finally had the opportunity to test the prototype of the monitoring system at the end of the development project. All of those who participated in the study embraced the idea of having real-time feedback of their hours through the monitoring system. Finally, they could "see" the wearing-time, was an often made comment. They could finally see what was otherwise complicated to perceive. And they enjoyed collecting kittens and found it a fun challenge to accumulate as many as possible. Yet there was also a quest for another form of visualizing their wearing hours. Interestingly, the test persons came up with a similar distinction of "wearing time" and "wearing duration" as the engineer I referred to above. In addition to the display of the overall duration within a 24 hour period, they wanted to "see" the actual times during which they wore the brace: e.g. from 8 a.m. to 4 p.m. As Aaron, a 17 year old teenager, explained, to understand at what point his counting "went wrong" he needed to have the actual time:

> The way it is designed now, I still don't know when I wore the brace and when I didn't. (...) If I had the exact times I could see, okay, every time I think I just took it off for a couple of minutes or so to do some exercises, but it was actually two hours, I would know when I got it wrong. If I had the exact time I would know the reasons.

What Aaron describes here emerged as a general theme in the interviews with children and teenagers. Recounting their wearing practices, children and teenagers did not offer a total number of hours per day, but related the wearing time with certain activities at concrete times. Typically school and sleep were central routines that they referred to. One teenage girl, for example, who had been advised to wear the brace for only 12 hours a day, preferred to wear it at night and not during school. "I need to put it on in the evening before we have dinner or I watch TV. Or else I don't reach the 12 hours, because I don't sleep 12 hours. But then in the morning I can take it off before I go to school." Others prefer to wear it at school and not at night. "I want to be well-rested for school, so I don't wear it at night. I never got used to sleeping in the brace, so I prefer wearing it in school. When I am sitting, it is rarely a problem."

In describing their daily routines, children and teenagers generally divided their daily rhythm into three blocks: sleep, school, and leisure time. This is a typically modern way of structuring time into labour time and leisure time. Most of them avoided wearing it during any leisure activities, when they "just wanted to chill", or were hanging out with friends. A consistent rhythm helped most of them to gain some sort of routine. But as their days differ from one day to the next, this kind of habituation is also a challenge. As Laura explained in detail:

> For example, when I am on the go the whole day and I know it will bother me, maybe while running to catch the bus or when I go shopping and will be walking around a lot, then I leave it at home and wear it at night instead. But when I know, okay I will be at school and have nothing else planned afterwards, then I wear it to school. And in the summer, when it is really hot, I sometimes do not wear it at all. But besides that, I wear it all the time. Yes, really, all the time.

What those children and teenagers indicate here is a strong link between hours worn, daily activities, and the requirements and conditions these activities bring with them. Moving a lot means you sweat (it is a rigid plastic brace); having to run is hard if you are limited in your mobility by the brace; participating in gym classes or other activities, means you have to find a place where you can lock the brace. Summer is different from winter. And so on. So while, interestingly, children and teenagers do refer to indexes similar to the data captured by the sensor – temperature, acceleration, moisture, pressure – they "process" these indexes in a different way. Their calculation practice is fundamentally embodied as it takes the moving body in the environment into account. While generalizing their hours they move back and forth from concrete contexts and activities to the whole day and the whole week. This reliance of children's and teenagers' calculation practices on concrete context and activities is inherently indexical. Certain weekdays, school schedules, seasons and their temperatures display a complex index for their calculation practices. While they attempt to arrange the wearing of their brace to add up to enough hours each day, the advised time is an ideal, which does not strictly order their day. Whereas the monitoring system sums up discrete time units to a daily number of hours worn and proceeds in a one-many ordering, children and teenagers related wearing-times to the course of a whole day. In Verran's (2013) words, they engaged in whole-parts relating. Their ordered/ordering microworlds are impacted to a large extent by the division of time along school and leisure time and co-constructed by a schooling system that operates on a five-day school week ontology. This became obvious through an interesting discrepancy between the way in which children and their parents calculated hours worn and the medical logic of calculating hours worn. Again and again, I came across the explanation that a lower number of hours worn on a weekday would be compensated on the weekend. Parents legitimized fewer hours on a school day with reference to higher wearing hours on the weekends. When I asked the orthopaedics on the team, they were rather surprised by this widespread misunderstanding and stated that this made absolutely no sense from a therapeutic point of view. Their medical bodies are not subject to the ontological distinction between workdays and weekends. Yet in the microworlds of patients and parents this made perfect sense, like catching up with homework on the weekend. The entirety of a day or even the entirety of a week was the reference for generalizing wearing hours.

Most importantly, this whole-part generalizing allowed for a rethinking and re-evaluation of compliance in many ways. Take the story of Jenny as an example. In an interview she recounted the time she actually had a temperature sensor built into her brace, which recorded her wearing behaviour similar to the monitoring system of the project. While the monitoring system aims to feedback the hours worn in real-time, Jenny's doctor read out the "temperature chip" in the consultation room once a year.

> When we looked at it, I saw that there were actually a couple of days where the hours were extremely low, where I wore it very little or not at all. So I was like: Eh, what happened there? So I investigated a bit and as it turned out, yes, that was when I did a lot of exercise or it was somebody's birthday, so I kind of exercised as we probably went dancing. So at first I thought, oh shit, I didn't wear it enough and when I looked at what I actually did that day and why I didn't wear it, I wrote it down, so that I have an excuse.

While Jenny and others are committed to achieving the advised hours, these numbers became neither symbols nor icons. Simply accumulating kittens might be fun and it might tell them where they stand (numerically) at a certain point in time, but it does not relate back to their daily activities and the concrete contexts of their daily routines. While children and teenagers (and their parents) were engaged in whole-part generalizations, the numbers of hours worn did not quite yet become icons either. The number of hours worn and the category "compliance" were not treated as one and the same, as for example in medical logic. Rather, their calculation practices remained unstable and open to rethinking and redesigning – and therefore remained in the 'indexical zone'. The lack of kittens could actually be reinterpreted as an index of a birthday party. This allowed them to "make excuses": it situated (non-)compliance in everyday routines and their impositions and affordances. The "correct" number was in a sense not the advised duration worn but the number achieved in an adjustment of brace therapy with everyday routines in the course of a day as a lived sleep-school-leisure rhythm and the school week.

This juxtaposition of children's and teenagers' time reckoning[10] to the monitoring system's calculation of time units and one-many generalizing was in itself an analytical time twist: I took the monitoring system as it was (more or less) finalized at the end of the project, confronting it with the on-going therapeutic routines of the patients unfazed by any real-time monitoring. While this gave me the ability to problematize the potential conventionalization of certain calculation practices and the use of numbers as symbols to make truth claims, it presents the monitoring system as a somewhat context-free technological device. It is a device engaged in manifesting numbers as symbols and is an example of the reworking of mundane calculation practices and interventions into problematic behaviour based on computerized processing. However, looking back into the development process, the system's processing of numbers was for the most part at least as messy and indexical as the children's and teenagers' juggling of hours worn during a day or week. Even though it is a rather straightforward example of a combination of sensors, algorithmic processing, and a smartphone application, it took the project team and the engineer responsible for the measurement and development of the algorithms a lot of effort to produce durable and robust symbolic numbers.

Algorithmic inference as incremental calculation

The question I will pursue in the following pertains to how one-many relations become so robust that they gain the capacity to impose their reasoning on certain microworlds. More concretely: How did the monitoring system's generalizing of (non-) compliance become so robust that it could reconfigure what compliance was in the microworlds of young patients struggling to achieve compliance to advised hours? This is not an argument for a deterministic framing of the monitoring system. The routine implementation of the device is still a project of the future, and as I have shown, children and teenagers already have to incorporate various demands into their calculation practices and were mostly able to create arrangements of therapeutic and other obligations that suited their specific needs. Yet, as the monitoring system is a device which is designed specifically to "correct" children's and teenagers' calculation practices, its appearance as robust, objective, along with its ability to provide immediate feedback in real-time has the potential to transform not only how compliance is experienced, but also how it is accounted for by doctors, parents and other parties, such as insurance companies. The aim of this research was to trace the configuration of this "robustness" which was usually framed by the project team as simply the (planned) outcome of technically and mathematically processing physical parameters. This robustness, I argue, is achieved in an incremental process: through a cumulative, shifting process and the gradual manipulation of data, the inferences made in these calculation practices shift from inductive to deductive.

The definite set of the system's indices and the classifications inferred from them take us to the core of what is at stake in algorithmic data-processing. Compared to other examples of algorithmic processing the system at hand might seem rather banal. The data sets are small, hardly Big Data; the algorithms implemented are very basic compared to the complexity of intelligent algorithms. Yet, by attending carefully to this developing process as an incremental process, I intend to show how algorithmic processing as an increasingly common component of one-many generalizations becomes effective in a specific way. As I show, this effectiveness is produced in a complex human and non-human intermingling of data, its clustering and reworking, and technological and mathematical procedures. At the end, these workings seem to become opaque, hardly understandable or questionable.

I came across this opaqueness when I was working with one of the engineers responsible for developing the data processing system. My overall aim, investigating along the development process, was to understand how the imaginaries of the project and the involved stakeholders, and the materialities of sensor, brace, and smartphone etc. would potentially reconfigure what compliance was (Suchman, 2007). I wanted to understand the architecture of the data processing system and found myself venturing into the world of algorithms and machine learning. I particularly remember my excitement in one of our conversations. The engineer had drawn (yet another) sketch of the different steps involved in data processing and machine learning to explain her work to me: producing data in the lab, developing features, training the algorithm with all but one data set, possibly adapting features, training the algorithm once again before testing it with the last data set, and finally evaluating the recognition rate of the algorithm. At some point she tapped on the drawing with her pen and said: "And this is where the direct link between data and decision [yes/1, worn vs. no/0, not worn] disappears. This is not comprehensible to our eyes and our human logic anymore. But with all those coefficients in our equation we can deal with the potential variance." Following this conversation, I spent a lot of time trying to trace and understand this moment where "the link disappears". I envisioned some magical moment of shifting where suddenly the algorithm took over. Eventually I had to accept I had fallen, as many others, for this mystical techno-fantasy of "the algorithm". As it turned out, there is no such magical moment. There are many small steps and there are some important transformations in the processing of data before those 16 kittens appear on the screen. No sudden flip but a continuous cumulative shifting process, where data, features, algorithms are manipulated and adjusted to finally generate the definite decision: yes - no. In striving to make sense of this incremental process of data processing and machine learning, Paul Kockelman's (2017, 2013a, 2013b) systematic focus on transformations of relation (of relations), his repertoire of concepts, and his specific interest in computerized interpretation processes helped me follow these transformations.

Like Verran, Kockelman draws on Peirce's theory of signs to set out his conceptual framework. In his elaboration of the workings of equations in the example of an algorithm for spam filters, he starts with the following semiotic categories and their linkages to elaborate on types of inference: the index, the kind, the individual, the interpretative agent, the ontology (Kockelman, 2013a). As Kockelman (2013a, 2013b) himself states, it is less about the terms but how they are defined; so to reproduce his definitions: "the term index will be used to refer to any quality that is relatively perceivable to some agent" (in the case of the monitoring system, the increase in temperature, sweat, movement among others); "the term kind will be used to refer to any projected propensity to exhibit particular indices" (in the case of the monitoring system, the two kinds are simply 'worn' or 'not worn'); "the term agent will be used to refer to any entity that can perceive such an index and thereby project such a kind"(the engineer in the lab or at the end the monitoring system itself); "the term individual will be used to refer to any entity that can evince indices to an agent and thereby be a site to project kindedness" (the wearer of the brace); "the term ontology will be used to refer to an agent's assumptions as to the indices, kinds, and individuals that constitute a particular world" (the assumption of validity and objectivity of the monitoring system's processing of sensor data and the resulting inference of (non-)compliance) (Kockelman, 2013b: 40-42, 2013a: 151). Using the workings of spam filters as an example, Kockelman (2013a: 45-48) elaborates five forms of equations that are the basis for algorithmic processes (and assumption making more generally) and that produce distinctive kinds of 'ontological transformativity'. Ontological transformativity encompasses both how interpretations (of an agent, based on indices, referring to a kind) mediate ontologies (assumptions concerning an individual and/or kind) and how, vice versa, ontologies mediate interpretations.

To explain the different transformativities and modes of inference involved in the working of spam filters, Kockelman concentrates on the three modes of inference which he refers to as (relatively) deductive, inductive, and abductive.[11] I will again repeat his definition of these as he uses them in various texts: in the relatively *deductive* kind or inferential profile, "indices may change an agent's ontological assumptions regarding the kinds that constitute a particular individual"; in the relatively *inductive* kind or inferential profile, "indices may change an agent's ontological assumptions regarding the indices that constitute a particular kind"; and in the relatively *abductive* kind or inferential profile, "indices may change an agent's ontological assumptions regarding the indices, individuals, kinds, and agents that constitute a particular world" (Kockelman, 2013a: 46-47, 2013b: 151-152). I suggest Kockelman's systematic semiotic vocabulary is productive for dissecting the inner workings of mathematical equations such as algorithms, which have the unique capacity to appear free of context given their technicality and high level of abstraction.

Let me return to the development process of the monitoring system. As I described above, the sensors capture data, which serves as the basis for the classification "worn" or "not worn". With Kockelman's vocabulary we have the sensor-system (the interpretative agent) which produces robust inferences concerning kind (worn – not worn) based on a fixed set of indices: a certain temperature range, rate of acceleration, humidity in the brace. What the algorithm needs to predict is whether – according to the indices – the individual belongs to the kind "wore the brace" or "did not wear the brace". To be able to do so, the algorithm needs to be trained. Just as the children and teenagers have to learn calculating the hours they wore the brace correctly, the monitoring system had to be trained to make the "correct" inferences based on the data produced by the sensors.[12] The engineer started with four types of sensors, which captured acceleration (through one sensor outside the brace), pressure (through one inside), moisture (through a sensor inside and one on the outside of the brace), and temperature (again through one inside and one on the outside of the brace). In her lab she equipped test persons with a provisional measuring system to produce data. In the lab situation, there is an observable link between data (on her screen) and reality (the test person doing motion sequences with and without a brace). The engineer sees that the brace is worn and what kind of data wearing the brace produces, e.g. the rise in temperature once the brace is put on, the change in the temperature difference between the inside and the outside of the brace. For the engineer this is the rather boring part of collecting data. She eventually got used to my fascination with numbers, and we sat at her computer one day to stare at rows and rows and rows of numbers consisting of nine digits. The rows are the output of the laboratory measurements. For each test person she has a folder with a number of spreadsheets, each spreadsheet sheet comprises the data of one sensor. For example, there are two spreadsheets for temperature. While the numbers in the spreadsheet table T1 (the temperature on the outside of the brace) only slightly change after the decimal point (e.g. from 24.6789012 to 24.8765432) the numbers in T2 (measured by the inner sensor) consistently increases (e.g. from 23.4567890 to 36.4637485). These rows of numbers were her starting point:

> I first collected data and then analysed whether the temperature differed significantly between 'worn' and 'not worn'. Then I took the respective average values [of all test persons]: average temperature of brace when worn and average temperature of brace when not worn. Then I calculated the mean of the two values, and divided the difference in two. That is my limit value. This is a method like any other method.

She draws a graph with an x and a y axis and draws a line representing the numbers, showing an increase in temperature. More than 32, it is worn, less than 32, it is not worn. This is the outcome of her observations in the lab: "So first you have a cloud of data. And then you make your first differentiation in the data. Which is based on simple logic. I made the decision simply according to our central question: is it being worn or is it not being worn. That's simple logic."

What she calls simple logic here is an inference starting with the kind or classification "worn"/"not worn" which she observes in her lab. From there she further elaborates which indices constitute that kind. Is 31,7°C an index of worn or not worn? Which is the limit value where a number could be an index of either worn or not worn? This "simple logic" can apply in real time or hindsight explorations. They do not project forward, but induce from what is or has been observed. Based on observation of the test person, brace, and data, the agent (the engineer) creates a range of indices, which potentially constitute the kind. To process the relatively small-scale data-based indices (e.g. 23.4567890 or 36.4637485) into groups, she uses "features" which help her to reduce the multitude of indices.

What the engineer described as a method based on "simple logic" is called "feature engineering" in machine learning. Even though not formally part of machine learning, but rather a prerequisite, feature engineering is often described as the most time consuming and essential part of machine learning (Domingos, 2012; Guyon et al., 2008). As features are domain specific it is difficult to describe them in an abstract way. Basically, the task of features is to "prepare" the data for algorithmic processing. Or to put it in another way, to establish some basic differentiations, which potentially cluster the data into certain groups (of indices). In the project the engineer worked with three relevant types of features: limit values, standard deviations, and one termed signal magnitude area. The latter was relevant for processing data pertaining to pressure and acceleration in order to distinguish between static and dynamic activities. Another feature would be the limit value of the temperature difference between the inside and the outside. For moisture, it is easy insofar as one can assume there is zero humidity at the beginning, so any change points to "worn" (the sensors have in-built heaters, so once the brace is taken off any moisture vanishes quickly). At the end of the feature engineering process there were altogether 14 features. Features are in a sense small-scale generalisers, enabling one to abstract from the multitude of data a cluster of indices. Yet, these features alone do not accumulate units. Rather, they describe what could be part of the unit. And they remain attached to the data and "their" objects, e.g. temperature. But an important initial disentanglement is produced in the process of feature engineering. While the engineer is able to relate the features and the processed data back to the timeline of the spreadsheets and the here-and-now of the lab situation, the features themselves have no direct reference to the time line anymore. The transformation from calculating along a course of time to the cumulation of time as duration begins. But it is not an abrupt disentanglement, for during the training process, data, features, and timeline are constantly connected and reconnected via the engineer.

In this first step of developing the data processing system we can see an inductive mode of inference: observing a phenomenon, defining features of the index. The starting point for inferring in the inductive mode is the observation of a case, where the relation between kind (worn/not worn) and individual is clear. What is potentially transformed is the relation between kind and indices. Once the engineer has a fair number of features, she trains the algorithm with a data set. The whole process is an iterative development, as the engineer explains. When she has a data set, she uses all but one sample to train the algorithm, or rather, various types of algorithms. She does not work with algorithms as mathematical equations, but with "ready-made" software packages in the programming languages of different types of algorithms, provided in an open source library of machine learning, which she can implement, combine, and modify. In a sense she draws from the accumulation of machine learning methods, which themselves are the result of incremental learning in computational science. As Adrian Mackenzie (2017: 22) points out, machine learning itself should be understood as "an accumulation rather than a radical transformation", taking shape "against a background of more than a century of work in mathematics, statistics, computer sciences as well as disparate scientific fields ranging from anthropology to zoology". The part of the development process I depict in this text is but a small sequence in a much longer inferential and indexical chain.

Based on her experience and some literature review, the engineer chooses a few relevant types of algorithms, such as the "*k* nearest neighbour", the "support vector machine" or the classic "decision tree" which seem relevant for the questions she intends to answer. After a phase of training the algorithm, she uses the last sample of data to see if the algorithm comes up with the right solution. This form of machine learning is called "supervised learning algorithm" in software engineering (cf. Mackenzie, 2017: 84-85). Based on the quality of the outcome she goes back to the features and "fiddles around" with them, as she calls it, then generates another data set and so on. Are the features chosen accurate enough, valid enough, the right ones so that the algorithm produces the right assumptions for the kind? Iteratively moving from features/indexes to assumptions/inferences, the validity of the system is ensured, as Kockelman (2017: 128) writes, "in long, tangled, indexical and inferential chains, mediated by machines and algorithms as much as by humans". The central goal is to find the combination of features – algorithm relation (or possibly a combination of several of them), which gets the highest recognition rate of the symbol worn and not-worn.

Learning needs to be completed before the algorithm leaves the lab. The iterative process needs to come to a closure. Kockelman (2017: 25) describes this as a prototypical form of enclosure in computer science (and beyond), involving "processes of objectification, formatting, stabilization, and containment (and sometimes even ways of escape)". In the case at hand, a final selection was made. At the end, one algorithm operating with one feature turned out to be valid enough to produce a recognition rate of 98%. A combination of two of this one set with another set of one algorithm with one feature reached 99%. From this point on, the system was working with a deductive inferential profile, based on the assumptions it was trained to make. The inference is finally disentangled from the observation of "worn" and "not worn" in the lab and the concrete time the actions took place. As the engineer summarizes:

> As long as we develop the algorithm we have a clear mapping with reality and we see what the algorithm spits out. Once we are done with the developing process and can't see the patients anymore, we simply do not know what really happened. We only have our assumptions.

The sensor-system and the algorithms filter out noise for signs in order to make inferences about the kind present. Much of the engineer's work is to train the algorithms with data, features, and then to compare the outcome to the phenomenon observed "in reality". Along the way, the algorithms learn to make increasingly valid assumptions; if they get something wrong, the features are reworked or another algorithm is chosen. The aim is to implement a combination of algorithm(s) and feature(s) that fits the data and the kind in question. While there is much debate on the opaqueness of algorithms and their relative autonomy in decision-making, I like to emphasise that this autonomy is fundamentally distributed: the features and data are assumptions that are very much based in the materiality of the sensor system as well as in the logic (and beliefs and desires) of its developers. Nevertheless, at the point of closure, the inferential profile has shifted from inductive to deductive. While for most of the developing process, the distinction between "reality" and "algorithmic reality" is crucial, once the learning process ends, they are conflated. Or, because they merge, learning stops. The numbers produced with the monitoring system have moved from the indexical zone to become conventions, or symbols as Verran calls them, now generating one-many relations, bringing forth kittens on the screen of a teenager's smartphone. One could claim that this is actually the moment where the algorithm "takes over". But as I have hopefully made clear, this gradual shifting towards deductive inference is distributed in a specific way and is the (preliminary) result of a long "inferential and indexical chain" (Kockelman, 2017: 128). This chain started long before the moment in the lab when I switched on the ethnographic light. The question remains as to how patients will interpret the monitoring system and the objective numbers and alter their assumptions on compliance.

Conclusion

This article addressed the sensor monitoring system and young patients' calculation practices, sieving through the empirical and building an argument using concepts developed by Helen Verran (e.g., 2010, 2013) and Paul Kockelman (e.g., 2013a, 2017) . Switching between Verran's careful attention to 'ontic / ontological tensions' to Kockelman's sieve of 'ontological transformativities', I was moving from the system as it was developed by the end to the microworlds of the patients and back to a retrospective dismantling of the step-by-step process of the system's learning. Throughout the analytical and especially the writing process, I had to actively construe those sieves which did not really fit at the out-set. Kockelman's sieve is fine-grained yet isn't able to capture the tension generated by the different "versions" of numbers' workings as one-many relations and whole-part relations. Or is there something like a double-sieve? At the same time his sieves are especially well-attuned to the dissecting of computer-based techniques of making interpretations. Moving along the incremental process of developing a robust, valid, seemingly autonomous calculation device enabled me to focus on the reworkings of numbers in a constant recombination of data, features, and algorithms. While much debate focuses on the deductive, reductionist and at the same time seemingly opaque workings of algorithmic processing, I mainly focussed on the moment before the deductive mode of inference was implemented. The management of complexity is performed in disentangling data from its empirical grounding to slowly transform numbers that perform indexicality into numbers that function as symbols. Becoming a deductive device is a process involving sensors and captured data (in the lab), its iterative manipulation based on "human logic", training and testing algorithms to make valid inferences. To understand what is at stake in algorithmic processing, this is the process at which we need to take a closer look. What are the assumptions that become embodied in the algorithmic inference and how might this alter ontological assumptions about compliance? I proposed introducing another performative property to Verran's repertoire: numbers' inferential profile and their capacity to make durable and unequivocal assumptions about the world and to interfere in the world.

According to the logic of the technology developers, the production of correct calculations can only be achieved through disentangling complexity and reducing potential nuisances on the way; juxtaposing the calculation practices of the system with the complexity management of the young patients brought into focus a very different form of complexity management. Recall the story of Jenny who was checking the outcome of the temperature sensor system against her actual activities at concrete moments. She interpreted the non-wearing of the brace during certain activities at certain events (dancing at a birthday party) not as a sign of non-compliance; rather, recounting the event and the activities served as an index for her inference: there is room for excuses, for a re-evaluation of what counts. It is a refusal to conventionalize calculating time according to a metric device (for a further discussion on refusal and resistance to numbering practices in evaluating schools' performances see Gorur, 2018).

Yet it is not a refusal or resistance to the device and the monitoring and visualization of hours worn per se. Children and teenagers embraced the idea of the monitoring system, they enjoyed accumulating kittens; most importantly, it enabled them to quantify and visualize what is otherwise hard to grasp. Having to count the hours one wears a brace (with many forgetting at some point that they are actually wearing it) is a challenging learning process. The question is how to integrate the monitoring system's symbolic numbers and one-many generalizations with the young patients' calculation practices and their microworlds. Obviously, an important empirical part is still missing here: the routine use of the system. What kinds of effects might emerge with the use of the system?[13] Verran would offer the right tools for carefully dissecting how the different modes of generalizing wearing-times encounter each other in everyday routines, how they possibly create frictions, merge and/or subordinate each other. I could speculate on what might happen if the monitoring system gets implemented, but cannot make an empirically sound claim. However, I insist on the potential of the chosen analytical tools not only in hindsight, when we can actually observe the workings of such systems in people's lives, but also to problematize the potential implications for the future. Juxtaposing the calculations of time worn based on my account of children's and teenagers' mundane calculative efforts with the development of the monitoring system, backed up by the feedback of the participants of the observational study, enabled me, for example, to propose an additional visualization of wearing-time to the development project. While children and teenagers embraced the kitten-version of the feedback, providing them with a time line similar to a school timetable would assist them in their struggles to achieve the advised hours and multiply the indexical range. Yet this small pragmatic supplement leaves a central problem untouched.

The incremental processing of deductive inference through the monitoring system will potentially have an amplifying effect as it reinforces the logic of psychological assumptions about rational choice and decision-making, medical assumptions about numerical evidence and evidence-based interference, and techno-scientific assumptions about the neutrality of mathematical and technological data processing. As I have shown, producing deductive inference is not a straightforward process, but messy work distributed between human and non-humans. The potential decontextualization of compliance and the reduction of compliance to absolute numbers are not produced "by the algorithm". Rather we have to pay attention to the configuration of numbers as symbols, with deductive inferential profiles, working as one-many relations, to potentially reinforce each other and make the case for only one logical way available to treat the issue at hand. As these numbers are created not only for inferring conditions from physical parameters within expert systems but with the intention to conventionalize mundane numbering practices such as those of the young patients, we need to carefully attend to the work done in the production of such unambiguous numbers and to their capacity to transform ontologies.

References

Asdal K (2008) Enacting Things Through Numbers: Taking nature into account/ing. *Geoforum* 39: 123-132.

Ballestero A (2015) The Ethics of a Formula: Calculating a financial–humanitarian price for water. *American Ethnologist* 42: 262-278.

Bauer S (2008) Mining Data, Gathering Variables and Recombining Information: the flexible architecture of epidemiological studies. *Studies in History and Philosophy of Science Part C: Studies in History and Philosophy of Biological and Biomedical Sciences* 39: 415-428.

Callon M and Muniesa F (2005) Peripheral Vision. *Organization Studies* 26: 1229-1250.

Cochoy F (2008) Calculation, Qualculation, Calqulation: shopping cart arithmetic, equipped cognition and the clustered consumer. *Marketing theory* 8: 15-44.

Domingos P (2012) A Few Useful Things to Know About Machine Learning. *Communications of the ACM* 55: 78-87.

Dourish P (2016) Algorithms and Their Others: Algorithmic culture in context. *Big Data & Society* 3: 2053951716665128.

Gillespie T (2014) 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.

Gorur R (2018) Escaping Numbers? Intimate Accounting, Informed Publics and the Uncertain Assemblages of Authority and Non-Authority. *Science & Technology Studies* 31(4): 89-108.

Guyon I, Gunn S, Nikravesh M, et al. (2008) *Feature Extraction: Foundations and Applications*: Springer Berlin Heidelberg.

Hacking I (1990) *The Taming of Chance*: Cambridge University Press.

Kockelman P (2011) Biosemiosis, Technocognition, and Sociogenesis. Selection and significance in a multiverse of sieving and serendipity. *Current Anthropology* 52: 000-000.

Kockelman P (2013a) The Anthropology of an Equation. Sieves, spam filters, agentive algorithms, and ontologies of transformation. *HAU: Journal of Ethnographic Theory* 3: 33-61.

Kockelman P (2013b) Huckleberry Finn takes the Turing test: The transformation of ontologies and the virtuality of kinds. *Language & Communication* 33: 150-154.

Kockelman P (2013c) Information is the Enclosure of Meaning: Cybernetics, semiotics, and alternative theories of information. *Language & Communication* 33: 115-127.

Kockelman P (2017) *The Art of Interpretation in the Age of Computation*: Oxford University Press.

Kockelman P and Bernstein A. (2012) Semiotic Technologies, Temporal Reckoning, and the Portability of Meaning. Or: Modern modes of temporality–just how abstract are they? *Anthropological Theory* 12: 320-348.

Lalor BJ (1997) The Classification of Peirce's Interpretants. *Semiotica* 114: 31-40.

Lave J (1988) *Cognition in Practice: Mind, Mathematics and Culture in Everyday Life*: Cambridge University Press.

Lippert I (2018) On Not Muddling Lunches and Flights: Narrating a Number, Qualculation, and Ontologising Troubles. *Science & Technology Studies* 31(4): 52-74.

Mackenzie A (2014) Multiplying Numbers Differently: An epidemiology of contagious convolution. *Distinktion: Scandinavian Journal of Social Theory* 15: 189-207.

MacKenzie A (2017) Machine Learners. Archaeology of a data practice. MIT Press, Cambridge, Mass.

Maurer B (2013) Transacting Ontologies. Kockelman's sieves and a Bayesian anthropology. *HAU: Journal of Ethnographic Theory* 3: 63-75.

Neyland, D (2018) Something and Nothing. On Algorithmic Deletion, Accountability and Value. *Science & Technology Studies* 31(4): 13-29.

Peirce CS (1974 [1906]) *Collected Papers of Charles Sanders Peirce,* Cambridge, Mass.: Harvard University Press.

Porter TM (1996) *Trust in Numbers. The pursuit of objectivity in science and public life*: Princeton University Press.

Suchman L (2007) *Human-Machine Reconfigurations. Plans and situated actions*: Cambridge University Press.

Verran H (2001) *Science and an African Logic*: University of Chicago Press.

Verran H (2010) Number as an Inventive Frontier in Knowing and Working Australia's Water Resources. *Anthropological Theory* 10: 171-178.

Verran H (2011) The Changing Lives of Measures and Values. From centre stage in the fading 'disciplinary' society to pervasive background instrument in the emergent 'control' society. *The Sociological Review* 59: 60-72.

Verran H (2012) Numbers. In: Lury C and Wakeford N (eds) *Inventive Methods: The Happening of the Social.* London, New York: Routledge.

Verran H (2013) Numbers Performing Nature in Quantitative Valuing. *Nature Cultures* 2: 23-37.

Verran H (2014) Anthropology as Ontology is Comparison as Ontology. *Theorizing the Contemporary, Cultural Anthropology website, January 13, 2014. https://culanth.org/fieldsights/468-anthropology-as-ontology-is-comparison-as-ontology.*

Zaloom C (2003) Ambiguous Numbers. Trading technologies and interpretation in financial markets. *American Ethnologist* 30: 258-272.

Ziewitz M (2015) Governing Algorithms. *Science, Technology, & Human Values* 41: 3-16.

NOTES

1 For a further discussion of these concepts see Verran (2012: 66). For a further classification of the interpretants (or 'significate effects') see also Lalor (1997).

2 It would be a misinterpretation of Verran's work to suggest she does not incorporate this potential gradual shifting and transforming of numbers. She points to this rather briefly in *Science and an African Logic*, where she describes numbers' relational capacity to "seamlessly connect" and "recursively juxtapose[s]" between various entities such as "a child sleeping on his mother's back in Ibadan with the ledger of the British Empire" (Verran, 2001: 100). In an example of the computerized processing of numbers into a hybrid of symbolic/iconic manifestation, she hints at the implications of model-based automated numbering processes (Verran, 2013: 31). Yet the overall focus in her work lies elsewhere, as discussed above. Overall, to elaborate a recombinant of the two analytic tools hinges on their principle potential connectivity to produce any combinatory benefit.

3 Note for example his attempt to synthesise in one article (Kockelman, 2011) what he calls biosemiosis, technocogniton, and sociogenesis, with examples as different as animal-signals systems and natural selection to lawn mowers and Turing machines.

4 In recent years there has been increased interest among scholars of STS on studying algorithms; for a critical discussion of what can actually be considered an algorithm and what it means to take algorithms as objects of analytical attention see for example Dourish (2016), Gillespie (2014), and Ziewitz (2015).

5 For a discussion (or mediation) of Kockelman's work as "transacting ontologies", also in contrast to other takes on ontology in anthropology, see Bill Maurer's review in HAU (2013).

6 The regional innovation cluster "BeMobil: Regain Mobility and Motivity" was funded by the German Ministry of Education and Research. The cluster focused on the development and improvement of rehabilitation technologies and therapeutic systems for patients with limited mobility after a stroke or due to amputation or scoliosis. For more information see http://www.ige.tu-berlin.de/bemobil/parameter/en/

7 In the lab situation, the system actually captured data more closely; e.g. temperature and humidity every five seconds. However, during the observational study with teenagers using the system "in the wild", data was captured every five minutes. The latter seems to be a more likely final solution, mainly due to storage capacity and energy supply. The underlying logic however – summing up intervals versus passage of time – remains the same with five seconds or five minutes.

8 Compared to the complex examples and hybrid numbers Verran is elaborating on in her numerous examples, this is a rather simplified elaboration of her concepts and arguments.

9 In a sense the intervention is rather symbolic, as there is no effect other than the numbers appearing on the display. In principle, the monitoring system addresses a self-reflexive subject, one who changes her or his behaviour based on the numbers. Yet in everyday life the brace-wearer and the monitoring system are not isolated from social worlds, where parents, but also therapists and – even though only once or twice a year – doctors, comment on these objective, technically produced numbers and participate in this new regiment of compliance. In what way the system will actually intervene cannot be answered at this point.

10 Cf. to Paul Kockelman and Anya Bernstein's (2012) work on time reckoning, with a systematic description of the portability of measuring systems.

11 He leaves aside the most common transformativities usually addressed in social sciences and anthropology: the speech act and the looping process (cf. Kockelman, 2013: 45-49).

12 But which sensors and what data? Actually, the answer to this question was part of the development process: Which (combination of) sensors will produce data that leads to the most robust mode of inference? And what is the relevant inference? And also more technical questions: How could they be attached to the brace? What about storage? Yet, even though this "relatively abductive phase" in the development process was an important prerequisite for the further development of the system and shaped machine learning in a fundamental way, for reasons of comprehension and space, I decided to leave this part out and start in the lab.

13 One could also draw directly from Peirce's concepts and differentiate between different interpretants generated by the use of the system. Peirce (1974 [1906]: 326-327) elaborated different interpretants: from the "emotional interpretant" evincing an emotional response (remorse, frustration or satisfaction maybe), the "energetic interpretant" to a "habituated response" (e.g. the number on the display triggers a certain change in performance, such as putting the brace on when the number was not yet high enough). Peirce's work offers a variety of ways of differentiating the potential effects of the system and I thank the anonymous reviewer for making me aware of this rich potential.