Transforming Data: An Ethnography of Scientific Data from the Brazilian Amazon

Antonia Walford
IT University of Copenhagen

Thesis submitted to the IT University of Copenhagen in compliance with the requirements for the degree of Doctor of Philosophy (PhD)

Abstract

This thesis is an ethnography of scientific data produced by a Brazil-led scientific project in the Brazilian Amazon. It describes how the researchers and technicians make data about the Amazon forest, and how this data in turn generates different scientific communities, scientific subjectivities, and claims about the world. It explores the limits of a representational idiom to describe such scientific practice, and in so doing investigates the reflexive and recursive repercussions of such descriptions for the anthropology of science.

Key words: Anthropology of Science; Brazilian Amazon; STS; data; climate

Acknowledgements

This thesis would not exist without the time and effort that the members of the LBA spent with me, answering my endless questions with the utmost patience. My special thanks go to E, David, Mark, R, J, all the Micro group, and all those at ZF2 with whom I spent the most time (who remain anonymous in the text). However, I am extremely grateful to everyone at the LBA for opening their doors to me. I would like to thank them all for sharing their work and thoughts: the Modeling groups, the Biogeochemistry group, the CliAmb students, the DIS team, the Logistics team, the technicians, the executive manager, those involved in PPBio, those at CPTEC - and everyone else, either permanent fixture or just passing through. This thesis stands as testament to the creative capacity of their skills and knowledge. Any errors this thesis contains are entirely my own.

I would also like to thank my supervisor, Casper Bruun Jensen, for being such a steadfast and calming presence, providing invaluable balance at crucial times in the argument and the process of producing it. I am extremely appreciative of his constant support and encouragement. My co-supervisor, Morten Axel Pedersen, also provided a great deal of intellectual stimulation and guidance, for which I am also very grateful. I was extremely lucky to have such dedicated and inspiring supervisors.

I would like to thank those in the TIP group at the ITU for providing such an open, friendly and fertile environment: Randi Markussen, Brit Ross Winthereik, Christopher Gad, Nina Boulus, Laura Watts and Pernille Bjørn. And to my fellow PhD students, at the ITU and beyond - Peter Lutz, Lea Schick, Naja Holton Møller, Birgitte Gorm Hansen, Helene Ratner - thank you for the spirited conversations and shared traumas. I also count myself lucky to have been able to hang out with PhD students from the University of Copenhagen and the University of Aarhus. I thank them for the formative conversations this permitted.

Although it is some time ago now, I also want to thank Eduardo Viveiros de Castro for the continuing impact he has on my work, as this thesis is born partly out of the time I spent under his supervision from 2004-2008. My time at the Museu Nacional, UFRJ, Rio de Janeiro, was hugely inspirational, and I thank everyone there who made it such an exciting place to be. I carry that time and those people with me.

Catherine Fitzpatrick read the whole thesis and whipped it into shape in record time, for which I shall be eternally grateful. I would also like to thank Kirsten Hastrup and Brit Ross Winthereik for insightfully critiquing an earlier draft. Julia Sauma, Alice Elliot, and Chloe Nahum-Claudel all read and commented generously on chapters. Martin Holbraad contributed to making the argument what it is, and Matei Candea provided essential advice.

I am lucky to have a very understanding family, and I would like to thank my parents, my brother and my sister, for their support throughout this project. My partner, Flora Berkeley, has been notably stalwart, and writing this would have been impossible without her constant love and insight.

This thesis was written with the financial support of the IT University of Copenhagen. I received a further grant from the Danish Research School for Information Systems. I would like to thank Jane Andersen in the ITU's Research and Learning Support, and Annette Jørgensen in the ITU's Økonomi og Personale, for dealing with all my plaintive requests so gracefully.

I dedicate this thesis to the memory of my grandfather, Jack Benzimra, one of the people who enabled me to carry myself forward into knowledge.

Contents

Introduction

"Yesterday's scientists studied nature.

Today's scientists study digital data."[1]

Why Scientific Data?

This thesis is based on thirteen months of fieldwork carried out in 2007, and in 2010-2011, with the Large-Scale Biosphere Atmosphere Experiment in Amazônia (LBA). In 2011, I also spent one month at the Centre for Weather Forecasts and Climate Studies (CPTEC), which is part of Brazil's Space Research Institute (INPE) in São Paulo state.[2] The LBA is one of the largest scientific experiments in the world (in terms of both personnel and duration) to focus on environmental science. A Brazil-led scientific project, it was initiated in response to the United Nations Framework Convention on Climate Change (*Convenção-Quadro sobre Mudanças Climáticas*) that was established at the 1992 United Nations Conference on Environment and Development held in Rio de Janeiro. This convention demanded that each nation determine its contribution to the global greenhouse gas budget, and the impact of its emissions, natural and industrial, on the global atmosphere.[3] After several smaller projects were conducted during the 1980s and 1990s, involving collaborations between different countries and Brazil, the LBA emerged formally in 1998 as an international collaborative scientific endeavour to discover the role of the Amazon

---

[1] Schröder, P. (2003) Digital Research Data as the Floating Capital of the Global Science System, 7-12. In Paul Wouters and Peter Schröder (eds.) *Promise and Practice in Data Sharing*. Available at: http://www.google.co.uk/url?sa=t&rct=j&q=&esrc=s&source=web&cd=1&ved=0CGMQFjAA&url=http%3A%2F%2Fdataaccess.ucsd.edu%2FPromiseandPracticeDEF.pdf&ei=4RM1UICxPOTS0QWgxoHoBQ&usg=AFQjCNG8fvUJrrc5M7BnouSe7LT7wGAC0g&sig2=I-rx2Uk5KZfqLU7kwikH8w

[2] Unfortunately, I was unable to include most of the CPTEC material in the thesis due to time and space constraints.

[3] http://150.163.158.28/lba/site/port/conciso/cpp04.htm. Accessed August 2012.

forest in the global carbon cycle and its relation to the “Earth System”. The LBA experiment has since then produced an enormous amount of meteorological, chemical and biological data, and many articles and publications.

After conducting nine weeks of fieldwork with the LBA for my master’s degree in 2007, I resolved to return for my PhD fieldwork in 2010, but this time with the specific intention of studying the LBA data. On several occasions I was told by researchers in both the modeling and observational sciences that there is large gap in many fields of scientific knowledge, including meteorology, due to the lack of observational data collected in tropical biomes. Temperate areas are well-covered with data collecting stations, but very few have been installed in the tropics. My informants pointed to three main reasons for this. Firstly, a large part of the tropics is made up of ocean, which is notoriously difficult to collect data on (and in). Secondly, where there is a tropical land mass, it is normally heavily forested, which also makes sustained data collection difficult. Thirdly, the countries in tropical zones are less financially stable than the countries in temperate ones, and so either do not have the money to invest in such scientific infrastructures such as data collection networks, or invest more heavily in other areas.

One of the broad reasons that world-wide coverage is sought after is in order to provide regularly-spaced observational data inputs for climate and weather models. These models work with Cartesian grid space and ideally need an "input value" for every point in that space; where there is no observed data at all, this leaves the model with only interpolated data as input (see also Edwards 2010: 251-286). Infrastructures that are available for data collection in the tropics, such as the LBA, are also highly-prized by those in the relevant scientific communities because they provide data on the largest tropical forests on the planet, and allow comparative studies to be conducted. There is accordingly a strong emphasis within the LBA, particularly in public presentations of the project, on the meteorological towers that have been erected in the Amazon forest to allow for the collection of profiles of carbon and energy flux data that are otherwise impossible to obtain.

Over the course of the fieldwork I had previously conducted, it had become clear to me that this data was not only considered to be crucial by the scientific community at large, but was also what most concerned the LBA's researchers and technicians on a day-to-day basis. In 2007, during the two or so months I spent with the LBA researchers, students and technicians, I had noticed that a most of their time was dedicated in one way or another to collecting and processing this data, using it, or ensuring it was available for further use by others. The LBA, as a distributed locale, was also used by many different foreign researchers (*estrangeiros* or *pesquisadores de fora*). It provided them with safe and legal access to the forest and with infrastructure in the form of lodging, internet access, trails into the forest, vehicles to get in and out of the forest, personnel, and equipment, including the meteorological towers. The researchers and technicians seemed to want to go into the forest solely in order to collect data and bring it out again. Nor was data only a quotidian concern: as the epigraph suggests, when these researchers and technicians showed me their work, they showed me their data. Obviously not natural, but also not exactly fabricated, data piqued my interest by (I suspected) conceptually inhabiting an interstitial space between categories that are normally to hand when social scientists try to describe scientific practice, such as "nature" and "culture".

This thesis thus explores the collection, processing and subsequent circulation of the scientific data[4] produced about the Brazilian Amazon by the students, researchers and technicians of the LBA. The LBA could be described as only a small part of a much larger, "global" network of infrastructures dedicated to ecological and meteorological data collection and dissemination (see Miller and Edwards 2001; Kwa 2005; Edwards 2010). But this small part, it emerges, has some striking peculiarities. Some of these are the results of its specific setting, others only became apparent because of the focused approach I adopted. Firstly, collecting data in the Brazilian Amazon is not the

[4] The word "data" in English can be referred to either in the plural or the singular simply by changing the subsequent verb: "the data are" or "the data is". In Portuguese, there are two separate words for the singular and the plural, *o dado* (sing.) and *os dados* (pl.). The LBA researchers almost invariably referred to *os dados,* because *o dado* might be better translated as "the datum", stressing the singularity of a particular data point or set. However, as I wanted to emphasize the LBA data in its entirety and in its singularity as the focus of my thesis, I opted to refer to it in the singular as "the data is", (but in contradistinction to "the datum is"), to mean the entire body of data that is associated with the LBA, as well as the particular data that was at any given point in time the subject of my (and my informants') scrutiny. Sometimes, it appears in the plural if translating a citation. Here I do not mean anything more by my choice, although the flexibility between singular and plural usage evidenced by the concept of data is certainly noteworthy.

same as collecting data in a temperate, easily-accessible and urbanized area of the world. The Amazon forest exerts a particularizing force on these activities in several different directions, as I explore more fully in the thesis. Secondly, instead of concentrating on the infrastructure or organization of the LBA, or the uses to which the data is put, I concentrated on the data itself. As such, this thesis is ethnographically aligned with work in Science and Technology Studies (STS) that examines databases and data as ethnographic objects in their own right, with the capacity to extend beyond being simply 'neutral information' – material, practical, ubiquitous and political entities that shape the contexts in which they are embedded (Bowker 2008; Hine 2006; Hilgartner and Brandt-Rauf 1994).[5] However, aside from recent studies of the impact that "big data" has had on the biological sciences (e.g. Leonelli 2012), there has been relatively little attention paid to the particular techniques and means by which data emerges as ‘data’. And this is, precisely, my interest: I want to examine what data might be in its own right, as it were. I have therefore begun not by looking at the data as one part of a much larger network - which would require one to start with the larger network and work backwards - but as an entity unto itself. Here, I take inspiration from a similar move made in recent ethnographies of documents as "artifacts of modern knowledge" (Riles 2008).

My fieldwork therefore focused, not on the scientists, technicians and researchers of the LBA, nor the Brazilian Amazon itself or any particular location within it, but rather on the data that the scientists, technicians, students and others produced about the Brazilian Amazon. As it was impossible to study the one without studying the others, the thesis is in one sense a description of the relations between these three. Nevertheless, its structure reflects the ways in which I took the data as a limit, and to a certain degree as the determinant of where I went and what I investigated. Taking data seriously in this fashion required me to trace its very first collection out in the forest, and the subsequent transformations that occurred as it moved out into the world from the forest. The data I investigated often did not reach the wider network of infrastructures that has come to characterize studies of climate science, or the institutions such as CPTEC that run climate models, and this thesis does not really

[5] This scholarship, rather than analysing scientific "inscriptions" (Latour and Woolgar 1979: 51; Lenoir 1998), concentrates specifically on scientific digital data. The earlier work focusing on inscriptions is however also an inspiration for this thesis.

deal with them. Instead, it asks what data is when it is extracted from the world, how this extraction occurs, what that implies about the world, and how data subsequently comes to signify and relate to other data and people as an entity. Simultaneously, it engages with the self-reflexive work of asking what relevance this has for anthropological enquiry.

Findings and Argument

As I proceeded with my fieldwork, three principle observations emerged that structure this thesis, and can be seen as its major findings. The first seems at first glance to be hardly a result at all. In confirmation of what I had observed during the previous fieldwork stint in 2007, I discovered that those researchers and technicians I was working with invested the world they were studying, namely the Amazon forest, with an absolute degree of reality. This was clear in the way in which they spoke to me about it, and conducted their work. There were no ways in which they indicated that their investigations were not discovering something already given. The importance of this was not clear to me until, near the end of my fieldwork, I was asked to give a presentation of my findings to the researchers and students. Nervously, I decided to conduct my own experiment, and settled on giving a lecture on the intricacies of Actor-Network Theory, stressing the points which I thought would cause the most alarm to a group of scientific realists - emphasizing, for example, how "the length of associations, and the stability of the connections through various substitutions and shifts in points of view make for a great deal of what we mean by *existence* and reality" (Latour 2005: 164). I used the work of several people present in the room as "case-studies". The lecture room was packed, as those I had spent the last year questioning and prodding were finally given the opportunity to do the same to me.

As I finished explaining the last slide, several hands shot up. Ready to deal with the table-banging that I had been primed for (principally by the texts I had been explaining), I took the first question. To my great interest, however, almost no-one took any issue with most of the contentious ideas I had taken pains to explain clearly and simply. They, in fact, asked much harder, side-on questions. One person asked why I had not included *his* project in my talk, was it not interesting enough? Another asked perplexedly what my master's viva had been like: if there was no such thing as right and wrong, how had I passed? Another asked if I had had any issues of cultural miscommunication, as she remembered happening when she had to conduct research with riverine communities. Another asked me what I had meant by suggesting there might be more than one way to do science, did I think that they did not do proper science in Brazil? And finally, a professor asked me if I had read any Isabelle Stengers (I had, and I use her work a great deal).

They did not misunderstand my lecture, (I had been very emphatic) but they simply did not find it interesting to talk about reality in the way that I had presented it to them. They raised questions that interested them, for instance asking how I had selected my data, or how others might evaluate my findings. Those questions, in turn, had unexpected and generative effects on my own enquiries. I realized, for example, that even though a researcher might tell me that "doing science in the Amazon is different to doing it anywhere else", if that were repeated back to them by someone *from* somewhere else, it became different information. This was all very stimulating, but it dawned on me that as far as those researchers and students are concerned, it is not interesting to ask whether the world is real or not, simply because the world just *is*.

The second observed finding of this thesis is that concomitantly, what interests researchers in those terms is not so much nature, but their data. This data inhabits an area of knowledge that is peculiarly its own. I have already suggested that as such it slips between opposed categories such as "real" and "fabricated". It has a life of its own, and as I followed its course from collection to processing to storage and dissemination, I became aware that it also has a very particular aesthetic. I use the term "aesthetic" to refer to "the persuasiveness of form, the elicitation of a sense of appropriateness" (Strathern 1991: 10; Riles 2001: 185). The data becomes data – that is, it becomes trustworthy and persuasive – by means of techniques and processes that elicit from it a certain form. I discovered that one of the important features of this form is that it is self-referential, because it carries with it its own referent; or conversely, it is real when it refers correctly to itself. Importantly however, although I imagine this conclusion would be mildly interesting to those researchers I worked with, I suspect that they might want to move on as quickly as possible to explain the *content* of *their* data to me. In so far as I also observed that the data needed to be given a specific form in order for its content to be understood and subsequently used, another important realization was that the data aesthetic is one in which oppositions, such as form and content, can elide and separate.

The third broad observation on which this thesis rests is that this form also has external relational capacities. It must not only refer or relate inwards to itself, but also outwards, to other people and other data sets. I was often told that data would only be shared when it was ready, or when it had been processed. An enormous amount of work goes into "cutting and polishing" (*lapidar*) the data in this way, in order to produce certified data that can then be shared, stored, compared and published. In investigating this, what was revealed was a mutual relation of co-production between the data and those who work on it. Just as the data cleaners, for example, work on the data, so too does the data work on the data cleaners. The final form of the data is not an end point. Data is important, in fact, because it can subsequently be transformed into something else – a publication, or another claim about the world. Thus, the relational capacities of the data, I discovered, often crystallized around this transformative potential, which itself was not comprehended in the same way by all those involved in its realization. Since relations can also separate (Strathern 1988), the social entanglements that occured due to this explained a great deal of the friction that arose in situations of international collaboration. I therefore saw a way to trace the contours of these frictions by viewing the data as relational entity. I also discovered the limits of my own study, in the limits of the data's relational form: not everyone who was involved in the data's life had any interest in it whatsoever.

This thesis is an exploration of three broad ideas provided by the ethnographic engagements I had in the field. The first is that nature is not real, because it cannot be false. The second is that the data can be real, because it comes to refer to itself. And the third is that the data has relational capacities that flow into and out of itself in different ways and have effects that are social. As well as following and drawing out the intricacies of these three findings, I also always intended this thesis to contribute to the methodological and conceptual discussions around the question of what it is to study science anthropologically after the post-modern (critical) turn. One of the ways in which I attempt to do this is to try to re-think the representational relation (which substantially underpins the dispute between so-called "realists" and "constructivists") in the terms of the data aesthetic that emerged from my research. I identify a series of elisions and separations as a governing dynamic that allows for a multitude of different oppositions (such as real/representation, data/error, skill/knowledge, form/content) to overtake each other. In this way, I hope to move away from the constrictions that the representational idiom imposes on anthropological descriptions of science, without losing what is essential to those descriptions - the possibility of countenancing scientific reality, for example.

An Ethnography of Scientific Data - Defining the Field

I have chosen to refer to this thesis as an ethnography of scientific data. What constitutes ethnography as anthropological methodology is continually being re-assessed within the discipline (for a recent example see Riles 2008: 1-40). Here, I am working loosely with the idea that ethnographic endeavours are ones of "immersement", in which "what must be taken into account is what has been overlooked" (Strathern 1999: 5). Ethnographic fieldwork in this rendering necessitates the suspension of the criterion of "relevance". Respecting knowledge's inherent expansiveness, it proceeds by gathering anything as potential information. What is relevant *becomes* so, continuously, as the anthropologist tacks between data and analysis, or the field and the desk, endlessly re-discovering both. My ethnographic methodology is, therefore, characterized not only by unpredictability, but also by awareness of the generative potential of what there still is to know; the partiality that resides in acknowledging the importance of what Bill Maurer calls "missing terms from the future that may arrive, or may not" (2005: 13).

This is not to say, as George Marcus and James Clifford (amongst others) did in the 1980s, that such an acknowledgement is necessarily intended to highlight the "artificial, constructed nature of cultural accounts" (Clifford 1986: 2). The field is only a construction in so far as it is a relation, as Eduardo Viveiros de Castro reminds us (2002: 113). And as Viveiros de Castro also reminds us, it is the anthropologist's task to discover what "a relation" might mean for those who he or she works with (ibid: 122). The field can therefore be engaged in as a relation between the anthropologist and their informants and fieldsite that is *in potentia*, constantly in the making. Again however, although this is a crucial step, it is thus not enough to acknowledge that the field is co-constructed through the act of fieldwork without paying attention to what differences this entails in each ethnographic case.[6]

George Marcus (1995) introduced the term "multi-sited fieldwork" in order to take account of the fluid and complex field-sites that anthropologists were increasingly finding themselves engaged in - focused not on one isolated locus, but on many that seemed to inter- and dis- connect on multiple scales. Certainly, in one sense, my field was constantly shifting. I accompanied the data as it travelled from the collecting platforms deep in the forest into the LBA central office in Manaus, and from there to multiple locations, virtual and actual, including the Brazilian Space Administration in the southeast of Brazil. I also accompanied trips to different research sites all over the Amazon, and attended conferences. However, discussions of "multi-sitedness" have come some way since the 1990s. In a recent re-appraisal, Matei Candea interrogates the ideas of limitlessness, complexity and expansion inherent in the idea of multi-sited fieldwork, and instead emphasizes the importance of an "arbitrary location" (Candea 2007, 2010) as a means of making the "cut" necessary to define one's field. He presents this as a development of the dissatisfaction with totalizing discourses that Marcus and others write of in the 1980s, but also as a move away from the conclusions they reached. Candea makes the point that "when it presents (un)boundedness as a real feature of the world out there…rather than a methodological issue, the multi-sited approach forgets the possibility, indeed the necessity of *bounding as an anthropological practice*" (2007: 172, italics in original). By highlighting limitation as a decision on the part of the anthropologist, an arbitrary location is "premised on the realization that any local context is always intrinsically multi-sited" (ibid: 175). At the same time, it also refuses to take that multi-sitedness as a totalizing discourse itself. It was through acknowledging the productive and specific partiality of the choice itself that I chose the data as my field-site, a (not-quite) arbitrary location.

[6] Here I follow Marilyn Strathern: "Tyler and Marcus remind us that every inversion we deploy is self-referential...; but the deployment of *particular,* concrete inversions is not...Any such contextualisation can of course be recaptured as in turn self-referential, in the same way as "other" can always be collapsed as a version of "self". But to regard this last position as a final one is to *hide the movement* through which it was reached" (Strathern 1987: 279).

Taking Science Seriously

The productivity of partiality is a powerful trope in discussions of ethnographic methodology. It is precisely because of this emphasis on the generative potential of partiality that sighting glimpses of hidden or unremarked totalities is an effective way for anthropology to destabilize its own dominant conceptual apparatuses. This is a manifestation of the fact that the anthropologist, as Anna Tsing points out (2005: 7), has become the champion of the local, particular and contingent. As Marcus claims, "[f]or ethnography, then, there is no global in the local-global contrast now so frequently evoked. The global is an emergent dimension of arguing about the connection among sites in a multi-sited ethnography" (1995: 99).

However, one of the problems posed to and by the anthropology of science, and the descriptive problem that this thesis grapples with, is that of *exactly* how one might talk about the universal, the global, reality or a totality at all in a disciplinary idiom that apparently no longer permits it (Choy 2005). If on the one hand science is that domain of Western knowledge that in large part bases itself on the necessary premise that universals are not created but given, and that aspires to totality, and if on the other hand anthropology is that domain that aspires to countenance partiality and particularity, then the anthropologist of science is forced into a strange version of the familiar "trap of having to believe either native meanings or our own" (Wagner 1981: 30). The strangeness arises from the difficulty of deciding exactly where to locate the "native" in this relation in the first place, for as Wagner continues "[T]he former alternative, we are told, is superstitious and unobjective; the latter according to some, is "science"" (ibid.:30). When studying scientists, the two sides seem to collapse into one another. In privileging the particular, situated and contingent - the "unobjective" - Anthropology is supposedly engaged precisely in drawing away from its natural scientific, Western tendencies (cf Ingold 2000; Franklin 1995). So how to take account of objectivity as ethnographic category? This is the analytical and ethnographic issue that might be called the problem of how to "take science seriously" anthropologically. This thesis is intended therefore not only as a contribution to our understanding of climate science in the Amazon, or of the "artifact of modernity" that is scientific data. It is also, simultaneously, intended as a methodological exploration of the particular issues that conducting such a study raises for a specific branch of anthropological theorizing about identity and alterity.

In a recent publication Eduardo Viveiros de Castro (2011a) evokes philosopher Gilles Deleuze to argue that to "take your informants seriously" is not at all to literally "believe" in or explain them, but rather to sustain them as the possibility of multiplying or increasing our world. That is, rather than having "sense" made of it – being "recognized" in the terms of Isabelle Stengers (2005: 85) – the Other should be retained and perpetuated as irreducibly outside our ability to understand. It should be incessantly *added* to our conception of the world, not reduced to it once and for all (see also Henare, Holbraad and Westall 2007). Elsewhere, Viveiros de Castro has expanded substantially on "taking seriously" as a means of "controlled equivocation" (Viveiros de Castro 2004). Drawing on his work with Amerindian peoples, he poses "controlled equivocation" as the simultaneous countenancing of all the inevitable and radically different misunderstandings that occur on both sides of any intercultural comparative act.[7] For Viveiros de Castro this is the necessary premise of any anthropological enquiry:

> "To translate is to situate oneself in the space of the equivocation and dwell there. It is not to unmake the equivocation, (since this would suppose it never existed in the first place), but precisely the opposite is true. To translate is to emphasize or potentialize the equivocation, that is, to open or widen the space imagined not to exist between the conceptual languages in contact, a space which the equivocation precisely concealed…to translate is to presume that an equivocation always exists, it is to communicate by differences, instead of silencing the Other by presuming a univocality – the essential similarity – between what the Other and We are saying." (2004: 10)

This means that what *could* bring anthropologists and those they study together is exactly what must be interrogated – we are not all "people", because notions of personhood are radically different across cultures; we are not all "men" and "women", because gender is likewise; we do not have the same figurings of nature

[7] It is "controlled", he tells us, in the way that walking is a controlled way of falling – a constant suspension of actualization.

and culture. But *in order* for these sorts of radical incommensurabilities to emerge, there must be an "ontological continuity" (ibid: 4) extended between us and the Other, a sort of symmetrical ontological determination. Otherwise, these differences would disappear, subsumed under our own categories. This is the first move, then, that posits or reveals alterity as analytical axiom; the Other of the Other is always other, as Viveiros de Castro puts it (ibid.: 12).

That these equivocations be potentialized in this way is vital in order for the second move of "taking seriously" to be effected - the creative distortion of our ideas by theirs: “anthropology is a conceptual practice whose aim is to make alterity reveal its powers of alteration” (Viveiros de Castro 2011a: 145). This has been, in other contexts, called "reverse anthropology" (Wagner 1981: 31), "recursive anthropology" (Holbraad 2012), "lateral reason" (Maurer 2005). What this implies is that the content of the anthropological exploration - the ethnography - should have a profound impact on its form. As Martin Holbraad points out, this commitment has often meant that anthropologist are forced to "think (and do) anthropology beyond representation" (2012: xvi), as representation is the abiding form that anthropological analyses take. This recursivity is the second aspect of "taking seriously", and one I think that distinguishes it from other descriptive injunctions - the “thick description” of Clifford Geertz (1973), for example. The double move entailed in taking seriously, as elaborated by Viveiros de Castro, requires the anthropologist not only to recognize that the conceptual capacity at his or her disposal is by definition insufficient to capture that which he or she intends to write about, but also to alter (at times radically) their descriptive and analytical repertoire as a result.

On both counts, however, studying science creates interesting and troubling torsions. As Viveiros de Castro points out, to "take the Other seriously" is in fact an asymmetrical task, given that almost all things that are "near to, or inside of us" (2011a: 133) must *not be taken seriously* in order for us to be able to take seriously that which is far from us (i.e. "them"). In so far as Western science is the epitome of what makes "us" distinct from other times and places, this creates a dilemma. One response is to question how one draws the line between near and far in the first place. Candea (2011a) suggests that what taking "us" seriously might do is return us always to the unsettled nature of that line. This implies that by taking ourselves seriously ethnographically we end up constantly "irreducing" (Latour 1988: 158) *ourselves*, in the sense that the endless discovery of internal differentiation that this engenders leads us to suspect that "we" are not who "we" thought we were. Thus it is precisely in the difficulty of establishing the elusive boundary between her own practices as anthropologists and those of her informants that Annelise Riles situates her analysis. In her study of Fijian women's rights movements and NGOs, Riles pictures this difficulty as a loss of the ability to "reveal", and with it, of whatever analytical purchase one feels one can lay claim to. However, she uses this apparent collapse of a crucial distinction as *itself* an ethnographic object, therefore reinstating the distinction even in its apparent dissolution. Its reinstatement consists in a refiguring, so that "what is problematized… is not so much my position in the field, as the way the field is both within and without myself" (2001: 20).

In respect of the ability to draw the line between near and far, which is necessary in order to be able to effect any sort of recursive movement, it seems that we are faced with an oscillation. On the one hand, there is Viveiros de Castro's emphasis on the necessity of ensuring "us" and the "other" perform directly comparable intellectual operations. On the other hand, for Riles, this is exactly what makes her analysis so difficult to grasp. The battle for "ontological continuity" appears differently depending on the what sort of relation there is between the anthropologist and his or her informants,[8] even as it is precisely this relation that determines what that continuity (or lack thereof) might be. Thus it might be said that *any* anthropology, whatever it studies, is involved in tracing the same shifting line between who "we" are and who "they" are: this is what is precipitated out of anthropological engagements with others. The negotiation of this relation between identity and difference could perhaps be seen as an exploration of the multiple permutations of the twin displacement: "they are not us" (exo), but "we are not ourselves" (endo-). In this

[8] The "trap" that Wagner refers to varies with this relation. There is a sense in which Riles' struggles to render knowledge as such are not the same as, for example, Strathern's struggles when writing about the people of Mt Hagen in Melanesia,. Strathern makes this clear at the end of *Gender of the Gift:* "[L]anguages themselves are not generalizable but specific phenomena. In expanding the metaphorical possibilities of the specific language of Western analysis, it can only be its own metaphors that I utilize" (1988: 343-344). Whilst Riles laments a common language, Strathern emphasizes that language is irreducibly insufficient to the task. Crudely put, whereas the former is trapped "in", the latter is trapped "out".

sense, studying home and abroad are just different versions of the same activity [9]. However, it must simultaneously be recognized that endo- and exo-anthropology are irreducibly different endeavours. Endo-anthropologists are potentially faced with a situation in which "equivocation" has no space, and "ontological continuity" is in fact the analytical problem. The oscillation lies in whether one chooses to stress the similarity or the difference between the two approaches, in terms of the relations of similarity to difference that they encompass.

How therefore might taking science seriously proceed? It may be possible to make a virtue out of what at first glance appears an impediment. Roy Wagner noted many years ago that "[A]nthropology will not come to terms with its mediative basis and its professed aims until our invention of other cultures can reproduce, at least in principle, the ways in which those cultures invent themselves" (Wagner 1981 [1975]: 30). If the objective of anthropology is to "reproduce" alterity in principle, then inherent in this very "reproduction" is the necessary contortion - what distortions *must* our concepts suffer in order to take account of an informant's explanation that, to take a famous example, witches caused his house to collapse? (Evans-Pritchard 1976). The very act of trying to establish ontological continuity, to "reproduce", forces the anthropologist to distort their own description, At the very least, our conventional ideas of causality must be revamped. The paradox here is that reproduction will necessarily result in change. If, on the other hand, we do the same to scientists, and "reproduce" what they tell us, what we will get is, presumably, science. One cannot achieve the same analytical result by applying the same approach to studies of, say, Amerindians, and of scientists. The same action has different effects.

This is the impetus behind Márcio Goldman's neologism "symmetrizations". This refers to the notion that "the symmetry between the analysis of scientific practices and African or *candomblé* ones can be obtained only by introducing a compensating asymmetry that is destined to correct the initial asymmetry of the situation" (2009: 113). Goldman suggests that it is by paying attention to what the *candomblé* practitioners are saying rather than what you interpret them to be doing, as he suggests Latour does with scientists, that the anthropologist might allow them to destabilize "our dominant forms of thought, while allowing new connections to be

[9] Like a Klein bottle (see for example Wagner 2010: 7).

made with the minority forces inside all of us" (ibid.: 123). The implication is that the anthropologist simply cannot take scientists seriously in the way he or she can *candomblé* practitioners.

However, taking my lead from Viveiros de Castro, I propose not to resolve this issue but rather to dwell in it, and push it as far as it will go in order to potentialize it. I will focus not on reproducing science at all, but *re-inventing* it.[10] This does entail paying attention to what the scientists say about what they do, not because science has the right to "define reality" (Goldman 2009:112), but because what must hold true, across the ethnographic spectrum, is the commitment to "irreduction" as the basis of anthropological enquiry. Of course, how one goes about that in the case of endo- and exo- anthropology is, as we have seen, different. That is to say, the anthropologist simply cannot take scientists seriously *in the way* he or she can *candomblé* practitioners. If, as Strathern suggests, certain sorts of ‘special knowledge’, as practices of self-reflection, have more potential to surprise and dazzle the anthropologist than other (home-grown) sorts,[11] then the task for the anthropologist of science becomes how to find that dazzle through other means. This is not a question of making the mundane exotic, so much as of learning how to "think" and not to "recognize" (Stengers 2005: 85), or how to "think with difference". In the case of the anthropology of science, what one is faced with is the *opportunity* to try to think "identity" with difference, for example. This can be very productive. If equivocation and misunderstanding are the hallmarks of an exo-anthropology that takes alterity as its axiom, then I suggest that the oscillation between recognition and estrangement – embodied in the paradox that the Other of the same may also be other - is the hallmark of an endo-anthropology that does likewise.

This is the conceptual hypothesis that underlies this thesis. The easily recognizable relation that I take as conceptual "case-study" for this is that of representation. What does it mean to take "representationalism" seriously? Do you find, as Candea

[10] Of course, to a certain extent all anthropological analysis is a form of re-invention. But this just lands us back in the oscillation that I previously delimited. See also Jensen 2012; Pedersen 2012.

[11]"It is worth remarking, however, that special knowledge which inheres, say, in theological or scientific expertise has never held quite the place in anthropological accounts as material which appear esoteric *because* they require revealing (beg immediate interpretation). An initial surprise becomes a suspension, a dazzle, and some kinds of ‘special knowledge’ are more likely to dazzle than others" (Strathern 1999: 11)

suggests, always another torsion, that distances you from that which you study? Does representationalism cease to exist when you start to look for it? On the one hand, my thesis demonstrates that it does, and that those I worked with were not representationalists, but something else entirely. On the other hand, as Martin Holbraad (2004) points out, scientists do indeed look and behave very much like representationalists when asked about questions of reality. In the case of those I worked with, the real-representation relation certainly figures importantly in their work. In the descriptions and arguments around this topic that animate the thesis, what I have tried to capture, rather than to mask, is the way the oscillation between identity and non-identity (as binary) and similarity and difference (as scale; see Mol and Law 1994: 660), plays out, both for in the practices of those I worked with, and in my own endeavours to understand them.

The Large- Scale Biosphere Atmosphere Experiment in Amazônia: some background

I will now turn to some important background on the LBA, and to the details of my fieldwork. The LBA is one of the largest of Brazil's projects for studying the environment. The overarching objective of the LBA is to "quantify, understand and model" the "physical, chemical and biological processes that control the cycles of energy, water, carbon trace gases and nutrients in Amazônia and determine how these processes are associated with the global atmosphere" as well as to predict the impact of changes in land use and vegetation cover inside and outside Amazônia.[12] The six research areas that the LBA was originally intended to focus on, according to the project documentation of 1996 available on their website, were the Physical Climate, Carbon Storage and Exchange, Atmospheric Chemistry, Land Surface Hydrology and Water Chemistry, Biogeochemistry, Land Use and Land Cover. Some time later, the category "Human Dimensions" was added, after the LBA came under fire for its failure to emphasize human activities and sustainable development (Lahsen and Nobre 2007; see also Schor 2011).

[12] http://150.163.158.28/lba/site/port/conciso/cpp05.htm Accessed August 2012

The LBA's scientific steering committee had considered it crucial from the outset to have strong collaborative international connections, in order to obtain the enormous amount of human resources and funding it required to function on the scale desired[13]. It has so far been responsible for 156 scientific projects (not including post-graduate projects), bringing together 281 foreign and national (Brazilian) institutions [14] in different areas of research from land use to biogeochemistry, although arguably its most important work has been with carbon flux. For the first ten years, the LBA was very explicitly an "international" project, that included researchers from countries including Brazil, Peru, Colombia, Ecuador, the USA, the UK and seven other European countries. It had particularly strong connections with the USA's National Aeronautics and Space Administration (NASA), and a separate section of the LBA known as LBA-ECO was established to deal with NASA-LBA collaborations, which generally focus on the research sites in Santarém, Pará.

In 2007, the LBA went through what my informants referred to as a process of "nationalization". Its contract with NASA ended, and with it the access to NASA funding, and it came under the aegis of the Brazilian Ministry for Science and Technology. This shift was accompanied by a consolidation of the seven research areas into three: physicochemical biological processes in aquatic and terrestrial systems and their interactions; physicochemical multi-scalar interactions on the biosphere-atmosphere interface in Amazônia; and the social dimensions of environmental changes and the dynamics of land use and land cover in Amazônia. This nationalization was also accompanied by a change in funding structure, that led to repeated financial difficulties for the project. During the time of my fieldwork there was also an ongoing controversy over the Codigo Florestal (Forest Code). This controversy encapsulates Brazil's long-running conflict between conservationists and developmentalists. It was against the backdrop of this increasing "nationalization" that I completed my fieldwork. My fieldwork was also bookended by the UN Conference Of Parties (COP15) in Copenhagen just before my arrival, and by the release of new data demonstrating an increase in deforestation in the Amazon around the time of my departure. Most people I spoke to were very aware of the charged eco-political

[13] http://150.163.158.28/lba/site/port/conciso/cpp04.htm. Accessed August 2011.
[14] http://150.163.158.28/lba/site/?p=intro&t=1. Accessed August 2012.

atmosphere in which they were going about their work, and also of its international dimensions (cf Barbosa 2000).

Over the years, the LBA project has involved hundreds of different researchers in different scientific disciplines. It has also included the construction of over twenty meteorological towers in the Amazon forest and the drier Central Plateau region (*Cerrado*). These towers sometimes stretch up to 60-70m high, and are bedecked with equipment that measure carbon "flux" (*fluxo*) between the biosphere and the atmosphere – how much and how fast carbon is being exchanged (normally in the form of carbon dioxide) between the forest and the atmosphere. They also have equipment that measures other meteorological variables, such as rainfall, humidity, air pressure, radiation, and wind speed and direction. Sometimes there are additional instruments installed on the towers that measure the concentration and flux of other trace gases that contain carbon, such as methane (CH_4). Only some of these towers, however, are still in operation: towers K34 and B34 at a research site called ZF2, near to Manaus; and ZF3, also relatively near to Manaus. There are also towers in Ji-Paraná (Rebu Jaru) and Ouro Preto do Oeste (Fazenda Nossa Senhora) in Rondônia; Brasilia (RECOR) in the Distrito Federal; and Santarém and Caxiuaná in Pará. Towers at Humaitá (Amazonas) and Cuiarana (Pará) will soon be built, as will a series of towers, including one 300m high, in the Uatumã Biological Reserve (Amazonas), a collaboration between the LBA/INPA, the Max Planck Institute, and the State University of Amazonas (UEA). There was a tower at São Gabriel de Cachoeira in Amazonas, in Northern Amazônia close to the frontier with Venezuela and Colombia, but it fell down in June 2010. The scientific rationale, I was told by one researcher, behind choosing different tower sites is one of representativeness. The idea is to capture as many different types of ecosystems as possible. Thus, whereas K34 was built in primary forest (un-managed forest), ZF3 was constructed on what had been farmed land; similarly, B34 in a basin, and K34 on a plateau.

The LBA has a central office at the National Institute for Research in Amazonia (INPA), in Manaus, the capital of Amazonas state; it also has several smaller offices in different Amazonian cities, as well as one in the Brazilian capital, Brasília. There were several different groups based at the Central Office in Manaus, where I spent most of my time. This office also housed the recently-inaugurated post-graduate degree course in "Climate and the Environment" (CliAmb) that offered master's and PhD degrees with specializations in different areas. The academic groups, including climate modeling groups, a hydrology group, a micrometeorology group, and a biogeochemistry group, were composed of students, data collectors and professors. Some groups were larger than others; the hydrology group consisted of only three or four people during the time I was there, whereas the modeling groups had eight to ten. There was also an IT group that took care of the LBA network, a logistics team, and an accounting team. There was thus a core set of people who were employed by the LBA either as professors, researchers, technicians, students or administrative staff. Nevertheless, the limits of the LBA were hard to hold steady, for several reasons. For one thing, there was an extended "LBA community", as those in the LBA called it, comprising of anyone who had ever taken part in a research project, collaborative or not. These were often people I did not encounter but who were on the LBA mailing lists and contributed to the LBA knowledge production, and who had access to the LBA data. There were also professors who sometimes taught on CliAmb, and different researchers who came temporarily to collect data and then left again.

Methodology and Research Sites

Despite the edges of the LBA project writ large remaining somewhat blurry, I spent most of my time with the core group employed, one way or another, by the LBA in Manaus. I conducted fieldwork mostly in the LBA Central Office in Manaus and the surrounding research sites on the forest, although I visited as many other research sites as I could. In 2007 I visited the research base in Santarém in the state of Pará, and the research site in São Gabriel de Cachoeira. In 2010 I spent time at the research site at Uatumã Biological Reserve, where the construction of the 300m tower is currently being undertaken. I also attended two large Conferences: the Brazilian National Meteorological Association Conference in 2010; and one expressly organized by the LBA in 2008.

The research site that I spent most time at was ZF2, which is comparatively near to Manaus. One of the towers there, known as K34, is the longest established of the LBA's towers. It has provided a ten-year series of data – the longest series of data for a tropical biome in the world, as several members of the LBA emphasized. At ZF2, there is a "lodging" (*alojamento*) where there is a cook, someone who looks after the lodging, and enough bunk beds and hammock space to have about twenty people there at any one time. A generator runs all day and night (although the energy it provides is "dirty" - it oscillates in frequency considerably so is not used to power the instruments on the tower). Due to the recent installation of radio transmission system that sends the raw data in real time to the LBA headquarters in Manaus, about 50 km away as the crow flies, it also has intermittent and incredibly slow internet access.

There is a trail that leads from the lodging to the plateau where the tower K34 has been built, which is approximately 3km of very rugged terrain through the forest that has been partially shored-up with planks, and which needs continuous upkeep. This is an infrastructure which, as far as this kind of scientific research goes, is remarkable. The non-Brazilian researchers who came to use the research site at ZF2 whilst I was there were often very complimentary about it, and even astounded by it. However, it requires the work and constant supervision of a team of people, including technicians and electricians, and the logistics team back at the LBA who arrange for food, people and any materials to get the site, and for rubbish and people to get off it.

Another tower at ZF2, called B34, was due to be reactivated during the time I was conducting fieldwork. It had been erected originally as part of a PhD project, but had not subsequently been maintained as a data collection platform, although the tower itself was left standing. Reactivation involved re-installation of instruments. However, although there was a lot of activity around the tower, it was hard for me to keep abreast of this reactivation as it happened in fits and starts, and I was not always sure when it was occurring. There is also another active tower near to Manaus known as ZF3, located at another different site, that I visited several times. I also spent about a week at the research site that was being constructed in Uatumã Biological Reserve, where there was already one tower built and several more in the pipeline.

At first I spent time with several different groups of researchers and students in the LBA, including the hydrology group, the modeling group, and the micrometeorology group. I tried to divide my time as equally as possible amongst them. I accompanied several students from different disciplines during the implementation of their PhD and master's experiments in the forest, and spent as much time as was possible with any foreign researchers who came to ZF2 to collect data. Sometimes, a well-known academic associated with the LBA community would arrive to collect data or for administrative reasons, but otherwise my contact with many of the more renowned researchers was limited to conferences (if they attended them). I did however conduct interviews with: most of the professors working at the LBA; the administrative staff; the logistics team; the technicians; the Data Informations System group who managed the storage of the data and the running of the network; and some of the CliAmb students.

As time went on and the focus of my research solidified, I ended up spending increasing amounts of time with the micrometeorology group, known simply as "Micro". The Micro group is responsible for the collection and processing of data from the towers at ZF2 and ZF3, as well as the maintenance of the equipment and instruments on the towers. I spent long periods of time trying to get to grips with their everyday work with the towers and the data, mostly at ZF2 but also ZF3. I also spent time following some of the data collecting technicians at ZF2 as they went about their work.

All of the interviews I conducted were digitally recorded when possible. All interviews were noted down in notebooks, and accompanied with photographs. At times, on asking questions of my informants I was directed to other sources of information, such as instrument user manuals provided by the manufacturers. I also supplemented some of the information from interviews with further material provided by unrelated sources. I have noted in the text when this occurs. During fieldwork, broadly understood as participant observation, most of my time was spent in a combination of watching and asking questions, and as I became more competent, sometimes participating. I lived in an apartment near to the Central Office in Manaus, in a block where several other LBA researchers and professors lived. I did not normally accompany any informants to their homes at the end of the working day. Even so, over time friendships formed and I was invited to dinner, to barbecues and out for drinks. During these events some opinions and views were aired which appear in this thesis, with the holders' permission. However, the majority of my research was conducted within the LBA office hours, or in the forest at research sites. I have made most of my informants anonymous, either because they requested to be made so or because they did not reply to my request to use their real names.

Organization of the Thesis - Divisive Divisions

In organizing the thesis, I decided to maintain two divisive terms, "Nature" and "Social". This is in recognition of two of the observed facts flagged above: first, that for the LBA researchers, nature [15] is considered not to depend upon their investigations, but to be given; and second, that to function, data needs to be made into something that is capable of generating and maintaining relations with other researchers and with other data, an observation that seems to chime with social scientific ideas about "the social". Data goes through a series of phases, from “raw data” (*dados brutos*) to "certified data" (*dados certificados*), so there was a very strong sense of a progressive "cutting and polishing" (*lapidar*) of the data, as one researcher put it. A particular form had finally to be revealed in order for the data to subsequently be shared. Researchers were generally very reluctant to share data that "was not ready", or had not been "worked on enough", and the LBA tower data was very rarely made available in raw state. Therefore the cleaning of the data – the process by which it becomes "ready" – assumes great importance in allowing the data to move, be used, and relate, and serves in my description as a threshold.

I therefore resolved to maintain the distinction between the natural and the social as an organizing feature of the thesis, but in order to enquire about that very distinction. I am using the terms "nature" and "social" here as what Annelise Riles has called "placeholders"; that is, only "in order to overlook [them] for the moment", as "a technique for working in and with the meantime" (Riles 2010: 803; see also Riles 2011: 157-184). A crucial aim of this thesis is to maintain this distinction, but in a way that allows it to do different conceptual work than might be expected – as it does in the practices of those I worked with. Because of the readiness with which dichotomous thinking in general has been deconstructed in the social scientific scholarship of recent decades, the definition of the specific terms in question can

[15] Or rather, whatever it is they are investigating, which they often refer to, not as "nature" at all, but in much more specific terms.

sometimes slipped out of sight.[16] What "nature" might be, for example, has sometimes become circumscribed, rather than elicited, by social scientific descriptions. [17] Despite these qualifications, I am aware that this division between the natural and the social sits uneasily with much current scholarship in STS that stresses the hybridized configurations of objects and subjects, or nature and culture as encountered in ethnographies of scientific laboratories and other sites (Latour 1993; Haraway 1991; Pickering 1992).

The study here presented draws on several different but related disciplines, including social and cultural anthropology, the anthropology of science, STS and the history and philosophy of science. I have already suggested how it relates to recent work in social anthropology. It also draws heavily on the work done in STS and the philosophy of science, most notably by Bruno Latour and Isabelle Stengers, and also by Donna Haraway. It is hard to disentangle their work from the intellectual context in which they and other scholars all sought to move beyond what Latour has called the "Modern Constitution" (Latour 1993: 13) that purifies the world into inanimate objects and human subjects. Both Latour and Stengers have problematized the relation of nature to culture within Western science itself, Latour by more recognizably ethnographic methods, and Stengers in a philosophical vein but drawing on a lifetime of association with science and scientists. Their approaches have developed in tandem and both were heavily influenced by the philosophers Gilles Deleuze and Félix Guattari.

Both Latour and Stengers insist on the importance of a symmetrical approach (Latour 1993) that positions Western science as a situated practice of working and reworking the (unstable) relation of objects to subjects, nature to culture and the non-human to

[16] "The more discretely and specifically we define and bound the units of our study, the more provocative, necessary and difficult it becomes to account for the relationships among those units; conversely, the more effectively we are able to analyze and sum up the relationships among a set of units, the more provocative, necessary and difficult it becomes to define those units" (Wagner 1977: 386).
[17] I am also aware that there has been much stimulating discussion in anthropology's trans-atlantic history as to the difference between notions of "culture" and "society", and I use the concept of society here rather than that of culture because in the second section of the thesis I want to summon anthropological scholarship concerning gift exchange, particularly in Melanesia, that gives rise to an idea of a sociality emergent from a configuration of exchange relations that separate as much as they integrate (Strathern 1988: 191).

the human. Both stress the necessity of recognizing and thus revealing the inherently political core of Western science as an ontological and not merely epistemological endeavour – that is, of recognizing the political machinery that allows Western science to claim ontological authority in declaring how the world is. Both draw out the ramifications of such recognition, emphasizing the impossibility of a world purified into an independent and inanimate Nature populated by objects, and a humanist Culture populated by subjects. They suggest instead a "pluriverse" (a term Stengers borrows from William James, Stengers 2011: 61) of hybrid entities and objects- and subjects-in-the-making that can relate to each other without any presumption of universality. Though their terminology is different – Stengers puts forward a "cosmopolitical proposal" (Stengers 2005), Latour proposes a (new) "political ecology" (Latour 2004) –they both aim at a radical re-appraisal of the way science is conceived and achieved, emphasizing uncertainty, contingency, multiplicity and symmetry.

This thesis lies at right angles to this metaphysical and cosmopolitical movement. This connected divergence is delineated by the friction that occurs when Latourian and Stengerian ideas concerning the constantly recurring impossibility of purified forms, specifically a form called "nature", are juxtaposed with the ways that those researchers and technicians I worked with go about what they are doing, and make sense of their actions. The "nature" of the researchers I worked with does in several important respects resemble the "nature" of many social scientific discussions: independent of human thought, universal and given. So, it seems, either nature is hybrid, or it is pure. However, as Marilyn Strathern notes, "it matters what ideas one uses to think other ideas (with)" (1992a: 10). In the thesis I suggest that there may be other ideas available to think science’s nature "with" than hybridity or purity. The entities, ideas and forms I encountered in the field elided and separated under their own steam - their hybridity was not an *a priori*, but merely a mark of a particular phase in the production of meaning. Therefore, although I draw on the work of Latour and Stengers considerably, I often do so in order to interrogate their arguments through the ethnographic material.

In the first part of the thesis (chapters one through four), I examine what I call "The Nature of Data". As the epigraph at the beginning of the introduction suggests, the central relation of this part is that between "nature" and "data". The first chapter explores the sort of nature the LBA researchers encounter. It follows a student at the LBA as she executes an experiment to collect data about small molecules called Volatile Organic Compounds (VOCs) for her master's project. The relation that she is interested in quantifying is difficult to obtain because of the inherent expansibility of relations understood to make up the world, or the "Earth System". There is no such thing as a single relation in this system, and so this singularization must be creatively forged. In this process of radical singularization, what is apparent is the relative unimportance of the subject-object relation, which figures as only one of the many relations that need to be excised or excluded in order for her to make the data she wants.

In the second chapter, I turn to the act of measurement. I examine the way in which data is made from the forest through a process of analogizing and discretization, and propose that this process may be more akin to a mythical transformation than to a process of abstraction of the "represented" from the "real". What is in fact being made in the production of data are a series of relationships, including the relation between "real" and "representation" itself. The world, or nature, is neither of those things. The third and fourth chapters carry this idea onwards, by describing the data "cleaning" process (*limpando os dados*), by which the "raw data" (*dados brutos*) collected from the instruments becomes "certified data" (*dados certificados*). These two chapters detail how the data undergoes different transformations, as it is detached from the world and made to relate to itself. These transformations can be parsed as phases of compaction (the elision of meaning) and separation (the extension of meaning) as the data finally becomes capable of making, or even manifesting, a certain form of relation that I refer to as "social".

The second part of the thesis is entitled "The Social Life of Data". That information can have a social life is an idea that I borrow from Alberto Corsin Jiménez. Jiménez enquires into informational transparency by comparing the incessant self-externalization or "emptying out" of the knowledge economy with practices of occultation and revelation in Yugakhir hunting transactions with the spirit world (Jiménez 2005). He argues that "[L]ike the pure, free gift (Laidlaw 2000), transparent information and real-time knowledge have no social life" (2005: 19). My analysis, however, takes the reverse tack, and asks to what extent informational objects such as data can have and indeed generate a social life, or sociality, and what this sort of sociality might be composed of.

In the fifth chapter, I examine the ways that the certified data subsequently moves around the world. I identify two different models of data dissemination system, that I call "flow" and "exchange". These two models imply, and indeed generate, two very different sorts of scientific community. That these communities can co-exist, I argue, accounts for much of the friction that occurs within scientific collaborations. I also examine what this might mean for ideas of scientific creativity and subjectivity, focusing on the inventive capacity of the data itself. This chapter thus attempts to sketch out an encompassing explanation for how and why the researchers and technicians of the LBA come together and relate in the way that they do, based essentially on the relation that each has to the product or focus of their work – the data.

In the sixth and final chapter, I point out the flaws of my own explanation by concentrating on those technicians who in fact have very little relation to the data at all. I examine the notion of exclusion that is often used in discussions of scientific technicians. I describe the ways that the different LBA technicians experience a sense of exclusion, but also the ways in which they might be seen to be indifferent to some of the exclusions sometimes ascribed to them. In order to investigate their indifference, I use movement as a descriptive trope. I argue that within their movement around systems of exchange from which they are putatively excluded there is another movement, that takes them in and out of the forest. I examine how from the perspective of this movement, shifts in scale can occur that can radically reshuffle relationships. The forest and the office therefore emerge as distinct and indifferent perspectives on each other, perspectives which are also, through the movement that mediates between them, inevitably linked and interdependent.

In the conclusion, I summarize the thesis, and return to the conceptual questions posed in this introduction, indicating some of the recursive reflections that this thesis inspires.

PART I - THE NATURE OF DATA

Chapter 1: Excision and Exclusion

Introduction

This first chapter will investigate in detail how an LBA student amasses data for her master's project. The student, E, intends to measure the effect of light and temperature on how much of a compound called isoprene is produced by the leaves of a certain tree species. The measurements she makes, and therefore the data she collects, are understood to describe (and in a sense are) the relation (*a relação*) between isoprene and the two variables, light and temperature. A 'relation' in this case follows a logic of cause and effect - as light varies, so does isoprene concentration; isoprene and photosynthsis are thus related to light intensity, and E's task is to give shape to these relations numerically by collecting data on them. This particular relation is just one of the myriad that the LBA researchers, and researchers in the Earth Systems sciences more generally, understand to constitute the world. This chapter describes how, given the way in which the Earth (and the Amazon forest particularly) is understood to be made up of incredibly complex interlocking webs of these sorts of relations between entities and processes, singling out any particular one is an incredibly arduous task.

The data that is collected by the LBA researchers and students does not come from a situation of controlled laboratory conditions; it is collected out in the "field" (*campo*), in the forest. Philosopher of science Isabelle Stengers points to what this difference might entail when she insists on a differentiation between lab science and field science (Stengers 2000 [1993]; cf Latour 1999a: 30). The field sciences are marked by the fact that they bring not "stable proofs" as a laboratory experiment might aim to, but "irreducible uncertainty" (Stengers ibid.: 144); and one of the most profound areas to which they bring that, Stengers tells us, is to the relation between subjects and objects. I would like to explore what it entails for E to singularize a relation in order for her to collect data, and what this might imply for one way that subjects and objects are understood to be constitutive of scientific practice.

Nature Out There

The researchers that I spent most of my time with are observational biologists, meteorologists, ecologists and climate scientists, or (self-denoted) "Earth system scientists". They generally do not manipulate anything in a laboratory, and if they do, it is a sample that was collected outside and brought in to the laboratory to be analyzed. Further, in the case of the LBA at least, normally someone else does this laboratory analysis, firstly because biochemical analysis is considered to be a distinct scientific activity and secondly because not many people have direct access to a laboratory. The LBA has access to a soil analysis laboratory at INPA that some people send their samples to in order to be analyzed by the technicians there, and is in the process of trying to build a gas analysis laboratory. They do sometimes use instruments, like mass spectrometers, that are more often found in laboratories. However, much more often than bringing the world inside a laboratory, they take their instruments outside, in this case into the forest, and they measure things out in the forest. In order to understand what the researchers of the LBA are doing in producing data on nature in this way, it is important to first understand what sort of nature they are dealing with.

The main point of this science, simply put, is to measure the world, and to discover new relations in it that can be quantified. To this end, they produce an enormous amount of data, which they subsequently analyze. On its website, the LBA describes itself as "one of the largest experiments in the world in the area of the environment", and speaks of its "integrated results" which have "allowed us to understand some of the mechanisms which govern the interactions of the forest with the atmosphere, as much in natural conditions (untouched forest) as in altered conditions" with the overarching questions: "how does the Amazon function as a region" and "how do changes in land use affect the biological, chemical and physical functioning of Amazônia, including its sustainability and its influence on the global climate?".[18] It lists among its results the measurement of pollution from hydroelectric dams in the Amazon; the provision of quantitative evidence of a strong relation between the formation of clouds in the Amazon and the occurrence of rain in the South of Brazil; the verification that the rates of carbon fixation and growth (of the forest) vary regionally within Amazonia; and the observation that solar radiation needed for the photosynthetic activity of plants can be reduced by 60% in some areas of the Amazon due to the smoke from forest fires. This is just a small sample of what the LBA considers to be its contributions to knowledge, and it serves to demonstrate the emphasis on discovering and quantifying interactions between entities. These interactions were variously referred to by my informants as "relations" (*relações*), "correlations" (*correlações*) and in some cases "feedbacks" (*feedbacks*). Relations here could refer to a causal relationship, where one entity or process causes the other one under investigation; or a relationship of correlation where two things vary in tandem but are not necessarily causally related. A feedback relationship is when the result of one process "feeds back" into the system that causes it, altering it by doing so. Although there are many more possible ways that connections can be characterized in the Earth Systems sciences, what was invariably impressed upon me by my informants as the salient point was the interconnectedness of the elements making up the "Earth System".

This image is not an unusual trope in the Earth system sciences – the most famous example is the Gaia hypothesis, first formulated by James Lovelock (e.g. Lovelock 1972; see also Kwa 1987), which emphasizes the self-regulatory nature of these webs of relationships that constitute the Earth System. In the same way as the Amazon can be understood to be connected to the Earth System, so to can the different spheres (Atmos-, Bios-) be conceptualized as being connected, as is the case for the different ecological and chemical domains within those spheres. It is impossible that on any scale an element could be independent. This is nicely evoked by one of the founders of the LBA, eminent Brazilian scientist Antônio Nobre, in a pamphlet intended for the public:

[18] LBA site: http://150.163.158.28/lba/site/?p=intro&t=1 Accessed 5th June 2012.

"In a useful comparison, consider that the more we investigate the mysteries of the human body, the more we discover about the microscopic and paradoxically gigantic complexity of this organism, a single species. Imagine, in Amazonia, the magnitude of complexity that the LBA researchers and students encounter in ecosystems with millions of species, millions of square kilometres, everything interconnected with the environment in a myriad of links, forming magnificent webs...despite its enormous impact and vast discoveries, [the LBA project] has just begun to uncover the “tip of the iceberg” of Amazonian complexity" (Nobre 2004: 5).

The world then, and particularly the Amazon, is teeming with relations waiting to be identified and quantified, and this excessive relationality is understood as giving rise to "complexity". This imaginary of complexity is a governing one in the LBA knowledge-practices. I was told on several occasions that working in the Amazon is much more complex than in other places, because it is a more complex environment - everything "interferes" with everything else.

All the researchers I spoke to told me that it was beyond their ability at present to take account of *all* of these relations; some told me this was because they still did not have the computational means to model such interactions, others told me that they still do not have the physical or biochemical knowledge to identify them. This "iceberg" of complexity is therefore largely invisible, or beyond our understanding, because its holistic pattern lies beyond our present comprehension: efforts to know it must concentrate on accumulating knowledge of parts of it. Thus the way that the LBA researchers, in the collective form they appear on the website, intend to understand the functioning of the Amazon is indicative of the way they already understand it to function: as a relational integrated whole that can be decomposed into processes through measurement, and then recomposed as knowledge. The emphasis on integrated and interdisciplinary studies was an important strategy behind the LBA exactly because the world is considered to be beyond any one discipline's perspective of it. As one researcher told me when we were talking about collaborating with other researchers in the Amazon, "it's just no good (*não dá hoje*) doing science today without being multidisciplinary". This is because "you just end up with very short answers, or very abstract answers: "given these conditions, this is like that". But show me these conditions! Because the tendency in nature is that these conditions interfere, they vary. This situation of classroom physics, it just doesn't exist! It isn't there in Nature, this ideal thing you might find in the museum in Paris where they keep the standards – the kilo and so on."

It also became apparent that for my informants, the act of knowing is itself also considered to be generative, because quantifying one relation only ends up in several more emerging from the effort. During one conversation, I had asked whether the researcher thought that they would ever answer the question of whether the Amazon was giving out more carbon dioxide than it was taking in or not. No, I was told, probably not. Every question answered created another ten to be asked. To measure how much carbon is being released into the atmosphere, you have to work out how much is being kept in the forest. And to know how much is being kept in the forest, you have to be able to take account of all the ways that it might escape from the forest without being released into the atmosphere. So, one would have to measure carbon transport through the decomposing leaves being swept away by streams, and this involves a whole different set of scientific investigations and theories than those being employed to estimate the exchange of gases in the atmosphere. The task seems endless. This position was stated to me in different forms repeatedly during my fieldwork. As a well-known Brazilian researcher, told me, chuckling, "but you know how it is, right? We start an investigation with one or two scientific questions, and we try to answer those questions. But almost always, these questions multiply themselves into more and more questions – the more research, the more questions there are." This was a motif of the way people explained their work to me: the more that was discovered, the more there was to discover.

Another common way that this potential for endless discovery was demonstrated to me was through the description of equations. Chemical cycles (such as photosynthesis) or physical patterns (such as convective events) can always be broken down into smaller and smaller units, and therefore the scale of any particular investigation must be chosen carefully. "Inside every term, there is another equation", I was told. Equations convey the importance of two governing metaphors in the researchers' work: balance, and flux. Carbon flux and energy flux - that is, the exchange of carbon and the exchange of energy between the biosphere and the atmosphere - were two phenomena the LBA had built the towers in order to investigate. Both of these systems of exchange are governed by rules of conservation: that is, what goes in, must come out. If a certain amount of radiation is measured as being emitted by the sun and reaching the Earth, then this is the total amount that should be measured being radiated or stored by the earth in different ways. There should be no remainders. What this means is that new discoveries are shown to have been already 'inside' old measurements; inside every term there is another equation. Thus one of the most important discoveries in terms of understanding the carbon cycle was the large concentration of tiny carbon-based molecules called Volatile Organic Compounds (VOCs) in tropical areas.[19] As one eminent researcher remarked, the contribution of these VOCs to the carbon budget is significant, and underestimated by the current calculations. Another LBA researcher had just defended his PhD on the role of newly-discovered horizontal flux of carbon through the forest. That is, the estimates of vertical exchange of carbon were not taking account of the horizontal movement of the air, which would mean that pools of carbon were collecting in certain places, and rolling from one place to another in the forest, rather than being exchanged with the atmosphere. Inside the carbon budget lies more carbon still. Not only do relations make other relations, stretching outwards as a web, but they themselves contain other relations, spiralling inwards.

The investigation of VOCs in fact uncovered their potential containment in all sorts of relations. As Antônio Nobre summarizes:

"love-making, defence against attacks, invitations to the banquet are only a few of the multiple functions of these chemical messengers….they have a critical role in the protection of the leaves against the sun. Also, they are antioxidants liberated into the air to remove dangerous pollutants, like ozone, nitrogen and sulphur. This role as tiny chemical hoovers explains why the air in the forest is so healthy, the cleanest air on Earth, purer even than the air in remote areas of the Pacific or Antarctica…Recently, research done by the LBA revealed an even more surprising role for them, the

[19] See Guenther et al. 1995. It was in fact an estimate based on a model, as there was a paucity of observed data. This article was recommended to me by E, the student whose work I will follow more closely in the following section of the chapter.

formation of clouds and the promotion of copious rain in the forest." (Nobre 2010: 41-42 my translation)

The revelation that VOCs are the precursors to aerosol particles, around which water condenses to form clouds (e.g. Claeys et al. 2004) provides a tangible physical and chemical link between the forest and the atmosphere. This relation then itself feeds into debates about the extent to which the Amazon is a self-contained system or not. The idea in simple terms is that the Amazon may create its own micro-climate, that then affects the global climate; [20] thus this relation between VOCs and clouds contains the relation between the earth and the sky, that may contain within it the key to understanding the relation between the Amazon and the entire Earth System. Relations have a habit of concertina-ing inwards and outwards.

This capacity for expansion necessitates therefore that the attempts to measure it also have this capacity. This might be seen as a question of precision. The ever-increasing world is thus matched by an ever-increasing necessity to measure it, across larger and larger areas and in more and more detail. One researcher explained to me that trying to measure the total amount of carbon in the forest simply by subtracting how much carbon enters from how much leaves is like trying to measure an inch by subtracting a mile from a mile and an inch. A difference of an inch or two for the larger distances will not make that much difference. But in terms of the measurement you are trying to reach, it can have a massive impact. Inches, therefore, start to matter a great deal. In fact, smaller and smaller units start to matter more and more. The instruments that the LBA uses measure in micromol, and parts per billion. The world is endlessly fractioned, and this quest for ever smaller and smaller units is understood to be limited by the instruments being used, rather than the world's capacity to *be* that small. The world is therefore understood to be on *all* scales. This world that the researchers are dealing with is endlessly expandable; relations beget relations, and inside every whole, there are more and more parts to discover. In this sense, the

[20] An extended controversy concerning this can be found around the so-called ”biotic water pump” – see A. M. Makarieva and V. G. Gorshkov (2007).

expansion of the whole here is in fact a result of it constantly being discovered to be bigger than the sum of its parts, by being in fact always lacking, or *smaller*. [21]

E's experiment [22]

As Chunglin Kwa (1987) has argued concerning ecological conceptions of nature in the USA around 1970, different metaphors for nature can "suggest very different understandings of how to control nature" (1987: 430). The model of nature that the LBA researchers work with likewise necessitates the employment of particular ways of getting to know it. It is against the background I have above described that those researchers I accompanied are trying to collect data and make knowledge. In an attempt to ascertain what this might involve, I devoted some time to following different students who were in different stages of their master's and PhD projects. The LBA had, at the time I was there, a newly-created Postgraduate programme called "Clima e Meio Ambiente" (Climate and the Environment), shortened to "CliAmb". It was intended as an interdisciplinary post-graduate course in different climate and Earth system sciences, and attracted students from all over Brazil and Latin America who wanted to study the Amazon forest and its biological, ecological and microphysical processes. One of the students I accompanied was E, who was at the time a master's student on the CliAmb course. She is from São Paulo state, and did her undergraduate degree in Biology. I spent quite a lot of time with E during the research phase of her dissertation, going with her into the forest, collecting data with her, and talking about her experiment with her.

E wanted to study the correlation (*correlação)* of the emission of the most common of the biogenic VOCs, called isoprene, with photosynthetic activity in certain tree species in the Amazon, as well as examine the causal relation of light and temperature changes with both isoprene and photsynthesis. As she writes in her final dissertation,

[21] I thank Martin Holbraad for this insight, and will return to this observation in the conclusion of the thesis.

[22] When the people I spoke to referred to the activities I am going to describe, they called them “experiments”, “projects”, or “studies”. Experiment seems the most fitting word to use here because I am going to concentrate on the primary phase of a scientific study, which is the planning of a methodology and the collection of data using instruments, rather than a latter stage such as analyzing data or publishing an article, or the project as a whole. It also allows me later to engage in a discussion concerning scientific experiments in other scientific disciplines.

knowledge of how isoprene emission changes "in relation to variations in light and temperature, could contribute to an understanding of chemical reactions occurring in the atmosphere" (Gomes Alves 2011: iv. My translation). It is understood, therefore, that what she will discover is potentially *already* linked to other interactions ("chemical reactions") occurring in the atmosphere. It is in the face of this acknowledgement that what she is quantifying as a separate phenomenon is actually interconnected, that she must execute her experiment.

The final methodology that appeared in E's master's dissertation reads thus:

"This study had as its objective the identification and quantification of the emissions of Isoprene and photosynthesis in different levels of intensity of light and temperature, in three phenological phases (Late Mature Leaf, Old Leaf, Recent Mature Leaf) of Eschweilera coriacea (Matamatá verdadeira), as this species is the most numerous in Amazônia. The photosynthesis measurements were taken between 8 and 12 in the morning using an Infra-Red Gas Analyzer (commercial portable system, LI-6400, LI-COR, Inc., Lincoln, USA).To measure the emission of isoprene, the air provided by the LI-6400 was directed into a sample bag (0.5-litre Teflon Bag). The concentrations of isoprene were determined using a PTR-MS (Proton Transfer Reaction Mass Spectrometer; Ionicon Analytik, Innsbruck, Austria). The measurements for both processes were taken at different levels of irradiation intensity (from 0 to 2000 µmol m-2 s-1) and temperature levels between 25 and 45°C." (Gomes Alves ibid. My translation)

It is important to note the specificity of the naming of the instruments, and each aspect of the methodology – 8-12 in the morning, 0-2000 micromol/m of radiation, 25-45 degrees, the species, the names of the instruments. These are the outlines of the relations she wanted to investigate. It is by holding these aspects of the experiment steady enough to control and to be certain of, that she can allow isoprene and photosynthesis to move. The work that goes into affording this control is extensive.

First, she must search the literature and choose the right genus of tree – one that previous studies have shown to emit isoprene. This already cuts away a vast amount of forest. Then there is her measuring method. Rather than just measuring the concentrations of isoprene in the atmosphere in the forest as other studies have done, E wants to measure it right "on the leaf" (*na folha*). This means she wants to take the air samples directly from the area around the leaf, at the same time as she is measuring photosynthetic activity, which has never been done before in the Amazon to the best of her knowledge. She would do this by hoisting an instrument that can measure photosynthetic activity, called an Infra-Red Gas Analyzer (IRGA), Li-6400, up into the treetops. Collecting the air that was expelled from the IRGA, she intended to take it to another instrument to be tested for isoprene concentration, a gas chromotograph (GC). Although there has been research already done on relating VOCs with photosynthesis in which leaves have been cut from trees, cutting the branches is known to "stress" (*estressar*) the leaves, causing their stomata (little holes in the leaf which control the flow of water and gas into and out of the leaf) to close. The closure of the stomata affects how much gas and water the leaves release. As photosynthesis is the process by which carbon dioxide, water and light energy is absorbed and converted into sugars and oxygen in the plant, the effect of the cutting will therefore interfere with a measurement of this process. Although this was a well-known interference effect in plant physiology experiments, and the careful researcher can take several precautions to try to reduce it, E wanted to make sure she did not have to cut the branch at all. This meant she would have to somehow reach the top of these 35m trees with the IRGA. As far as E was aware there were no other studies done in Amazonia employing her particular methodology, probably because the IRGA is so heavy and expensive. It was very important, she told me, that the same air sample from around the leaf was collected to be taken back down the to the ground, so it could be analysed by the GC for isoprene concentration.

E's methodology is the first procedure by which she starts excising the relations that she wants to measure from the enormous expanse of relations, interconnections, causes and effects, that is the Amazon forest. Her methodology gives her the outline of that relation; its specificity serves to cut away vast sections of the forest, the day, and in fact the rest of the world. This tree, not that; this time and not any other; this light, not that; this specific instrument, and not another.

Behind the Scenes

This tight control, however, is not always successful, as E soon realized. What E's formal methodology does not include is all the work it takes to obtain that relation that has been so carefully delimited. In this sense, then, there is an enormous amount of work "behind the scenes", that does not appear in her dissertation, but is crucial in order for her to obtain the data she needs. This is not only something that social scientists have pointed to for some time (e.g. Latour 1987), but also something that most scientists themselves recognize. As one told me, "science is always messy". However, it was also clear that a lot of the work that E had to accomplish and many of the obstacles that she had to overcome are in an important sense *specific* to setting up an experiment in the Amazon.

Money was one such general obstacle that took on specific contours in the context of E's research. E had already done one collection of air using *cartuchos* (cartridges/canisters) when I met her, but unfortunately she had not been able to analyse them at INPA (the national research institute of which the LBA is a part) because of the lack of a laboratory. INPA had been planning to build a laboratory for gas analysis for some time, and when she had first planned her experiment, the Executive Manager of the LBA had told her that it would be ready by May 2010. The works were only just starting in September 2010. It was a question of money, E told me – they did not have the money to start until now, especially as this sort of laboratory is very expensive to build as you need to install three different gas pipes in order to provide the gases necessary to do the analysis. INPA had been given a GC instrument from the Max Planck Institute in Germany several years ago, but it was collecting dust in a cupboard. The Executive Manager suggested installing it in a laboratory at INPA used for soil analysis, as the GC only needs one gas pipe – but still, "it's going to take a while, there's no money", E warned.

As the laboratory was not ready in time, E took her cartridges to a laboratory in São Paulo to be analysed with a mass spectrometer instead of a GC. However, the mass spectrometer was also out of action at the laboratory in São Paulo. So she carefully packaged her cartridges, full of Amazonian air, and sent them via airfreight to her co-supervisor in the United States, whose laboratory had a GC he could use to analyse them. However, when her supervisor got back to her, he told her that her samples had spent too long in the cartridges – 12 days in São Paulo and then 15 days on the way over to the States – and they were contaminated. It had been impossible to separate out isoprene from all the "noise" – there were just too many compounds. E suspected that it hadn't just been the time spent in the cartridges. When she had gone to do another trial with the IRGA, she had found the values to be impossibly high compared to what was expected. She suspected that, as the IRGA is shared by several students doing different studies, the tubes were full of molecules stuck onto the tubes of the IRGA from the other student's experiments, which then contaminated the next experiment. She decided she had to change the tubes of the IRGA to ones made of Teflon, which is much less "sticky". But this still left her without any sort of instrument to measure the Isoprene concentration. Luckily, a visiting researcher from the USA took pity on E, and told her that she could use his instrument, called a Proton Transfer Mass Spectrometer (PTR-MS) that he had brought with him from the USA, to analyse her samples. This would allow her to perform a much more precise analysis than she would have been able to with a GC, she told me. The shortage of laboratories and instrumentation and the problem of contamination were suddenly resolved by her friendship with a foreign researcher who had brought an incredibly high-tech instrument into the forest. I asked her what she thought of this; "I'm lucky!" she said with a grin.

It was not only the threat of contamination that E had to deal with. In order to perform these measurements in the trees, she had to have scaffolding erected that would allow her to sit in the top canopy with the IRGA, measuring photosynthetic activity and simultaneously collecting the air to be sampled by the PTR-MS in little plastic bags. She told me that one of the major problems was finding the people to put these scaffoldings up. As she wanted to repeat her experiment on several individuals of the same species, these structures had to be put up and taken down at various different times. At one point I noticed that there were several up at once, surrounding the tower at K34. They looked incredibly precarious to me. However, aside from their doubtful structural integrity, their presence caused a great deal of concern amongst the micrometeorologists from the LBA who monitor the equipment on the tower, because one of the scaffoldings had been erected right on top of their soil heat flux sensor, which caused some grumbling. They had to ask her to take it down and relocate it in order not to interfere with their measurements. Taking these structures down and putting them up safely necessitated someone experienced in this sort of work – "you can't replace these people, people who are prepared to risk themselves like that" - and the *técnicos* (technicians) who did this at the LBA were short on the ground, especially at the time because there was a large project going on in another part of the Amazon. E was desperate at the time I met her to find someone to put her scaffoldings up because her trees were losing their leaves; not only that, but she needed someone to help her lift the heavy IRGA up and down the scaffolding. It was impossible, and dangerous, to do it alone.

In the end, it took her so long to begin taking measurements that the trees belonging to the species that she wanted to study started losing their leaves. They would grow new leaves, but the problem was that young leaves are not the same as old leaves. Young leaves do not emit so much isoprene – they use all their energy for growth. The second problem was that they would be hanging downwards as they emerged and grew, which meant that she would not be able to assume that they were getting equal and maximum radiation, as she could if the leaves were mature and lying flat. In her final methodology, therefore, she had to specify exactly that she had measured three stages of leaf growth, young, mature and old.

E's experiment, therefore, not only necessitated controlling the light and temperature by means of the IRGA - but also the time of day, the height she is at, the condition of leaf she is measuring, the tubes of the IRGA, the weather conditions she measures in, where she measures, the *técnico* putting up her scaffolding, and the instruments themselves. Other researchers' interests have to be negotiated, and resulted in this case in E having to take her scaffolding down, and further negotiate with others, such as the *técnicos* who have no personal interest in her measurements, in order to put it back up again. This labour does not appear in her final dissertation, but the researchers I spoke to about it shrugged it off as a given of all scientific work, especially in the Amazon forest. There are massive logistics teams for programmes such as the LBA - "science is always messy", and luck has a large role to play sometimes. Because of this, the invisible controlling work that is done is intensive,[23] and must remain stable and constant for her to make the comparisons she wants to. If the conditions vary – if Adilson is not there to put the scaffolding up, if it rains, if she is at the wrong height – then she cannot make any comparisons at all, and the relation she wants to investigate is out of her reach. Furthermore, she has to get it right there

[23] As actor-network theorists and feminist theorists have often emphasized (e.g. Latour 198; Law 2004; Star 1991), and which I shall return to later in the chapter.

and then – in this sort of science, you only have windows of opportunity, up the scaffolding on a dry day at the right time with everything ready. If she does not measure between 8 am and midday, she has to wait for the next dry day to do it again, and in a rainforest you can never be too sure when that is going to be. It is not so much that certain factors are out of her control, and others are within her control; rather, the impression that I had from accompanying her over several months as she installed her experiment was that her control had to be extendable if necessary. She needed to be able to artificially control everything *around* the phenomenon she was interested in in order to allow it to vary naturally.

Keeping Things Stable

Stability was an overriding factor in maintaining this background of control. This was brought home to me on one occasion that I accompanied E into the forest to collect data. She climbed precariously up the rickety scaffolding, and Adílson the *técnico* hoisted up the IRGA, hand over hand, waiting every so often to ensure that it did not hit anything on the way up. The IRGA never travelled unaccompanied – the head of the INPA plant physiology group insists that one researcher, C, who had worked the longest with the IRGA, accompany it at all times. I sat in the tower with C, and E and Adílson crouched on the top of the scaffolding, which wobbled alarmingly. Using pure nitrogen, E had already very carefully cleaned each bag that she was going to use to transport her air from the IRGA to the PTR-MS which is waiting back at the lodging. She had numbered and labelled them, and she laid them out beside the IRGA.

She carefully chose a leaf that had no fungal infections that she could see, or holes in it, dried it carefully with a paper towel as best she could, and fitted the IRGA chamber around it. "There are more mature leaves now, that's good news!" she exclaimed. The LI-6400, the infra-red gas analyser E was using, has a chamber which you close around a leaf, providing a sealed and controllable space around it. The instrument allows you to measure the changing CO_2 (carbon dioxide) and H_2O (water vapour) in the chamber as you vary the temperature and light input into the chamber, as well as calculate indexes such as stomatal conductance (how open the stomata are). These variables are the ones known to indicate photosynthetic activity, which chemically proceeds by absorbing CO_2 and producing H_2O (water vapour) and O_2 (oxygen). This means that photosynthetic activity can be measured through differences in CO_2 and H_2O concentration in the chamber over time. The water vapour and CO_2 are removed from the incoming air, so that the exact concentration of CO_2 can be provided using an external source, such as a canister. However, as E did not have an external source – it is expensive and you can only get it from the USA, C told me - instead she just used the ambient air and its average amount, trying to control what she could by tying a plastic bag over the input to the IRGA so that no tiny wind eddies outside the IRGA could affect the flow rate and therefore the concentration of CO_2. Adílson sat with his legs swinging over the side of the scaffolding platform, making sure that the bag was attached to the input pipe.

As we sat and waited for E to get the instrument ready, C explained to me how the IRGA works. Inside the IRGA, the air passes through a flow valve, which controls the flow rate (which has to be constant as well), and then splits into two channels. One of these goes into the chamber where the leaf is sealed, and where there is a detector called "sample"; the other goes past the "reference" detector. It is the difference between these two readings that provides the information about photosynthetic activity – the difference between what you know about the air, and what the leaf has done to the air. In this case, how much CO_2 has been taken out of the air as the plant photosynthesises gives you the rate of photosynthetic activity. That day, E had decided that she was going to vary light and see what effect that had on the activity of the leaf - how CO_2 concentration (that indexes photosynthetic activity) varies with light. The air that is expelled from both detectors is the air she collects in her bags to analyse for isoprene on the PTR-MS - how isoprene varies with light. She will therefore be able to analyze how both photosynthesis and isoprene production varies with variations in light (the "relation" between them).

We waited for the CO_2 to stabilise, and set the light to 0, the first setting in her range – to simulate night time. E carefully noted down the number of the bag and the level of the temperature and radiation. But something seemed wrong – the input CO_2 refused to stabilise, it just oscillated wildly. "Have you matched?" C shouted across from the tower. "Matching" is a technique to ensure that there is no difference between the two detectors, without having to remove the leaf from the chamber – you can allow the air that is flowing into the chamber to be measured by both detectors, and adjust the sample to "match" the reference. In the manual, it suggests this should be done several times a day when using the IRGA. But that could not be the problem, she had matched already. Perhaps the bag is too small, she shouted to C, it keeps emptying - and the CO_2 was still oscillating wildly. "There's something wrong" C said, frowning. Perhaps she has to use an external CO_2 source in order to get the CO_2 input to be stable, he mused (this indeed turned out to be the case when we returned to collect more data on another day). But on Sunday, just a few days ago, it was functioning well, E said. E tried in vain to work out what she might have done which may have affected the instrument since then. After an hour or so of tinkering, they decided that it was just not going to work, and rather forlornly she climbed down the scaffolding, and we trudged back to the basecamp to do tests on the IRGA. She had not managed to collect any data that day.

I suggest that the reason that E was unable to collect any data that day is because she was unable to single out the relation she wanted - the relation between CO_2 and light - from all the background noise. She might have recorded something, but it would not have been the relation she was looking for. In experiments such as E's, comparisons will always be made, whether one is fully aware of what one is comparing or not. The danger in a world that is so full of relations is unknowingly making the wrong comparison, and unwittingly investigating the wrong relation.[24] E was so concerned about the state of the leaves because if she picked a leaf with fungus on it, she would end up comparing isoprene production and photosynthetic activity not just to each other but to the fungus as well. If she did not take her measurements at the right time of day, this would have an effect on the photosynthetic ability of the leaf, which she needs to control. If it rained, she would no longer be comparing photosynthesis in the conditions she specified, but under conditions that may have a very different effect. If she mixed up the bags, or mislabelled them, then she could not say for sure that this particular isoprene production is related to this particular photosynthetic rate. The point is, *whether she maintained control or not*, a relation of some sort would still be elicited, or cut out of the world. But it is up to her to make sure it is the *right* one, because the comparative potential in the forest can be overwhelming. Rather than

[24] One researcher told me that for years when measuring ammonium nitrate concentrations, some of the molecules were sticking to the tubes they were using to suck their air samples into their gas analyzers. The data they had about ammonium nitrate concentration turned out to be data about partial ammonium nitrate concentration.

intervening in the world, what is always threatening in this case is the world's intervention in the experiment. As Isabelle Stengers describes it, the world here is putting E to the test (2000 [1993]: 134). But it is exactly because of this that it has to be ingeniously excluded.

Making Zeroes and Reductionism

The way therefore to ensure the right relation is being cut out is to exclude everything else. One of the most common ways that exclusion operates in this case is through stability. Keeping everything else constant and unchanging is the principle operation through which E can allow isoprene's relation to photosynthesis appear. This stability is effected in various ways: that the air used for both the photosynthetic measurements and the isoprene measurements is the *same* air is vital, for instance, because then it ceases to make a difference. Nitrogen, which E used to clean her bags, is such an ubiquitous gas in experiments of this kind because it is inert – being stable and unreactive it provides a starting point for subsequent variation. Another student researching emissions of gases from the soil had several large, clear plastic chambers made for him by a mechanic in Manuas, that he inserted into the ground to a certain depth so that he could literally cut a section of the soil off from the surrounding atmosphere in order to ensure that it did not get wet, but was still affected by the sun. When I accompanied researchers on an experiment into the north of the Amazon, I was asked to help construct a flat, dry, clean area of about 3 by 4 metres in the middle the forest, in order to build a plastic child's Wendy house, which they had purchsed in Manaus and brought with them up the river to the research site to provide their instruments with a clean, dry, protected shelter. On whatever scale, eking out these spaces of stability is an essential first step towards being able to record data. The most important prerequisite for any comparison made in the forest is that everything that is extraneous or outside be held steady, be it the flow rate of the gas or the concentration of CO_2 in the air.

If the researchers are not successful in this task, then comparability disappears. At the end of that particular day in August, E had failed to get any data because she had not been able to stabilise the CO_2. Not being able to hold steady the CO_2 concentration flowing into the IRGA meant that there was no way to get at the difference between the sample and the reference. It is this difference, between the amount of CO_2 in the air before the leaf photosynthesizes and after, that contains the relation that E wanted. But without having a fixed and stable CO_2 concentration beforehand, it is impossible to measure the change. Likewise, her co-supervisor had not been able to tell which of the compounds was isoprene in the samples that E had sent over to the USA because there had been too much contamination - there was too much "noise". "Noise" is the name given to those relations that the researchers do *not* want to investigate - that is, the rest of the forest.

It is because of this extraordinarily excessive relational potential that the ability to make *no* relations is, I would argue, one of the most important capabilities which the instruments provide researchers like E with. Much of the literature on the subject has focused on the role that scientific instruments play in mediating and constructing what is discovered using them. There are many convincing arguments from different disciplines problematizing the notion of instruments as transparent observational conduits to Nature. [25] But I would like to instead emphasize the way that the instruments the LBA researchers use keep certain things *invisible*. Making something visible requires that lots of the world be made invisible. The question as to how much the instrument mediates what it shows, and the provenance of that mediation is often

[25] The ability to make invisible things visible is a frequent feature of the historical characterisation of scientific instruments, especially those revolutionary optical instruments of the 17th Century such as the microscope and telescope (Schaffer 1989; Bennett 1989; Schickore 2001), although ideas of transparency and opacity are recurrent in the studies of many different sorts scientific instrumentation and technologies (cf Gooding, Pinch and Schaffer 1989). A particularly frequent observation across this literature is that instruments are not transparent mediators between world and scientist. They are not only themselves "reified theorums" (as Gaston Bachelard denoted them, cited in Rheinberger 2010: 27), but what is seen is only seen via whatever intellectual framework is in vogue at the time. The instrument may determine how we see the world, but how we see the world also determines what they show us. A concern with technology has been very prevalent in feminist theorising (Suchman 2009; Franklin and Ragoné 1998; Thompson 2005), for example, particularly when such an intellectual framework concerns gender specification, such as with New Reproductive Technologies and medical technologies such as ultrasound (e.g. Barad 2007). As Barad writes "the sonogram does not simply map the terrain of the body; it maps geopolitical, economic, and historical factors, as well" (2007: 194). The foetus therefore becomes a gendered person, even a consumer product (Taylor 1998) through the application of this technology. Drawing attention to the way that these identities are technologically constructed therefore draws attention to the contingency of the construction, rather than the inevitability of the identity. This concern with the simultaneous co-production of technology and society and the resulting hybrid entities, cyborg bodies (Haraway 1991) and heterogeneous networks is pervasive in STS as well as feminist theory (cf Bijker and Law 1992).

taken up in debates as to the nature of what is being revealed. [26] However, concentrating only on what is revealed can blind the investigation to the particularities of the actions involved. It is as crucial that the instrument is able to "reveal" *nothing*.

Instruments allow the researchers to propagate 'zeroes'. This term, to "zero", is taken from the process of calibration and verification that all instruments should undergo before being taken out to the field. I observed this process several times being performed by members of the micrometeorology group. Verification and calibration, for the technicians and researchers of the LBA, are ways of giving the instruments zeroes and differences. In the case of the IRGA for example, a calibration performed by the LBA micrometeorology team first involves "zeroing": running nitrogen through the instrument and explicitly telling it "this is zero CO_2", and "this is zero H_2O". Afterwards, a known concentration of CO_2 is passed through the IRGA, and it is told "this is air with 338.98 parts per million of CO_2" (for example). [27] During the time I was at the LBA, a great deal of time was spent trying to "zero" the instruments in this way. This zero, 'zero difference', is as important as being able to register 'difference as data'. In fact, you cannot have the one without the other.

Whereas this might seem obvious, what is perhaps less obvious is that 'zero' is a not a normal state of affairs, from which one simply elicits information such as gas concentration. It is in fact very much the production of sustained and concerted effort, as I witnessed when E struggled for an hour to keep the CO_2 flux into the IRGA steady, or as I experienced when I sweated trying to erect a Wendy house in the middle of the forest. 'Zero' is in fact one of the most crucial measurements that is made, although it does not directly count as data. It is in fact the necessary condition for data.

---

[26] Here W.D. Hackmann's distinction between "passive" and "active" instruments in the 17th and18th centuries is perhaps relevant to think with, although not without its problems. In Hackmann's formulation the former type of instrument measured and revealed nature, whereas the latter were "philosophical" instruments that interacted with or reproduced nature in the laboratory (Hackmann 1989:39-40). My quibble would be with the passive role here assigned to revelation or measurement, as I explore in Chapter Two. See Bennett 1989 for a critique of Hackmann's typology.

[27] Which of course begs the question of how one measured *that* concentration, without calibrating. I will return to the circularity inherent in calibration as a more general feature of scientific practice in Chapter 4.

The instruments that I have mentioned all manufacture this in different and ingenious ways. As C explained to me, different molecules absorb different frequencies of light, so the IRGA emits infra-red radiation (IR) of a single, particular frequency, that will be absorbed by the molecule being measured (in this case H_2O and CO_2). There are two detectors, that measure the reference gas in one chamber (which holds the incoming CO_2 that E had such a hard time controlling) and the sample gas in another (which is the air around the leaf that has been sealed in this chamber). The two IR light beams that pass through each chamber are then absorbed by the CO_2 in each chamber, so what the detectors on the other side of the chambers pick up is the light that has *not* been absorbed in each case. It is the difference between these two – before (reference) and after (sample) – that the IRGA 'detects'. But the only way it can detect this difference is by making constant reference to the known, stable CO_2 concentration in the reference chamber. [28] Thus, according to the manual, the most important prerequisite for using the IRGA to measure photosynthetic activity is that "incoming concentrations must be stable". [29] It is only this control you have over the incoming air that allows for the right information to be extracted from the activity of the leaf. You must know how much is coming in to be able to measure how much is being produced, and in order to know that, you must keep it stable.

The PTR-MS, [30] the other instrument E used, works on the principle of mass spectrometry. This method relies on the fact that if you subject a moving object to a sideways force, how far it is deflected from its original path depends on its mass, as

---

[28] This information was obtained from the IRGA manufacturer's website: http://www.licor.com/env/products/photosynthesis/technology.html Accessed July 2012

[29] This is achieved by passing the air that has been sucked into the IRGA over a dessicant, normally silica, which controls the amount of water vapour in the air. It also passes over a $C0_2$ "scrubber" which takes the CO_2 out of the air, so that the exact concentration can be provided by an external CO_2 source. It was this externally provided stability that E did not have.

[30] The GC, that E did not use in the end, uses a principle known as adsorption, which simply means the adhesion of gas or liquid molecules to a surface. Different molecules have different adsorption rates, so when the sample air is sent through the GC, it passes along a column that is full of a substance that causes the different molecules in the air sample to "stick", slowing their passage down at different rates. This means they exit the instrument at different times. The air sample is thus differentiated into its constituent parts - but this will only occur if the temperature, vapour pressure and column length remain steady. If the column length is unknown, or the temperature and vapour pressure oscillate, there is no way of knowing what is coming out of the column at any one time. Two different compounds could come out at the same point in time at different pressures and temperatures. That is, unless you can hold the temperature and pressure completely stable to within tenths of a degree, with zero change, you do not know what you are looking at.

long as the speed and the size of the force are known. For example, one explanation online[31] suggested thinking about it as if you have a cannon ball moving through the air, and you try to deflect it with water from a hosepipe. Nothing much will happen to the cannonball`s line of flight. However, if you add a tennis ball moving through the air, then the tennis ball will definitely move, even though the cannonball keeps on its original path. This is the principle applied at the atomic level – that deflection is related to mass - so you can tell the difference between cannonballs and tennis balls even when you cannot see them. In a mass spectrometer, the deflection is provided by a magnetic field. For molecules to be affected by a magnetic field, you need them to be charged – so you have to first ionise them.[32] Spectrometry works on the idea that the substances to be sampled therefore have a mass, and a charge – and that these will govern their movement though a vacuum with an electro-magnetic field (it has to be a vacuum so that they do not bump into other air molecules). Two ions with the same mass-to-charge ratio will move in the same way through this vacuum. But this only works if they are subjected to the same electric and magnetic forces. If that were to oscillate or change, you would not in fact know what you were looking at.[33]

In short, the instruments E use are all able to perform the minute excisions that they make through creating backgrounds that are unchanging. However ingenious the act of molecular isolation is, it can only be registered as against something that inherently provokes zero difference. These are substances and states that can be characterised as being simultaneously present and absent - the state that is implied, for example, in the attempt to manufacture zeroes, or nothings.

---

[31] http://www.chemguide.co.uk/analysis/masspec/howitworks.html Accessed May 2011. I used this example because it seemed to explain the process in a way that a layperson could understand, unlike the complex information that E directed me to.

[32] In the case of the PTR MS, this is done by using H30+ (hydronium) – this is the "proton transfer" of its name. The PTR MS produces these hydronium molecules from the air, and then, when the sample VOC is injected into the initial chamber with the hydronium ions, a proton transfers from the hydronium to the VOC, giving the VOC a charge of +1 – so it now has a positive charge. This charge attracts them further into the machine, separating them out from the other molecules in the air. This information was sourced from the PTR-MS manufacturers website, as I was directed to by E (who was also trying to work out how the instrument functions at the same time): www.ionicon.com. E also showed me information that an Ionicon engineer had sent her, but I found this very difficult to understand, and therefore did not use it.

[33] When I asked her, E explained how the PTR-MS works to me in basic terms. I found however that I needed to supplement this for my description, and found an online resource: http://www.astbury.leeds.ac.uk/facil/MStut/mstutorial.htm Accessed February 2012

Making nothing where there is everything could be seen as a very particular form of reductionism. In it, what is known is analogous to what is held steady, and therefore what is ignored or invisible. One 'knows' in order to ignore. This is also a dynamic that is apparent in the methodological decisions that E makes. E had to try to ensure she was taking every factor into account, exactly in order to discount each one. Thus, the smaller the relation one is cutting out of the world, the bigger the world must become. This general principle was apparent when E noticed that the leaves she wanted to use were in different growth phases. This discrepancy had to be incorporated into her methodology; it was by including this observation that it could be 'discounted', as it were, as a discovery. In fact, it comes to refine the relation, singularizing it further. It is no longer just the relation between isoprene and photosynthesis in all leaves, but this relation in three different phases of leaf. In this sense, the effect of singularizing a relation through what you know makes the relation singular through an additive process. Reducing the world in fact necessitates that you include as much of it as you can. The act of singularizing a relation holds within it a complicated choreography of potential and multiplication, and, I maintain cannot be understood only as a form of simplification (*pace* Star 1983), or construction. Cutting a relation out of the world is a complicated process: refining a relation means including more of the world, and some of the world that you include you do so in order for it to be ignored. What you know becomes invisible, and what you do not yet know, with all its internal multiplications, is brought into sharp relief.

Subjects and Objects in E's Experiment

If the preceding analysis is accepted, it has some important repercussions for the way that certain aspects of scientific practice are understood, particularly relating to ideas of objectivity and subjectivity.

E conducted a specific sort of experiment. [34] Francis Bacon, who proposed the "experiment" as the "royal road" to knowledge, was adamant that experimenting

[34] Scientific experimentation, as a crucial aspect of understanding scientific activity and knowledge production, has received increasing attention from philosophers of science, historians of science and in the practice-orientated approach of STS over the past 20 years. Unsurprisingly, in each discipline it is considered to be important for different reasons. Philosophy of science sees it as a turn towards a more nuanced understanding of the reality of

involved some degree of manipulation of "Nature", famously urging 17th century experimentalists to "twist the lion's tail" (cited in Hacking 1983: 149). Registering or measuring the effect of instigated perturbations in order to learn about how something works seems to be an important part of making knowledge from scientific experiments. The extent and type of this manipulation, however, can vary widely and it has been argued (Galison 1997) that the term "experiment" in fact has no stable definition over time.

Despite this, one particular interpretation of experimental activity has come to predominate in contemporary STS and Philosophy of Science: that, in one way or another, scientific experiments in fact *create* the phenomena that they purport to investigate (Hacking 1983, Harré 2003, Barad 2007). One of the reasons that establishing the emergence of different types of novel entities in science is of enduring interest for the social study of science is because it points to the plausibility of constructivist interpretations of science. The "construction" of entities in physics such as quarks, or neutrinos, is an example (cf Pickering 1981). Of course, the terms of the debate itself are not new; the wrangling about the nature of science in constructivist/realist terms has been ongoing for decades, and as such there are many different strong and weak stances taken on all sides (Smith 2005).

According to philosopher of science Ian Hacking, for example, an experiment is an "intervention" in the world (Hacking 1983), not just a manipulation of it. Hacking gives the example of Hall's effect (1983: 226), in which the "photoelectric effect" is created by the particular apparatus and action of that apparatus. It is therefore this experimental set-up that allows for the emergence of an entity that otherwise would not exist. Hacking, however, would not agree with the suggestion that "the electron" was created as the photoelectric effect was (See Hacking 1991, 1992). Karen Barad, on the other hand, suggests the more radical concept of "agential realism", which bases itself on a reading of Niels Bohr's dictum that "the nature of the observed phenomenon changes with corresponding changes in the [measuring] apparatus"

science ("experiments have a life of their own" Hacking 1983). Historians of science track its emergence as the hallmark of the new breed of empirical scientists of the 17th Century (cf Gooding, Pinch, Schaffer 1989). STS scholars take understanding it as potentially a methodological necessity in order to "follow" scientists ethnographically (Latour 1987, 2005).

(cited in Barad 2007: 106, italics removed). Thus for Barad, the electron would exist only as part of the apparatus that 'perceives' it, and in the "cut" that that apparatus in so doing makes. The phenomena that populate the world therefore are apparatus-object complexes that can be "agentially cut" in different ways to give different material configurations of "human, non-human and cyborgian forms" (ibid.: 178). Bruno Latour, another prolific author on the topic, has argued extensively through many different publications and permutations for the co-construction of the two domains of object and subject, and finds it useful to employ philosopher Michel Serres' terms "quasi-objects" and "quasi-subjects" to take account of these hybrid entities that take the place of "two pure forms known as Object and Subject/Society", which are only "partial and purified results" of processes of hybridization that characterize the sciences (Latour 1993 [1991]: 79).

Although there are many different positions taken on this question, then, what interests me is that these debates are framed as if the world can be arranged along a scale. This scale has "objects" at the realist end, and "subjects" at the constructivist end (for an example, see Latour 1993 [1991]: 51). One extreme of this scale privileges the object and objective knowledge: the world exists independently of thought and is waiting to be discovered. At the other extreme it is the "subject" and subjectivity that is privileged, such that the world is made by the relation that human beings (or subjectivities) have with it, and thus cannot exist independently of human thought and action (cf Meillasoux 2008). Western science generally, in as much as it can be taken as a coherent and unified body of knowledge-practices,[35] is often understood to be orientated towards the "object" end of the spectrum - that deals with matter, physical reality or "Nature". When that object is held to be independent of all practices of representing it, this version of the distinction in question has often been called "representationalism" (Barad 2003: 804). As science is understood to assign objects this form of independence, it is therefore considered also to be inherently representationalist.

[35] The idea that Western science is a unified body of knowledge-practices has been challenged explicitly (cf Galison 1988, 1999; Hacking 1992). Nevertheless, this broad characterization of "objectivity" would probably still hold generally across the differentiations suggested in these challenges, such as that between the laboratory or experimental, theoretical and instrumental sciences.

However, an abiding prerequisite for this configuration seems to be the image of a 'sliding scale' between objects and objectivity, and subjects and subjectivity. Thus in Barad's rendering, one can "cut" at any point along that scale, refiguring what is a subject and what is an object each time you do so. In Hacking's, the line has been shifted a little bit towards the object end in order to encompass the creation of new phenomena in the laboratory (but not so far as to encompass the electron, that remains obstinately on the other side). In the case of Bruno Latour, the two terms that structure his philosophy likewise both depend on the notion of a continuum. In the first, "purification", reality is understood as a process of accumulation, and described as a progressive move towards either the "pure object" or "pure subject" at opposite ends of the scale. The second term, "mediation", is understood as the inevitable hybridity that this purification necessarily commences from (and also engenders). Latour explicitly and graphically demonstrates this by indicating that mediation starts from the middle of a scale and works outwards towards the poles (Latour 1993 [1991]: 51).

There is one particular repercussion of understanding the subject-object relation as a spectrum or scale that I would like to focus on. This is how "objectivity" is therefore conceptualized.[36] Objectivity is often described as the necessary exclusion of the self – the "subject" - from the world being studied, or at least the claim to be able to do so. A facet of this exclusion is the concern that scientists have "to eliminate the mediating presence of the observer" (Daston and Galison 1992: 82). As historian of science Peter Dear summarizes:

"'objectivity' tends to be conceptualized in terms of its opposite, 'subjectivity'. 'Subjectivity' connotes variability and contingency, the perspective of an individual human 'subject', prey to local circumstances and moulded by them into a distorting mirror. 'Objectivity' is, by contrast, the 'view from nowhere', in Thomas Nagel's words, with no local circumstances- and no perspective- to distort the mirror." (1992: 619).

[36] Although historian of science Lorraine Daston suggests that "objectivity" is a rather confused notion, referring to "metaphysics, methods and morals", concerns that have varied enormously through history and vary even in present usage: "[W]e slide effortlessly from statements about the 'objective truth' of a scientific claim, to those about the 'objective procedures' that guarantee a finding, to those about the 'objective manner' that qualifies a researcher" (Daston 1992: 597). See also Porter 1995.

Also implied in this exclusion is the more mundane sense in which scientific practice is understood not to permit any sort of "self-expression" or "creativity", as anthropologist James Leach explores (Leach 2011). In these studies and many others, a great deal of attention has been paid to the practices that remove the scientist (subject) from that which he or she is studying (object/world) or that which he or she produces (scientific text), giving both object and text therefore an authority deriving from the 'view from nowhere'.

The scale that underlies such a presumption is therefore zero-sum, or conservative in a thermodynamic sense. That is, the terms involved are related in such a way that as you decrease one term, the other necessarily increases. As the 'object' side goes up, the 'subject' side goes down, and vice-versa. Getting 'closer to nature' or 'closer to the object' therefore necessarily means the progressive eradication of the subject altogether. Adherence to scientific objectivity in this sense implies reductionism: that there is in the end only one singular object, the world, and concomitantly only one correct way of explaining and describing that world – an objective, scientific one. Objectivity thus comes to indicate the exclusion of the subject and subjectivity; it represents an insistence on ontological singularity at the expense of subjective, or espistemological, plurality.

However, this trope of a sliding, conservative scale cannot cater for one very important fact. This is that it is a specific object-subject *relation* at stake, not just a particular object, or a particular subject. As historians and sociologists of science have taken pains to document in their emphasis on the contingency and situated-ness of scientific work: this relation has changed over time (cf Daston 1992). Thus what is actually being removed, in the conservative scenario I have suggested, is not 'the subject', but all sorts of different subject–object relations, leaving one particular one remaining. The image of a sliding scale is in fact the end product of this process, as it presents itself as the *only* relation an object and subject could have to each other. However, in describing objectivity in these terms, what is indirectly implied is not that 'subjectivity' is absent, but that different and specific subject-object relations are absent.

On the basis of the ethnographic description of this chapter, I would like to argue that for E and her colleagues, it is not in fact 'the subject' that needs to be removed, but *all* *the relations other than that which is of interest.* These may be between subjects and objects, objects and objects, subjects and subjects, in any form, because the emphasis is on the singular relation that is to be excised, and *not* on that which is being cut away. The particular relation of the scientist to the object of study, for example, becomes only one such relation amongst many that has to be expunged, in order for E to make the data she needs - unless that is the very relation of interest, which is of course sometimes the case.[37] Indeed, excluding her 'self' is the least of E's worries. Taking measurements in the Amazon, E is beset by all sorts of threats from other potential relations that will remove this ability to produce scientific knowledge. The removal of the self is actually just one very small part of what needs to be made invisible. E also has to make all the other relations she *could* be measuring in the world temporarily disappear. Thus this deletion is far more radical than has perhaps been previously described.

This has repercussions also for understanding what it means that E did not mention the trials and tribulations she underwent to execute her experiment in her dissertation, beyond mentioning the change in instrumentation. Such omissions are standard in scientific dissertations, and indeed in scientific texts in general. I was not present at E's master's viva as it was after my return from the field. However, having seen others' defences and having talked to E about hers afterwards, I found there is a place to discuss some, but not all, of these aspects omission in the space provided by a viva. People mention problems with sample contamination, for example, in the discussion section of a dissertation or thesis to explain results that have gone awry.

Despite this, the removal of other details nevertheless seems to be a clear indication of the purging of perspective and the mechanization of the process. It seems to suggest that *any*one could have done the experiment, when as everyone knows, it was in fact someone in particular who actually did. Thus it feels like "human agency is written out of these accounts" (Gooding 1990: 3). This 'writing out' has been a key piece of

[37] See Matei Candea's work (2010) for the distinction between behavioural data and weight data collection, for example. I would add here a comparative remark that from the analytical point of view of what is excluded, what is included becomes multiple; from the point of view of what is included, what is excluded becomes multiple. Thus in Candea's paper, concentrating on what is excluded, i.e. the animal as a whole, demonstrates that what is included, i.e. the animal as partible, is more than suspected. On the other hand, here concentrating as I am on what is included, i.e. the singular relation, demonstrates the mulitplicity of what is excluded in order to do so, i.e. the rest of the world.

evidence in attempts to overturn universalistic conceptions of objectivity and to stress instead that, as Donna Haraway puts it, all knowledge is situated knowledge (Haraway 1991: 183-201).

However, if knowledge genuinely cannot escape the situation of its production, as Haraway and others have argued, then it follows that there must always be a context. What seems to cause concern is not the absence of any context whatsoever, but the removal of a very *specific* context. 'Decontextualisation' may mean removing a particular context – in this case the context of production perhaps. But it necessarily also implies 'recontextualisation' – into the context of academic publications, for example (Jensen and Winthereik 2013; see also Strathern 2002: xiv). When I asked E about including the specific obstacles that she had had to overcome in her dissertation, she told me that she thought it would indeed be very useful to have a separate place where such information could be shared, so that others who might conduct similar experiments to her in the future could learn from her experience. But, she told me, it would not contribute to her knowledge production about isoprene. Information about her experience is a different sort of information to that about isoprene, and it therefore has a different context of production. Further, if E is mostly concerned about excluding an enormous array of interfering relations from her measurement of one particular interaction, then it might indeed seem strange to single out only one of the excluded relations - for example, that between herself and the IRGA - to include in her description. Therefore, I am tempted to echo Marilyn Strathern in asking what if this problem were also a fact? (1992b: 92). It is of ethnographic importance that E cuts away what she does in the way that she does. Making the conditions of the production of her data textually invisible fits in to a larger dynamic of singularization that is not simply about the removal of a particular subject, but the removal of a myriad of different relations. In this sense, her methodology is part of the singularizing apparatus as much as the IRGA is. It is exactly the *indifference* between the relations that each of these different instruments exclude which is the most telling.

All sorts of relations have the same sort of ontological valency when it comes to producing knowledge through exclusion; it is as important to control Adílson, as it is to ensure she does not breathe on the IRGA chamber, as it is to ensure that she chooses the right leaf, or that it does not rain. Failing to ensure any of these could in the end elicit the wrong relation. This is in a certain sense what actor-network theorists have suggested in their emphasis on the symmetry between the agency of humans and non-humans in the networks that produce scientific facts (Callon 1999 [1986]; Latour 1993, 2005). However, what my analysis suggests is not a binary between pure forms on the one hand and hybrids on the other, as might be suggested by a schemata that arranges the world along an axis with subjects and objects at either end. The putative opposite of a particular subject-object relation is in fact a whole mish-mash of different relational configurations that is not that easily captured by the category of "hybrid", and might better be thought of - *temporarily,* for these relations might well be the subject of subsequent experiments - as "noise" or nonsense.[38] It is that which must be controlled, kept stable, and ignored, for it threatens any chance of making meaning out there in the forest.

Conclusion

The idea that an experiment might create a new phenomenon would not make sense to the LBA researchers with whom I worked. They do not consider themselves to be in any way constructing the world that they are measuring; the relations they measure are given in the world, and this world pre-exists its measurement. It is generally accepted that these relations are given in the world independent of the action of the scientist, although as far as the researchers I worked with are concerned, this given-ness is rather unremarkable - they were generally indifferent to my questions about the nature of that given-ness: 'given-ness' is itself given. E therefore must transform

[38] This comparison has further elements to be explored. Latour suggests that even as the sciences purify the world into objects and subjects, they also "mediate" all sorts of objects and subjects to form these "imbroglios", "hybrids" of "quasi-objects" and "quasi-subjects". There is therefore a peculiar doubleness to Latourian analysis, such that even as the sciences purify the world, they *also* create hosts of these hybrids: "The less the moderns think they blend, the more they are blended. The more science is absolutely pure, the more it is bound up with the very fabric of society" (Latour 1993: 42). The way in which in order to reduce or singularize the world E in fact added 'more' to it might be a correlate of Latour's insight. However, it is not that E added and subtracted at the same time, but that addition *is* somehow excluding. Knowing more is in order to reduce further. In this sense, then, in Latourian terms, purification (or singularization) would need to somehow also *be* mediation (addition). That is, such a formulation would have to think around the way that purification is opposed to mediation in Latourian analysis.

the world, but not in order to create it. [39] Drawing attention to the crucial process of singularization, I have tried to demonstrate that it is constituted by practices of excision and exclusion, that simultaneously perform a strange double move of addition and inclusion. This move is more complex than a mechanical reduction. I have tried to demonstrate that the emphasis is not on the creation of new entities, but on this incredibly complex practice of revealing relations 'enough'. The forest, understood here as the 'world' out there, is constantly expanding and contracting, and insistently intervenes. It is taken as pre-existing its own description, but this *a priori* is expressly *not* stable. It is this stability - often a zero - that the scientists must add to the world in order to measure it. The forest must briefly be given the relation which is "no-relation", in order for E to countenance it.

In a certain sense, then, E's experiment sidesteps constructivist-realist debates altogether. E is not really "constructing a fact" (Latour and Woolgar 1979), even if it could be argued that she is contributing to the construction of some fact or blackbox about isoprene in some way. I would rather here draw attention to what she produces there in the field. When I ask to see E's results, what she shows me is her data. This data is the form that the relations she is investigating take when they leave the forest, and is the point and product of the whole exercise. What she creates is data, then, and not nature.

It is important however, in order to learn from my own description, that I now proceed with caution. It is tempting to suggest in light of this observation that in a 'representational' practice such as Western science, data and nature are inverted - so that the representation now can be seen to take the place of the referent. What we thought was a representation (data) is actually *real*, and what we thought was real (nature) is actually the artifice. This is a type of figure-ground reversal, such that the

[39] This argument deserves a proviso, however: the experiments that Hacking, Harré and Barad (amongst others) all refer to are experiments in physics. The most infamous of physicists, such as Richard Feynman, Albert Einsten, Niels Bohr, Werner Heisenberg to name just a few, are notorious for having had intense debates amongst themselves concerning the nature of reality. For this reason, they have been the subject of numerous brilliant and thought-provoking studies by scholars from outside physics who are interested in the same questions, including philosophers and sociologists of science (cf Barad 2007; Kumar 2008). I suggest that physics, as a discipline studied by other disciplines, lends itself particularly well to constructivist-realist debates; as Isabelle Stengers points out, the vocation of the physicist may be "inherent in the art of fabricating "factishes", which singularizes physics" (2010 [2003]: 20).

data (representation) and the nature (referent) swap places, and implies a constructivist position – the data 'makes' nature - rather than a realist position – nature 'makes' the data. But such a reversal only works in an imaginary in which data and nature are related representationally. This representational axis, or scale, as I have suggested, is a particular understanding of the relation of subjective to objective, or representation to represented.

I have tried to demonstrate in this chapter how this representational relation of object-subject may not be a governing trope for the scientists I worked with, and how in fact they were indifferent to this particular relation. I have argued that with reference to what needs to be achieved, it is simply one relation amongst many. It would be strange to proceed on the assumption that those I worked with proceed by de-subjectifying themselves and their surroundings, when they so clearly de-objectify it as well. But it would be equally strange to proceed by assuming that what they produce are representations, simply because that is what is presumed to lie at the other end of the scale. Data, I suggest, does not stand in relation to nature as the subjective does to the objective. If the analytical purchase of such figure-ground inversions in fact belongs to a representational rhetoric that may hide more than it reveals, then so too might the imaginary which envisages subjectivity and objectivity as the finite content of a single relation, such that as one increases the other decreases.

I started the chapter by suggesting that I wanted to explore the world that the LBA finds itself in, inspired by Isabelle Stengers' suggestion that the field sciences bring uncertainty to established relations between subjects and objects. I have endeavoured to demonstrate that E's work out in the forest does certainly bring into question some established social scientific ideas regarding the subject-object relation in field science. But Stengers in fact meant something else. Her argument is that the field sciences invert the established roles of those who ask the questions (the subjects), and that which meekly submits (the object), by allowing the *object* itself to object. The 'givenness' of nature here is not in question, as she also points out: "no-one in fact doubts that the terrain exists, that it pre-exists the one who describes it" (Stengers 2000 [1993]: 144). It is this pre-existence that "forbids mobilization...In effect, the terrain does not authorize its representatives to make it exist other than where it is."

(ibid: 144).[40] What this implies is that field science must follow the world, rather than interrogating it in a laboratory. She also maintains that as a corollary, the risk, or "putting to the test" that the subject-object relation implies needs to be conserved (ibid: 134) - but not in any static fashion. The end point of her argument is to establish the imperative of a constant re-beginning ("*recommencement*" ibid: 70-87), in which subject and object are constantly re-figured anew, in uncertain and unpredictable ways.

Even as we grant this, however, it is worth noting that the researchers I worked with actually have a very clear idea of what uncertainty is and what form it takes. That is to say, they have also singularized uncertainty. The uncertainty that Stengers is talking about is an unpredictability that stems from relinquishing control of what the object might say. Among the LBA researchers, uncertainty takes the form of gaps, that are always presumed to exist. The balances they work with never close - as several researchers explained to me, there is always some part that does not add up. It may get close, but there is always a lacuna between what is measured going in, and what is measured coming out. As one researcher told me "The difference is veeery small. It's 90 and something percent...you'll never have 100%, right? It's impossible to have that. But it's the best closure (*fechamento*) you're going to get." On enquiring about this at another time, I was told offhandedly that "we know we are not looking at the really really real (*o real real real*). But it's what we understand as real". These gaps are the subject of the next chapter.

[40] This is another point I would have liked to elaborate on if there had been time, as there is an interesting distinction to be made here between the "representation" of Stengers and Latour, and the 'representativity' that Stengers is here I think speaking to - that is, the extent to which you can try to make the "terrain" stretch as far as it can.

Chapter 2: Myth and Measurement

Introduction

In the last chapter, I examined the radical singularization that is necessary to cut a single relation out of a forest of relations. The forest, understood here as the ‘world’ out there, is constantly expanding, and insistently intervenes. Its given-ness, therefore, is expressly *not* stability. It is this stability that the scientists must add to the world in order to measure it. One of their most important initial tasks is therefore the creation of a position of zero relations. I suggest this is one of the most important and often unnoticed affordances that scientific instruments provide. However, these instruments are only used to afford such a perspective in order to produce data. This production is carried out through the act of measurement.

Measurement is a crucial moment in the production of the LBA data, as it is the moment that the data comes into existence as such, as a set of digital numbers that can then be processed. There have been several important historical studies that explore the gradual construction of systems of metrological standardization across different countries and in different periods. These studies tell tales of conflicts and compromises that have arisen between countries, ideals and institutions in the effort to quantify nature (Schaffer 1992, 2000; O'Connell 1993; Alder 1995; Porter 1995). That these numerical efforts go hand-in-hand with particular political and social configurations has been well-documented in this literature: measurement emerges as a means to control and govern socio-historical realities. Thus the history of quantification is one often told from the perspective not of mathematicians or scientists, but of administrators and bureaucrats (Hacking 1991). As historian Matthew Norton Wise remarks, "when we ask about the most general source of the desire to quantify, we find it more nearly in the requirements for regulating society and its activities than in the search for mathematical laws of nature" (1995: 5). Other scholars go so far as to suggest that in addition to regulating society such practices might even create that which they purportedly merely numerically describe (cf Bowker and Star 2000).

In this chapter I take a different route in an attempt to understand what measurement does in the practices of the LBA. I trace the different stages that constitute the act of measurement as it occurs mechanically on the level of the instrument, (as it does for the most part in LBA experiments). The first stage I identify as 'analogizing', and the second 'discretization'. In the first, an analogue of a natural process is created, and in the second this analogue - often an electric current - is transformed into a discrete and digital medium: data. This mechanistic approach is deliberate. I am attempting to describe the creativity of measurement without resorting to social constructivism (cf Fortun 2011: 9). My description demonstrates that a very clear distinction is maintained throughout between nature and that which is created - but that this distinction is not between real and representation. What this distinction might be, and what it lies between, is explored using the anthropological analysis of myth, specifically in Amerindian cultures, as a heuristic. I then suggest that this may throw a different light on one notable model of data collection that is provided by Bruno Latour.

Instruments as Analogy Machines

To explore the mechanics of measurement, I will concentrate on the instruments on the tower K34. Over the course of the time I spent with the Micro team, I often went into the forest to the tower with them. This was more often than not in order to perform routine maintenance work on the tower or to collect data, but also frequently to investigate an instrument that was behaving strangely or to replace one that was broken. On these occasions, we had to climb the towers to get to the instruments that needed tending to, and on the way up I made enquire as to how the various instruments worked. These trips were quite relaxed affairs and interspersed with my questions, people would gossip and chat as they worked, sometimes dangling from harnesses from the side of the tower. Although most of the members of the team had some idea about the principle behind each instrument, there were only a few, including L the electronics technician, who could explain to me in any detail how the instruments operated. Over these fragmented but frequent excursions, I became familiar with how each instrument functioned (to a certain level of technicality), information that was often supplemented by reading the manuals that someone in the Micro team gave me. I also got to know when an instrument or measurement system was considered to be malfunctioning - when it was not producing data that was good (*dados bons*), as was often the case.

At the very top of the tower, sticking up into the air above the trees, is a carbon flux measurement system, known as an eddy covariance system after the method used to calculate the flux. This comprises of a sonic anemometer and an infra-red gas analyzer (IRGA). As one of the Micro team explained to me, the sonic anemometer measures wind direction and speed in three axes ten times a second (10 Hz) by pulsing sound waves between its six prongs and measuring the interference caused by the wind eddies that pass through it. The IRGA Li-7500, the model on the tower K34, measures carbon concentration also at 10 Hz, through an 'open pathway'. This means that it measures just what passes between its detectors, which are out in the open rather than in a chamber as with the IRGA that E used. When the data is brought back to the LBA office in Manaus, those in charge of processing the data use a specially designed programme called Alteddy to calculate the carbon flux using the wind eddy speed and direction data and the carbon concentration data. Also on the tower K34 is another IRGA – the Li-820 or Li-840 – that measures what is known as the "profile" of CO_2 concentration by sucking air in from different heights up the tower, to give cross-sectional measurements of the forests' atmosphere. There are also all sorts of meteorological instruments installed along the length of the tower: a barometer that measures atmospheric pressure; a thermohygrometer that measures relative humidity and air temperature; a pluviometer, that measures rainfall volume and intensity; a cup anemometer that measures wind speed; radiometers, that measures different sorts of radiation, emitted and reflected and photosynthetically active in this case, (although one of the Micro team was keen to impress upon me that different towers can have different radiation sensors). There are soil latent heat flux sensors, and soil temperature and humidity sensors in the soil. All of the instruments are connected to dataloggers, simple computers that store the data, that are installed at different heights on the tower.

Thus the tower sits in the forest gathering moss and mould, whilst the instruments on it register changes in the relevant properties of the environment continuously. The result of these measurements - data - is transmitted in real time (as long as the connection is working) from the dataloggers to the LBA central office via telemetry. The instruments are often called "sensors" (*sensores*) - because they sense the world they are in. I would like to dwell here on what this sensing, that takes place sometimes at an incredibly high frequency, mechanically involves. Some phenomenon outside the machine - in 'nature'- has a property. Let us take humidity of the air, for example. That property has a relation to a property of a material inside the instrument. Originally, for example, in the instruments that measure humidity (called hygrometers) there was a human hair.[41] Human hair expands and contracts according to humidity in the air. In these original instruments, a stylus was attached to the hair, with its tip resting on a revolving drum of paper. As the hair lengthened and contracted, it caused the stylus to move up and down, recording the change on the revolving drum of paper (Knowles Middleton 1969: 81-132). This provided a continuously varying line that could then be 'read' using a scale. The modern electrical hygrometers that are installed on the towers have a semi-conductor[42] in them rather than a hair, but the principle is basically similar. As L explained during a short course he gave in electronics to the entire Micro team, it is the conductivity of the semi-conductor that is affected by the property of nature in question. Lithium chloride is a common semi-conductor in electric hygrometers, as its resistance (how much current is let through it) changes depending on the amount of water it has absorbed from the air. There are also other substances used, called capacitors. In both cases, however, the conductivity of the substance is correlated with the relative humidity of the air.

Most of the instruments that are on the tower convert one property into another property using semi-conductors. They almost always convert a property of the

[41] As a standardized instrument, it was stipulated that the hair had to be blonde (Knowles Middleton 1969: 85).
[42] Semi-conductors have revolutionized electronics, L impressed upon us several times during the course he gave. Through different processes, a pure, stable chemical (such as silicon) has impurities added to it that affect its chemical structure in such a way that it has either one free electron, or one free space for an electron - making it a ‘semi’-conductor of electricity with very particular properties. This means, for example, that current can run in one direction, but not the other; or that you can switch current on and off. The whole of computing is dependent on this property of semi-conductors, L told us.

environment into an electric current, which, L told me, is simply the movement of electrons.[43] So, what is the relation of the property outside the instrument, and the property inside the instrument? What is the relation of humidity in the air to lithium chloride's resistance? One does not stand for the other, nor is one a reduced or abstracted version of the other. More correctly, one could be seen as an *analogue* of the other (cf Edwards 2010: 205). The current is proportionate to the humidity, but structurally different from it (unlike a *homologue*). Furthermore, the current is the given, natural property of lithium chloride – it is the movement of lithium chloride electrons - as much as the humidity is considered to be a given property of the air. The one does not represent the other by being somehow less "real" than it.

The LBA instruments seem to sit at what might be considered to be the interface between the forest and knowledge, or even nature and culture, as philosopher of biology Hans-Jörg Rheinberger discusses. He describes different interfaces between nature and instrument, and tends to place these interfaces at the intuitive separation between instrument and world – the cut of the biological sample that the microscope peers at, the skin-to-machine interface between test animal and testing instrument (2010: 220-222). He asks, when describing a test-tube centrifuge, where exactly nature would be considered to begin and culture to end (ibid: 224). But the boundary between nature and culture in the case of the instruments I am discussing does not seem to lie where the instrument meets the world in this intuitive sense. In the case of the LBA instruments, this boundary where sensor meets world is one that produces analogues, not the different ontological domains of "nature" and "culture"; I would therefore be reluctant to assume that it is "culture" that is being mediated at this interface. Rheinberger's question in the context of the LBA points not to the difficulty in pinpointing where the distinction between nature and culture is being fabricated, but rather to the difficulty in describing this interface, and what it mediates, at all using these terms (cf Wagner 1981 [1975]: 24). The moving electrons in the lithium chloride are not any less natural, real or physical than the humidity in the air, nor do they symbolise the humidity, or abstract it from the world. Even with the more complex instruments, such as mass spectrometers and IRGAs, the transfer of

[43] In fact, although current is conceptualized as the flow of electrons in a certain direction around a circuit, it is actually the creation of spaces for electrons to jump into, L informed me – so the current in fact "flows" the other way to the way that the electrons move.

world to instrument is through a current that is instigated by the ions or charged molecules in the world themselves actually hitting a detector – they are already a current, as it were. The continuous process that is the constantly varying natural property produces another continuous property, a current, that is also a property in the world - a property of a physical substance and not a "sign" (Hankins and Silverman 1995: 113). Therefore, I would argue that these interfaces do not mark an ontological divide between nature and culture, but mediate the creation of an analogous version of that phenomenon that is being measured.

The question then is not a case of determining where nature and culture meet, but of investigating what sort of process this analogizing is. Several authors have suggested that ("philosophical") instruments of particle physics, such as the air pump or the cloud chamber, work by "reproducing nature" (Hackmann 1989: 42; cf Galison 1997:46). As Hackmann points out, the phenomena created in 17th-century laboratories were considered to be the same as those that occurred in nature. The lightning of the lightning machines was the same as the lightning during thunderstorms; Newton's rainbow from his prism was the same as a rainbow in the sky. The underlying logic is one of a continuity of principles, based on the Aristotelian "order of natural symmetry", which stated that everything in the world behaves according to the same principles, giving rise to "the intuitive feeling about the underlying regularity of natural processes" (Hackmann 1989: 42). This allows for analogy, understood as extension, to work as a way to make knowledge. Thus, famously for René Descartes, a heart works like a motor, an eye like a camera. Analogy here then relies on the continuity of principles across a divide.[44] It is worth, however, qualifying the sort of continuity in question with the meteorological instruments I am talking about. It is not that the continuity of principles inside the

[44] I should point out that Hackman is sceptical of this form of scientific argumentation, remarking that "logically there was no reason why the phenomena recreated in the laboratory with models should be the same as the natural ones. This was the fundamental weakness of the analogous argument" (1989: 57). But, as science historian J.D. North writes "[M]y concern is with analogical *argument*…Analogy is the basis for much scientific conjecture, but even conjecture is an art, which can be done well, done rationally, even though it might prove in the end to have yielded a false conclusion…Not all nonsense is equally foolish" (1989: 285). North also points to the “very great role played by the *theological* debate in the history of analogical thought...the *justification* of the analogy between human nature and the nature of God has always been at the centre of Christian theology” (1989: 292), an interesting avenue for further investigation of the role of analogy in observational science.

instrument and outside the instrument creates both the humidity and the current, for example, as it is with Hackmann's example of lightning being produced by 17th century instruments. It is rather that one of the properties in question – electrical resistance – is correlated to humidity in the air in such a way that it reveals it, but not as the working of a pump can reveal the working of the heart. The form of the revelation is not a mechanistic imitation, nor does it reproduce nature. Electrons do not work "like" humidity, nor does humidity resemble a current. The form of the correlation is different, and so is the form of the revelation. I would suggest, given this, that what Hackmann describes could instead be called *homology* machines, in reference to the biological term that indicates a common underlying structure to different physical forms; whereas what I am describing are *analogy* machines, in reference to the very different structure that humidity and electron have that nevertheless reveal something about each other. In the case of the LBA instruments, the analogy machines of the instruments are sometimes called "transducers".[45]

Measurement as Discretization

Thus before any measurement occurs, the instruments that the LBA use first create an analogue in the form of a current. This current, however, must still be measured; it is still not revealing of anything about the humidity in the air. It must become data to do that. As a researcher, who specializes in building low-cost sensors, told me "the world is in analogue; everything is analogue, life is analogue. The digital world does not exist. You have to transform one into the other, because computers are digital. When they pass the information to us, computers make it analogic again. And each time you do this, you lose information." The transformation of the electrical current into data takes place in another element of every electrical sensor: the analogue-to-digital converter (ADC). The ADC is what converts the analogue signal – the current - to digital data – the number. L explained to me that an analogue signal is converted into digital code by assigning a value in the continuous signal to 0 in the binary code, and

[45] A fact I garnered also from L’s electronics course. See also here Webb Keane, 2011, "On Spirit Writing: The Powers of Transduction Across Semiotic Modalities", key note address, Danish Research School and Anthropology and Ethnography Annual Megaseminar.

another to 1 in the binary code. Then, whenever those values were registered, they would be converted into 1s and 0s (see figure 1).

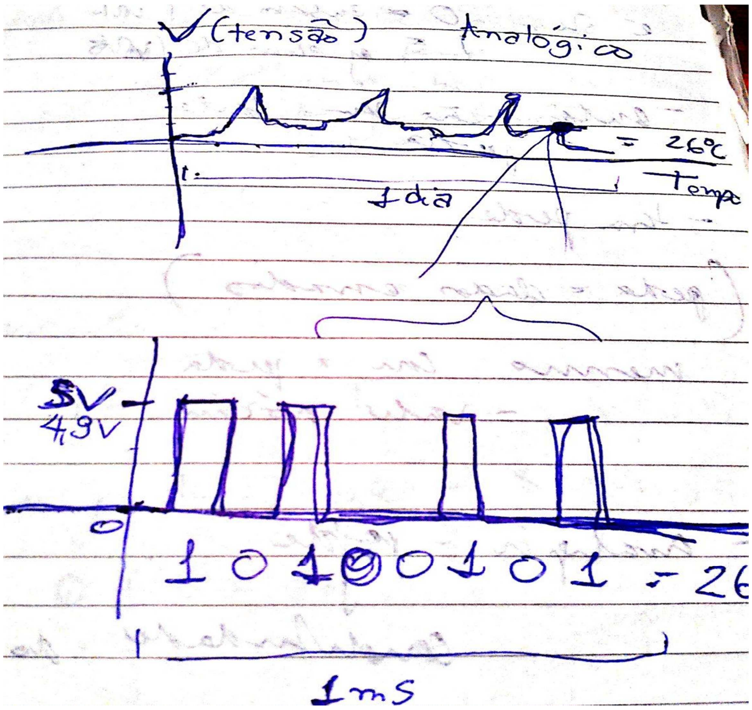

Figure 1: Diagram that L drew for me to demonstrate the conversion of an analog signal (26° C, above) into digital (also 26°C, below). Note the difference in scale between the two graphs. Taken from my fieldnotes, August 2010.

These 1s and 0s then can be processed either as they are in binary form by computers, or transformed into numbers for the people in the Micro team to process. The digital form is one which allows computers to act upon it and manipulate it. The data that the LBA produces is thus the product of a series of stages: from the creation of a continuous analogue of a continuous property in the world to a discrete digital number. Not all the technicians or researchers I spent time with had specialist knowledge of this process. However, it is a testament to the importance attached to the process of electronic transformation that they all considered L an integral part of the team, even though he was newly-hired and so lacking experience, testifies to the importance of this process of electronic transformation. All of the researchers I spoke to who were conducting, or who had previously conducted, research with the LBA confirmed the absolute necessity of having someone, normally a technician, who knew how the instruments worked electronically – especially in the Amazon forest, where the humidity and the dirt can damage the electronic equipment in a matter of days.[46] During the time I was at the LBA, L gave several classes on basic electronics to members of the team that maintain the instruments on the tower. A more formal class was also arranged on the more complex aspects of programming a datalogger and installing instruments correctly with specialist technicians who came up to Manaus from the Laboratory of Meteorological Instrumentation (LIM) in INPE, in the Southeast of Brazil.

Talking to these specialist technicians from LIM about their non-stop schedule, I realized how important this electronically-mediated process is. Not only were the technicians asked to install experiments from scratch for researchers or students, but often they were contacted by researchers who did not know exactly what had gone wrong with their instruments - only that they were not producing data anymore. "Some people from [a state in Brazil] want to study fluxes in sugar cane plantations, but they don't have the material, they don't have the knowledge, the technical knowledge, how to install the equipment. We go there and we give support, as much in the theory as the installation and programming" one LIM technician told me who had worked as a meteorological technician for several decades. He also told me that sometimes he had to ask the researchers to take a picture of the equipment and send it to him by email, because he could see immediately what the researcher could not, and could direct them how to fix their instrument by email. What is interesting is that these technicians had no interest whatsoever in the *content* of the data, only in the process of its production. As the head of LIM told me about the data itself, "I don't use that information, there's no way for me to keep up with the data like that", whereas the scientific researchers, another member of LIM tells me, "know what to do with the data but have no idea about the instruments" and so they "might know what data [an instrument] gives, but not how it functions". The importance of the technicians' work to the production of data is significant despite this lack of interest in the data itself because, I suggest, they mediate this crucial process of electronic transformation from analogue into digital - from the continuous into the discrete. It is the transformation, not the data, that counts here.

[46] Interestingly, the other role considered of utmost importance for the success of a campaign was the *mateiro*, the local person with many years' knowledge and experience of the forest. I will return to explore the significance of these different roles in the last chapters.

In describing measurement as a transformation, what I am trying to capture is the moment at which 'nature', or the property of nature that the LBA researchers are interested in measuring, becomes *something else*. Thus I am suggesting an anthropological theory of measurement of the LBA researchers and technicians, that focuses on these transformative moments. This is not the same as a scientific theory of measurement. For example, there is in basic statistics a difference between "continuous" and "discrete" scales. Whereas one can have continuous measurements of degrees Celsius, such that theoretically one can measure to an infinite number of decimal points (that is, the units are infinitely divisible), one cannot have a continuous measurement of people in a certain area – one can only have a headcount. That is, people are indivisible and discrete units; there is no such thing in this latter case as a decimal point.[47] This was also apparent in the slight slippage between terms that occurred with the technicians and electronics specialists I spoke to. They sometimes spoke as if numbers are analogue when contrasted with "binary" (1s and 0s), but when I enquired, they told me that binary is digital, though a digital watch that displays numbers is also digital. However, the latter of these is the sense of the contrast that I wish to employ here. Numbers are also digital because *any scale is by definition discrete*. The conversion of the current into binary digital format, and the conversion of the current into numbers, effectively performs the same operation of transforming a continuous entity into a discretized one. Thus discretizing the current into a binary code is just another level of the same scaling process that would discretize the line drawn by the stylus in the original hygrometer into separate measurements. Because one has to use scales with discrete points in order to make meaning in measurement, (even if the points are very close together making it a very precise scale), my argument therefore is that the "something", (which is "data" in numerical or binary form) that is imparted or created by measurement is thereby inherently discrete. I propose that this property is important in order to understand what measurement does.

Discretization presented itself to me as a crucial aspect of the process of collecting data in general only whilst I was going over my notes on my return from the field.

[47] Although, of course, this does depend on a particular notion of a person - anthropologists have described how in Melanesia for example, persons might not be thought of as in the singular or plural but rather "fractally", a notion that in turn has profound implications for the indigenous mathematics (Wagner 1991, Mimica 1988). Thomas Crump has made a detailed comparative investigation of the mathematics of different cultures (Crump 1990).

Although I had several conversations at different times with different researchers as to what a "parcel" or "packet" of air might be, or how one decided on the limits of one's research area, how to fraction air samples into different molecular quantum frequencies using laser technology, or indeed about the difference between digital and analogue signals, it was not until some time afterwards that I realized what an abiding motif the problem of discretization had been. There was one particular conversation that sticks in my mind concerning this, which I had with one researcher about the method by which "flux" (*fluxo*) is calculated, known as the eddy covariance method. It is the same method if the flux is of carbon or if it is energy flux, even though they require different instruments. One of the most important features of this method is the notion of a "footprint". The footprint is the area that the meteorological tower is receiving information from. As the researcher told me, who works with energy rather than carbon, there are mathematical models that calculate this "angle of attack", as he put it, given the prevailing direction of the wind, and the height of the tower.[48] The basic idea is that for every metre of height, the tower is influenced by 100 metres of horizontal area in the direction of the prevailing wind, which "carries information to the sensor". According to this model, you can tell whether a particular parcel of air could arrive at the tower or not – only parcels of air within the footprint can be contributing to the information it is collecting. Each tower is therefore surrounded by a footprint, the limits of which mark the limits of the information the tower provides in the form of data.

I was however, confused at how one might separate the air into parcels. "But what counts as a parcel?" I asked. "Well", the researcher said, "it's in the moment of measurement, what was measured." "So it is defined by time?" I suggest. "Actually, it is magnitude" (*grandeza*) the researcher said, after a pause. "The more you measure, the smaller chance you have of error. So, the sensor picks up all the measurements (*medidas*) which are passing by it (...) The footprint will give you the point up to

[48] This researcher also provided me with a power-point presentation that he had on his hard-drive that explained how the eddy covariance method functions: "Introduction to the Eddy Covariance Method: General Guidelines and Conventional Work Flow" by G.Burba and D. Andersen, produced by the manufacturer Licor BioSciences, and available at: http://www.google.co.uk/url?sa=t&rct=j&q=&esrc=s&source=web&cd=3&ved=0CDcQFjAC&url=http%3A%2F%2Fwww.instrumentalia.com.ar%2Fpdf%2FInvernadero.pdf&ei=m5U_ULGSAbDY0QXxn4DABg&usg=AFQjCNHCLSO34L_s2YCxN-k_J79OfZ5E1Q&sig2=wS80U-CeaQbREdAZ6jF0vw Accessed August 2012

which your tower is influenced by the region around it. After that line, your tower isn't influenced anymore. It doesn't pick up anything else (*não pega mais*)." "So is it the instruments that determine the size of the footprint?" I asked, suggesting it was their sensitivity that defined how far they could 'see'. "Exactly" said the researcher, but then he continued, "the air is continuous. But the measurement is not continuous. The measurement is broken (*quebrada*). Measurement does this. Ten times per second, the instrument measures the parcel. It doesn't measure continuously." I realized then that it is measurement that actively "breaks up" the air into parcels, or the world into the "footprints" of towers. Further, I would suggest that the eddy covariance method is simply a complex version of a mundane operation. As Paul Edwards notes of reading the temperature from a mercury thermometer, "[Y]our act of reading the thermometer transforms a continuous, infinitely variable analogue quantity into a discrete number, such as 75° F: a set of digits, which vary discretely or discontinuously" (2010: 105). It is the instruments that perform this operation at the LBA, not only the ones that create the flux data, but every single instrument on the tower. These instruments "break" the properties in the world into discrete sections in order to produce data; any scale can be an instrument that measures because it performs exactly this operation.

An important correlate of the emphasis I am placing on this interpretation of measurement is that the data that the instruments produce is not a reduced, abstracted or elemental version of the world, in the way a Law of Nature might be imagined to be. Thus my description might add another (contemporary, digital) dimension to Thomas Hankins and Robert Silverman's description of 17th and 18th century instruments. They explain that whereas analogy was the fundamental principle in 17th century natural *magic*, "analytics" took over in the 17th and 18th century natural *philosophy*.[49] Whereas in the former, the natural magicians' instruments (which were the envy of the natural philosophers) *reproduced* nature, as I have previously discussed, in the latter, analytical instruments *took nature apart* "to reveal the rules by

[49] Hankins and Silverman (1995) posit a very interesting historical continuity between natural magic, and natural philosophy, if one takes instruments as the focus of the enquiry. Both practitioners used the same instruments for different ends. Whereas Anasthasius Kircher, a natural magician, preferred to convey what he knew in the form of allegory, the Enlightenment scientists, Hankins and Silverman tell us, preferred tables and graphs, that abstracted nature. Graphs are a pivotal form in the LBA science, and I will come onto their importance for the LBA data practices in the fourth chapter.

which it operated" (Hankins and Silverman 1995: 114). This does not map neatly onto what my analysis of instruments demonstrates. In Hankin and Silverman's analysis, 'analogy' seems to imply continuity between the world and the act of knowing it, and 'analysis' seems to imply a discontinuity, such that the act of knowing the world deconstructs or reduces it in the process. The act of producing data that the LBA researchers are engaged in, mediated by the measurements their instruments make is autonomous of either of these two positions. Nature is not taken apart through being measured, nor is it created by it; it is, rather, somehow *recomposed* as data. Analysis happens only afterwards, on the *data*. Digitization or discretization is a process of recomposition, rather than representation or reductionism. It is important to distinguish digital data as a specific product of scientific practice in this way, because it takes into account the constitutive role that uncertainty plays in the distinction between data and nature, to which I now turn.

Error and Uncertainty

A crucial aspect of this recomposition is the role played by error and uncertainty. Returning to an earlier citation, one researcher had told me that each time you convert the analogue into the digital, information is lost (see also Edwards 2010: 106). The problem, L said, is that when one converts a continuously varying signal into a discrete code, the "in-between" is cut away. One cannot transform a curve into a series of points without losing some of the curve.[50] Returning to the example of reading a mercury thermometer, "the fluid is almost never precisely on the line marked on the scale, but instead slightly above or below it. By setting aside that fact, *you* make the observation digital" (Edwards 2010: 105). That is to say, fixing the point, at any one scale, will always be incomplete. The transformation of analogue into digital has a built-in loss.

This incompleteness permeates the act of measuring natural properties out in the world, as I observed when talking to researchers about the eddy covariance method:

[50] This is the same problem, in fact, as the climate modelers I conducted interviews with at the end of my fieldwork encounter when they have to try to mathematize the weather, which is understood to be made up by non-linear processes: they have to try to discretize what is essentially continuous.

"when you use eddy covariance, you see that the balance doesn't close. What goes in is meant to come out, but it's as if something is lost, as if a bit of energy disappears." This is the case even in homogeneous terrain, such as pastureland, which is considered to be relatively easy to quantify because the footprint is easy to calculate, due to the fact that the terrain does not vary and the air can pass smoothly over it. Even so, the eddy covariance method cannot measure all of the energy that is presumed to be in the system - there is always a gap. It manages to capture around 95%, one researcher said "which in this case is pretty good". But in the Amazon forest, this gap increases in size dramatically. The Amazon is completely heterogeneous in terms of terrain and vegetation, and the vortices and eddies that pass by the tower therefore are of all different sizes and speeds as a result of the irregularity of the surface it is passing over. Many of these vortices, it is surmised, are missed because a single frequency of measurement – for example measuring 10 times per second, as the systems at the LBA do – cannot take account of all the different speeds and sizes. It captures some, but not all, of the variety of the vortices. This means that measurement can never be complete. As another researcher told me, "the tower data is not going to be faithful (*fiel*), but it is enough to validate the capacity to simulate fluxes. No-one works with perfection, but with approximation (...) In general, observed data do not represent the real world." The loss of information as one tries to discretize a continuous signal in an instrument is analogous to the loss as one tries to discretize the air by "breaking" it with measurement. There will always be something missing.

However, I would like to suggest that the impossibility of effecting a complete transformation is a constitutive aspect of measurement, and takes the form of uncertainty. The extent to which this is so was made very apparent to me when I spent some time with metrologists[51] at the Laboratory of Meteorological Instrumentation (LIM) at Brazil's Space Institute in São Paulo state, in the southeast of Brazil. LIM is responsible for the maintenance of the scientific instruments used to collect data on meteorological variables throughout Brazil. These instruments are either part of a network of long-term meteorological data-collecting stations spread across the country, like the LBA, or belong to specific and normally shorter-term projects linked

[51] Metrology is the science of measurement.

to other institutions. The LBA would often send their instruments to LIM to be calibrated; during the time I was there, several people in the micro group had several long conversations with the technicians at LIM, after sending different instruments to be calibrated. It was also the specialist technicians from LIM that came to give a course in electronics to the LBA researchers and technicians. Alongside mechanical maintenance, LIM is also in the process of establishing Brazil's first accredited calibration laboratory (which I will refer to as "CL") for scientific projects.

Calibration is the process by which the error and uncertainty in an instrument is calculated precisely according to – that is, by comparing it to – the universal standard for the units it measures in.[52] There are already several such laboratories for the Industry sector in Brazil, all accredited by the Brazilian Institute for Metrology (INMETRO), but none exist to cater for the very specific calibration demands of scientific instruments that work in specific derived units, such as watts/m-² (radiation) and require the simulation of specific environmental conditions. During the time I was there, the CL at LIM was in the process of being set up in order to rectify this, and is run by husband and wife team, P and M, who have both previously worked in metrology in the Industry sector in another town in Sao Paulo state. In order to be able to measure environmental variables as the LBA does, the instruments have to be calibrated to a standard. This is akin to the act of "zeroing" that I discussed in the first chapter - the instrument needs to be *given* a scale by which to measure. I will briefly here describe what a metrological standard is, and what calibration involves, in order to explain how error and uncertainty are constitutive of measurement. As M told me, "nothing is pure; the minute you start measuring, that's it, there's no way round it".

Within metrology generally, there are two different types of absolute standard. The USA's National Institute of Standards and Technologies (NIST), one of the leading national institutes, works with artefact standards and intrinsic standards (O'Connell 1993: 152). An artefact standard is for example, the kilogramme, which resides in the

[52] The process of "calibration" that I briefly talked about in the last chapter was called "verification" by the LIM metrologists. That is to say, they did not consider it to be a proper calibration merely to "zero" the instrument, without being in controlled laboratory conditions and having an accredited standard to perform the calibration with. This was almost invariably impossible in the forest, and in the LBA in Manaus, which is why instruments were sent to the CL at LIM as often as possible.

International Bureau of Weights and Measures (BIPM) in Paris. Artefact standards exist as material instantiations of themselves, and metrological practice consists in ensuring that any kilogramme measured can be traced back directly to this standard kilogramme. This is called "traceability" (*rastreabilidade*), and these traces exist in the case of the metrologists that I worked with, as stamps of accreditation that are stuck onto any instrument that is calibrated in this way. These stamps effectively say that the instrument is now part of "a documented unbroken chain of calibrations".[53] One of these stamps is gained either by being visited by a "touring" standard, which is periodically taken around the various smaller laboratories (as used to happen with 'the volt' in the USA), or by sending equipment to National or International metrological laboratories to have them calibrated to the standard there (O'Connell 1993). This is what the CL at LIM do, for example, in the case of radiometers. Their radiometers, that they use to calibrate the radiometers that are sent to them from other parts of Brazil, are sent periodically to Davos in Switzerland to be calibrated to the absolute radiometric standard unit that is generated there, up in the Swiss Alps.

The other form of metrological standard is the intrinsic standard. These are physics experiments that labs can perform themselves, to "create the volt, second, ohm or various temperature points right on their premises" (O'Connell 1993: 153). The volt, metre, and ohm were formerly all artefact standards, but have been converted into intrinsic ones[54]. As I was conducting fieldwork in a metrological laboratory that specializes in the calibration of meteorological instruments, the units in question are all derived units. Nevertheless, the logic of traceability, (as a chain of marks of accreditation leading back to a standard), is the same irrespective. The CL at LIM has a variety of standard instruments, ranging from standards that they themselves are testing to accredited climatic chambers that they use to calibrate thermometers.

Calibration is the act of comparing the instrument in question to the standard for the units the instrument measures in. The calibration of any sensor at the CL at LIM

[53] A definition taken from the BIPM vocabulary of metrological terms, which I was directed to by my informants at LIM: http://www.bipm.org/en/bipm/calibrations/traceability.html, Accessed March 2012.

[54] As Joseph O'Connell (1993) suggests, this implies a major conceptual systemic shift, in the sense that direct, unmediated contact with the absolute standard is now available to all. O'Connell likens it to a Calvinist reformation, especially as one does not "confess" or "redeem" (that is, calibrate) intrinsic standards, as they are not thought to "drift".

requires at least: the standard you will be comparing it to; some sort of computer to register the data (normally a little computer called a datalogger); and the sensor in question. It also requires some way of controlling, or ensuring variation in, the phenomenon that the sensor measures. For temperature, P and M have a climatic chamber, (which is accredited by INMETRO), in which one can vary the temperature at different points. The thermometer to be calibrated is put in the thermal chamber along with the standard, 'traceable' thermometer. Different parts of the chamber can be heated to different temperatures, depending on the range of your target environment (measuring in the Arctic does not have the same range as measuring in the Amazon – so you can calibrate just ranges of temperatures, rather than all temperatures). What is remarkable about calibration however is that at every stage, uncertainty and error enter into the system. The metrologists are, I suggest, specialists not only in measurement, but in uncertainty.

First of all, the thermal chamber has a certain level of uncertainty – it can only be precise concerning the degrees to which it heats itself to a certain order of magnitude; likewise the standard thermometer has an inbuilt uncertainty. These uncertainties accompany the chamber and the thermometer and are ineluctable. The datalogger also has a degree of uncertainty due to the conversion of analogue signals into digital, and in fact P and M had to join forces with their old metrological laboratory to find some way to calibrate their own dataloggers, as they could not find anyone else in Brazil who could do so. Secondly, there are physical problems to be taken into account, such as that of hysteresis. Hysteresis is a non-linear effect that makes a thermometer heat up differently from the way that it cools down - where increasing temperature has a different mechanical effect on the thermometer compared to decreasing temperature[55]. As P tells me, "mechanics, electronics, physics, they all vary, there is always a deterioration". The temperature and the humidity at which the standard you are using

[55] If the ambient temperature increases by 1.05°C, the thermometer reads an increase on 1°C; but when the ambient temperature drops by 1.05°C, the same thermometer records a drop of 1.1°C. This means that every temperature point that one is testing against needs to be done from zero – the thermometers must be cooled back down to zero before being heated again to the different temperature points. If one were to go straight from 5 degrees to 10 degrees, you would end up with a different set of temperatures on the way up and on the way down. I was originally told about hysteresis during an interview, and subsequently supplemented that description using the site: http://wattsupwiththat.com/2011/01/22/the-metrology-of-thermometers/ Accessed 4th September 2011.

was calibrated must also be taken into account – if the calibration in question is performed at a different temperature and relative humidity than the one that the standard was calibrated in, that must also be included as "embedding local error" (*embutindo erro local*). Even the longitude and latitude at which the standard was calibrated is important to include in the calibration of subsequent instruments.

Finally, there are some sources of uncertainty that are only discovered through experience. Pluviometers (that measure rainfall intensity and volume), for example, work via a system called the "tipping bucket". The Micro group at the LBA had been having particular problems with the pluviometers on the towers, and had sent several to LIM to be calibrated. When I asked about these, M explained that when the little buckets in them are full to a known volume, the weight of the water causes them to tip and touch a metal pad below them. This completes a circuit that causes a current to run through the sensor, registering data. But in different intensities of rainfall, this can fail to work sufficiently; either the rain is too intense for the bucket to register enough 'tips', or the rain is so light that it never quite manages to fill the buckets. That is to say, this type of pluviometer works with one level of uncertainty in heavy rainfall, and another one in light rainfall – and presumably in all the sorts of rainfall in between (Braga and Fernandes 2007). Thus when one calibrates a tipping bucket pluviometer, one has to be sure to recreate all the environmental conditions, simulating intensity of the rainfall as well as volume of water.

I asked P one day if anyone calibrates the standard. No-one does, I was told. Then she paused and explained that the standard is "normally the materialized form of the thing". The notion of a "material form of a thing" implies that the "thing" itself – in this case the standard unit – is still as such, *immaterial* – that is, not present. The absolute standard is, in this sense, at one step removed from the "absolute absolute".[56] The chain of traceability that characterizes the science of measurement could thus be seen as having an infinite regress on both ends in different ways. One is endless repetition of units (in the form of measurements), the other is eternal displacement (the absolute unit lying always beyond its own materialization). Thus the singular

[56] An understanding that I think clearly resonates with the citation with which I ended the last chapter - "we know we are not looking at the really really real (*o real real real*). But it's what we understand as real".

standard is the universal reference at the same time as being itself part of its own universal expression. There is inherent in the idea of a standard the fact that it is always incomplete. If the standard is the materialized form of the thing, what lies in between the thing (unit) and its materialized form is the necessary and inescapable uncertainty that accompanies *all* metrological activity. "Everything is going to interfere", P tells me; and Q, the head of LIM, also impresses on me "you're always going to have a certain error which is uncertainty, the uncertainty of calibration, you can't escape it."

The result of the calibration is a certificate of calibration, which reflects this. It has two different sets of values on it. One is "error", and the other is "uncertainty". These two are importantly related, but can be separately defined thus: Uncertainty is the range within which the "true value" is asserted to lie – so as we have seen, uncertainty is the error embedded into the process of calibration by your calibration system itself. This uncertainty is defined as a percentage. It means that the standard value itself has a range around it – or rather, is itself a range. It is often calculated using standard deviation – so, for example, if you are calibrating the temperature point 15°C, using a standard accredited thermometer, and your personal thermometer reads 14.9, 14.6, 14.8, 14.5 each time you repeat the measurement, then you have a high standard deviation – how much the numbers are spread out - which is an indicator of a high uncertainty. If, on the other hand, it reads 14.9, 14.9, 14.9, 14.9, 14.9 then it is systematically wrong or in error, but has very little deviance and therefore a much lower uncertainty. This uncertainty then would be added to the uncertainty that surrounds the standard, along with all the other types of uncertainty, in order to give the total uncertainty of the system. Error, on the other hand, is simply this *difference* between the "true value" and the measured value – the extent to which your instrument being calibrated errs from the standard. The error is what you can try to remove from your system, by correcting for it mathematically, or, if you can isolate the cause, mechanically. This is what calibration allows for. The researchers who receive the instruments back from the CL at LIM then are expected to correct their own measurements according to the coefficients provided them. They are also provided with the uncertainty percentage that is built in to their system. This cannot be removed. As such, uncertainty is a strange measurement: it singularizes how much one *cannot* know what one *does* know.

This uncertainty, lying between the unit and its materialization, has the property of being both in the measurement process and in the world being measured. In metrological professional vocabulary this is captured neatly by the phrase "the conventional true value of a quantity". This is "a value of a quantity which, for a given purpose, may be substituted for the true value", as that true value is *by definition* unknown (Mallard 1998: 571). And this uncertainty is therefore not so much an epistemological error, but an inevitable, and therefore constitutive, aspect of measurement. As M told me: "look, for [electrical] frequency, the primary standard today is the oscillation of caesium – how the caesium atom oscillates – that's the standard today. But it's got an error already of 10 to the minus 20 – and it's the *standard*. Because if you put a lot of caesiums together, this is the difference between them – there's no way of improving this, it's nature. The error comes from nature." Thus from the perspective of the metrologists, error and uncertainty accompany every act of measurement.

Metrologists study the science of measurement, and so they are keen to seek out the most precise ways to talk about uncertainty. They act as "support" (*apoio*) for the LBA researchers and Micro team. These researchers may not be so aware of the uncertainty that surrounds their every act of measurement, and this causes a great deal of worry for the metrologists. As the head of LIM explained, "my biggest worry in all of this is the following: you calibrate everything and so on, and reach the conclusion that your thermometer has an uncertainty of 0.2 degrees and you start collecting data. Only, the researcher [whose thermometer it is] who gets this data, does he use this information? He doesn't use it at all. If it says 25.3°, he's going to use 25.3°, he's not going to take that 0.2 into consideration. He's going to use it as if this value is exact (…) So even if you calculate the uncertainty of the equipment, it's not going to be used in the subsequent proceedings".[57] Whereas the researchers care about the content

[57] M told me in a subsequent interview that this was not the case for researchers from the USA and Europe, who were more accustomed to factoring uncertainty and error into their research. He told me that a French researcher had come to conduct an experiment and had shocked his Brazilian colleagues by asking where he could calibrate his instruments before installing them. "You just don't have that here," he told me ruefully; not only because Brazilian researchers are not in the habit of a more militant regime of calibration, but also because there is only one laboratory in the whole of Brazil where these instruments can be calibrated.

of the data, they are sometimes not aware, or do not take into account, the data's uncertainty. But irrespective of this I would still, with the head of LIM, argue that uncertainty and error are essential to understanding measurement. Indeed, researchers themselves are also convinced of this: P told me that one researcher who had brought his instruments to be calibrated had been so surprised at the difference it made that he vowed to continue with calibrating his instruments as often as possible. Thus I am quite happy stating this not as a question of "false consciousness", but because this is how the researchers themselves would consider it - that the uncertainty is there whether they knew it or not at the time. The question then is for those researchers (who are not metrologists) how important it is. They may not need the most precise measurements for their particular experiment. In the case of the LBA, the centrality of uncertainty was clear in the work of the Micro team members I worked with. The Micro team is very aware of the uncertainty that surrounds the data they produce because they often have to liaise with the metrologists and technicians at LIM in order to keep the towers and the instruments on them up and running. However they do not have reason to explicitly deal with this uncertainty aside from writing calibration factors into datalogger programmes, or correcting data using it, because they only process the data, they generally do not analyze it. Even so, as they often told me, they know that uncertainty is inevitable.

The Continuous and the Discrete

Thus far, I have tried to describe what measurement entails, generally and in the more specific setting of the Amazon. I presented an image of the progressive "breaking" of continuous processes in the world. This transformation - that I have glossed as the transformation of the analogue into the digital - is composed itself of two processes: the production of an analogy, and the transformation of that analogy into a discretized code. The qualification to this is that in this process, at whatever scale, there is always a loss of information, that is called "uncertainty". The size of that loss can be elicited through calibration, but the fact of it is inherent in the act of measurement, or at least in the act of making meaning that is inherent in measurement. It is also more apparent in an Amazonian setting than in other places where measurements might be made, because the Amazon is great deal more heterogeneous than other environments.

In terms of relevant studies for a discussion of measurement in this setting, perhaps most empirically interesting to serve as a comparison is Bruno Latour's study of pedologists (soil scientists) in Brazil. Although perhaps better know for his work on laboratory science, in a chapter of his book *Pandora's Hope* (1999a) Latour describes how he ventured out into the Amazonian forest to accompany a pedology field trip that a group of scientists embark on in order to investigate the savannah-forest interface. He describes in detail the way that the pedologists measure and characterize the soil samples they collect, paying close attention to the instruments they used, as I have done in my description. The result that Latour (self-consciously) brings back from his trip is that the ontological gap between "world" and "word", and the concomitant emphasis and interest in "correspondence" between the two, are misconstrued. He suggests instead a series of transformations, so that the world – in this case the soil from the Amazon – serves as "matter" for the instrument that will covert it into a number or "sign", which will then serve as "matter" for the next transformation by a different apparatus into a "sign" – all the way along the chain until we get to publication, at which point we stop, but one imagines only for sake of brevity. At the same time, "something" is maintained during this series of transformations, and kept constant. This something seems to be the meaning given to the world, or its "truth value" (ibid.: 69). In a nutshell, what Latour demonstrates is that the gap being investigated is repeated at each transformation; so the gap in the philosophy of language between "world" and "word" is the same as that between the forest in Brazil and the publication in the library in France, which is then reduced to a series of gaps that replicate that lacuna between "real" and "represented" at every level, in fractal manner. And it is across these gaps that the pedologists "conveyed the meaning of each phenomenon by making matter cross the gap that separated it from form” (ibid.: 57).

Latour thus proffers two models - the first is that of a gap itself made of gaps; the other is of a movement across these gaps that is "indirect, crosswise and crablike" (ibid.: 64), that in fact does away with the idea of there being a gap altogether. He has also elsewhere referred to this movement as "articulation" (Latour 1999a, 2004a), in contradistinction to "correspondence" (Latour 1999a: 142-143). The direction of this movement is given through a "dialectic of gain and loss": the transformation of the soil into a diagram and then a set of numbers implies a progressive loss of, amongst other things, materiality and particularity; and a gain of, amongst other things, textuality and "relative universality" (ibid: 70-71). So in this movement, we have a progression towards abstraction that substantiates his other, and perhaps more famous, dynamic, which is the construction of a fact through the black-boxing effect of an actor-network (Latour 1987: 103-144). Latour maintains then a double, and paradoxical, position; the gap between the world and words does exist, and it does not.[58] Although Latour mentions the contradictory nature of what he is describing several times (ibid: 70, 77), he does not do anything in the text towards explaining their apparent incongruity, nor does he directly address the use he makes of the very relation he is trying empirically to do away with.

I would like to suggest that given the ethnographic descriptions presented in this chapter, one way in which measurement might be approached analytically is not in fact as a progressive process of abstraction from "world" into "word", but along the perhaps unlikely path provided by anthropological studies of origin myths. This is because these myths focus exactly on the transition from the continuous to the discrete. I do not intend, nor am I able to provide, a comprehensive point-by-point comparison between experimental measurement and mythological thinking and discourse. Rather, I want to take that discourse as a starting point to offer an alternative model for understanding the production of meaning in the case of the LBA, specifically meaning as produced through measurement. Based on a mythic narrative by the Yanomami leader Davi Kopenawa, anthropologist Eduardo Viveiros de Castro suggests an analogy between Amerindian shamanism, (the indigenous practice that is inherently associated with myth, see Kelly 2012: 8) and 'White' writing. It might be that this analogy is even more apt for the quantitative rather than the qualitative, given that the former is expressly concerned with discretization, as I hope to have demonstrated.[59]

Although the Lévi-Straussian distinction between the continuous and the discrete ("the celebrated transition from the 'continuous' to the 'discrete' which constitutes the meta-mytheme of the structuralist cosmology", Viveiros de Castro 2007: 158), has

[58] “[N]o step – except one – resembles the one that precedes it, yet in the end, when I read the field report, I am indeed holding in my hands the forest of Boa Vista” (ibid: 61). I return to this doubleness in Latour’s thinking throughout the thesis in different ways.
[59] See also Bowker 2008: 201-220.

been substantially elaborated on by other anthropologists (notably Eduardo Viveiros de Castro), particularly by applying the insights from Roy Wagner's semiotic anthropology, it serves my present purpose to merely shift attention away from the relation between the world and its measurement as one of representation. My suggestion is that the "gap" between world and words, that Latour is intent on removing, is in fact the gap between the continuous and the discrete and is continuously created anew through measurement. This means one *could* do away with the real-representational relation, as Latour wants to - but *keep* the gap. What is on either side of it is not the real world and representations of that world, but a continuous nature in which there are no gaps, and a discrete world of data in which there are. Thus what Latour's contradiction provides us with is perhaps the beginning of a discussion not about the ways in which to collapse the gap between the "world" and "word", but another way to think about that gap altogether. When Latour presents us with an image of a gap that is itself made of gaps, and yet one that can be crossed nonetheless, I suggest that it might be important to take the "gap" seriously.[60]

Following Antônio José Kelly's analysis of Amazonian myth, that draws heavily on the work of Roy Wagner, the gap may not be, in fact, between the "real" and the "represented", but between a spatio-temporal plane in which that distinction does not exist, and one in which it does: "meaning proceeds from metaphor to denotation; from symbols that stand for themselves (Wagner 1981) to a distinction between symbol and referent; from analogy to its deterioration into homology... none of the first terms of these pairs disappears with the appearance of their contradictory complements, they are their origin and in a way immortal" (Kelly 2012: 8-9). In Amerindian mythic discourse, the pre-cosmological world is characterized as one of undifferentiation, or as Viveiros de Castro puts it, a domain of "qualitative multiplicity" such that "everything seeps into everything else" and in which difference is internal, rather than external, to each mythic entity (Viveiros de Castro 2007: 157-158). In very general

[60] In broad terms, this is reminiscent of Zeno’s paradox. As mythologist Gregory Schrempp has pointed out, this propensity for “dividing up” that he situates in the Greek philosopher Zeno’s work may be the basic trope in Western cosmology (1992: 8). Dividing up the world as Zeno does results in “pairs of seemingly equally necessary but mutually exclusive propositions about the character of the universe (Schrempp 1992: 9). One of the particularly generative suggestions Schrempp makes is that logical paradox is not an anathema to thought, but constitutive of it – the fundamental way in which “reason ends up at odds with itself” (ibid: 9).

terms, it is from this primordial substrate - where transformation rather than essence constitutes existence - that the cosmos is created. This is done by the hero of the origin myth in one way or another introducing differentiation, which essentializes the different human and animal groups that today populate the world. Thus in one Yanomami myth that Kelly takes from anthropologist Bruce Albert, we are told that the mythic hero Omame made the Yanomami what they are by "putting an end to transformation" so that they no longer turned into "tapirs, armadillos, and red brockets" (Albert in Kelly 2012: 3). Although Kelly uses this story as a starting point from which to investigate some complex articulations of the figure-ground dialectics of specific Amerindian myths, here it is enough to take this basic point as suggestive of a means by which to understand another transformation of the continuous to the discrete, as I have suggested measurement is. Such a comparison would imply that "nature" here - the continuous - is a "symbol that stands for itself", as Wagner puts it (1986), but only in so much as it is a domain that is "at once proposition and resolution" (Wagner 1986: 11). The world is "given" not because it is not "made", but because asking whether it is given or made is an impossible and non-sensical question. This notion of a continuously variable world in which the putative relation between real and represented, or objects and subjects, does not exist, fits with the description of the world that I presented in the last chapter. In that chapter, I noted that the world that is cut away is not a hybrid, but a sort of relational mish-mash, or "noise"; a place where no meaning can be made at all. Nature here, then, is not so much "referent" or "reality", but the continuum or transformational substrate that in a sense pre-dates such a position. Taking the comparison between myth and measurement to its logical conclusion, it is this position, or rather relation, of "referent" or "real" that measurement continuously introduces. Nature is therefore not "real" until it is recomposed as measurement; but it is simultaneously as given as mythic time, for example, is.

To suggest that measurement might be seen in this way seems a rather grand claim for a very banal act - measuring the ambient temperature seems a long way from the creation of the world as we know it, which is what the origin myths being referred to tell of. But I think it is enlightening (and this incongruity itself may even be of note, in as much as it speaks to a capacity of everyday actions to invent the world anew). Although I do not wish to stretch the comparison any further concerning the relation between nature and a mythical time [61], I would like to dwell on the hypothesis that within the framework I have suggested, it is in the *data* (i.e. the discretized 'world' recomposed through measurement) that the relation between real and referent might be found, and not in the 'world' pre-measurement. The relation between representation and reality, and even subject and object, is shown to perhaps only even appear (if it does), *after* the act of discretizing, or singularizing, and does not inhere in the world/nature at all.

Light is thereby also shed on the importance of uncertainty in measurement, not as an afterthought but as constitutive of the world that measurement produces. In the process of transforming the continuous world into measurements that are data it might be said that what is introduced into that world *in order to do so* is uncertainty. Drawing on my description of the work of the metrologists at CL, it is relevant that uncertainty straddles the continuous (nature) and the discrete (data), or, in another sense, is the gap between them. In a world pre-measurement, as it were, there was no such thing as uncertainty: to paraphrase Latour, 'Nature' had never been asked, before Western science came along, whether or not it was the "true" or "real" one (Latour 2009: 472); measurement cuts through this "relaxed attitude towards truth" (Latour ibid: 472) by creating a specific transformational relation between nature and data (the act of knowing). Uncertainty, in my description of measurement as the metrologists I worked with tell it, inheres in the act of knowing the world, but once

[61] I am not suggesting that the world of the LBA researchers is the same as that of Amazonian "mythic time". There are of course many important differences that deserve to be explored, not least of which those intimated by Wagner concerning the distinction between ecological and indigenous Daribi modes of meaning, in which they appear as the asymmetrical inverse of each other (Wagner 1977). In fact, to a certain extent Wagner's reading of science implies the opposite of what I am here suggesting, as he argues that ecological science is concerned with "relating the perceptibly differentiated" rather than, as in the Daribi case, "differentiating the perceptibly relational" (1977: 391). However, as my ethnography suggests, the LBA researchers do understand the world to be continuous, and it is the relation between their knowledge of the world and the world that is the question. I would argue that in fact, ecological science, as practiced and understood by those I worked with at the LBA, operates at more than two levels of the "real" and the "symbol". It accommodates the idea of what nature is in contrast to an idea of what a representation of nature is, as for example between the climate and a model of the climate. But there is also the added complexity that what nature is and what a certain representation of nature is can also theoretically correspond. Thus there is the world, a model of the world - and observational data, which is not the same as either. Although I am unable to explore this triad fully here, what I am pointing to in this chapter is the difference between data and the world, whereby it is the *data* that in fact that takes the position of being the "perceptibly differentiated" substrate of Wagner's analysis of science.

there, becomes a constitutive aspect of natural knowledge itself. Uncertainty is in the standard, and in the measurements made using that standard. The world, or truth, is always at one step removed. Although Alexander Mallard refers to this as the "social" and "natural" character of precise measurement, such that metrology consists in what he calls a "pragmatics of approximation" (1995: 594), I suggest that this indicates that uncertainty is integral to meaning exactly because it is neither approximation nor distancing but the gap that permits both. The uncertainty that surrounds every measurement refers to the inescapable difference - the gap - between discrete entities (caesium atoms, for example); this is "in nature", I was told. But these atoms are only in the world in discrete form because they have been measured; it is their measurement that introduces the uncertainty between them. Uncertainty is also simultaneously therefore what lies between the caesium atoms and the measurement of them - between the world, and the numerical recomposition we have of that world in the form of the data. From one perspective therefore this difference is the gap between discrete entities - i.e. that which makes them what they are; from another, it is the shadow of the continuity that has been removed - i.e. that which makes them unknowable, undistinguishable, uncertain. Uncertainty thus constitutes measurement by being the index of the discrete act that measurement is, at the same time as working against measurement by ensuring that the "continuous" is still somewhere in the "discrete" - by turning discrete measurements into ranges of deviation. Uncertainty captures both aspects of what a gap might be - what has been taken out, and what permits two discrete entities to be joined. Uncertainty thus also captures both aspects of what knowledge is. It is itself a measurement of how much of what is known is not in fact knowable. [62]

[62] This could be elaborated on further by comparing this simple analysis of what uncertainty in the form of a gap is by comparing it with the way that Viveiros de Castro explicates the different notions of ‘gap’ in pre-cosmological and everyday time: "I just stated that pre-cosmological differences are infinite and internal in contrast to the external finite differences between species. Here I am referring to the fact that the actants of origin myths are defined by their intrinsic capacity to be something else; in this sense, each mythic being differs infinitely from itself, given that it is posited by mythic discourse only to be substituted, that is, transformed" (Viveiros de Castro 2007: 158). Thus pre-cosmological differences are intensive and internal to mythical entitites - they are not different to each other, but rather different to *themselves*. In contrast, the process of transforming the continuous into the discrete necessitates the introduction of extensive differences *between* entities, crystallizing "molar blocks of infinite internal identity...separated by quantifiable and measurable external intervals (the differences between species are finite systems of correlation, proportion and

I would like to suggest that even if mythic though as I have briefly presented it here does not present itself as an immediate ethnographically-sourced notion by which my informants understand their actions, it fits much more snugly with how they do so than conventional dichotomous models of reality and representation. Not only does the importance of discretization in science resonate with its cosmological importance in mythic thought, but how my informants would speak to me about the data and its relation to nature (as well as the means by which subsequently data was treated), made me aware that it might be important to separate the idea of nature from that of reality. At the very least, then, what making a comparison between anthropological analysis of myth and scientific measurement (perhaps even quantification in general) permits is the possibility that there might be more terms to account for than expected: nature, uncertainty, and reality/representation. This idea is taken up again in the next two chapters.

Conclusion

My aim in this chapter has been to try to examine measurement not as a tool of social control and standardization, but as a basic and creative semiotic substrate of scientific knowledge, taking as the focus the work it does in the mediation between nature and data. As such, I have described measurement as composed of two stages. In the first, analogues of the world are created, and in the second, these analogues are discretized to give digital data. What this suggests is that the relation between the world being measured and the measurement is not accurately described as one between nature as "real" and "given", and culture as "abstracted" and "made". The data is not an abstract representation of nature, but a recomposition of it. Given the importance of this process of discretization not only for the production of the LBA data but for the production of scientific meaning more widely via measurement, I suggested that an

permutation of characters of the same type or order" (ibid). This is a world where "each being is only what it is, and is only what it is by not being what it is not" (ibid: 159). Thus the gap that is introduced between entities in the creation of the cosmos is that gap that is understood to define what those entities are, it is "external" or lies between them; the gaps between entities in pre-cosmological time are instead "internal". In my argument, I only discuss the different aspects of what an external gap might constitute once introduced through measurement.

interesting means to think about measurement is through anthropological analyses of the continuous and the discrete in Amerindian origin myths. The main benefit of this is to move away from a discussion that circles around a particular question of "representationalism", and in this way to open up the description of scientific practice in new ways that capture aspects of it that other frameworks might not.

One of these aspects is the important constitutive role played by uncertainty in measurement. My ethnography suggests that uncertainty is the key notion in metrological practice because it becomes, in the work of the metrologists, not a hindrance to knowledge but itself an unavoidable form of knowing. It is the knowledge of how much of what is known is in fact not known. Within the logic of discretization, uncertainty could be seen as what is introduced into the world by the act of measuring it, and as such can be seen as constituting the "gaps" between entities that measurement necessitates (the "loss" of information in the transformation of analogue to digital), as well as the gap between the natural world and the data world that measurement produces (the difference between the absolute standard and the world). Bruno Latour's suggestion to rethink the "ontological divide" between nature and culture, as evidenced in his analysis of the way meaning is made by pedologists and their instruments in the Brazilian Amazon, might therefore be better construed as the suggestion to rethink what a gap might be at all, rather than do away with it altogether.

This approach is only cursorily demonstrated in this chapter, and there remains more mileage, I believe, in the comparison between myth and measurement. However, another interesting ethnographic lead from the framework thus proposed moves towards the subject of the next chapter. In the case of the LBA (although not only), what the act of measurement produces is data that is known as "raw data" (*dados brutos*). This is data that is "data" at the same time as being "not yet data", because whilst it may have an inevitable uncertainty, it has yet to be "cleaned" of error. It is therefore simultaneously *both* error and data. In terms of the analysis of this chapter, the relation between data and error can be seen as a transformation of the representational relation between real and represented that the act of measurement introduces. The raw data is therefore the subject of the next chapter.

Chapter 3: Raw Data, Unique Ambiguities

Introduction:

The data that the LBA micrometeorology team collects and processes in fact has several different forms; it evolves and changes as it is transformed from raw data (*dados brutos*), into certified data (*dados certificados*). The Micro team refer to the "data" in the plural (*os dados*) all the time without necessarily having to refer to any specific stage, because the context makes clear which stage is being referred to – the question "did you collect the data?" implies one sort of data; "did you send the data to so-and-so?" implies another. This chapter is about the "raw data" (*dados brutos*), which is how the data is referred to before it has been "cleaned" (*limpar*), or put through a "quality control".

Raw data is the data that is collected directly from the datalogger, where it has been stored after being generated by the instruments. When I asked what raw data was, it was almost invariably explained as something still unfinished – "data that has not been cleaned yet" - or as one researcher put it, yet to be lapidated (*lapidar*). It is data with noise and errors in it, which contains "impossible" measurements, has gaps in it, and does not display the relations between variables that it needs to in order to be considered "clean" and ready for use by other researchers. Until the raw data goes through a process of "cleaning" (*limpar*), it is thus in some sense not yet data; but nor is it 'not' data, either.

In this chapter I will explore some of the properties exhibited by the raw data and investigate what collecting the data in the Amazon means for the LBA and for those researchers who travel to the Amazon expressly for that purpose. As I have already discussed, singularizing a relation sufficiently to measure it takes an enormous amount of effort. This is particularly the case in the Amazon forest. The difficulty with which the raw data is collected is what makes the raw data unique. However, the raw data is far from singular in meaning; it is "raw" exactly because it is potentially full of uncertainty and errors. One of the conditions of doing research in the Amazon is that many of the phenomena that the LBA researchers are trying to measure interfere in the process of measuring them. This is almost a reversal of the 'observer's paradox': it is not the presence of the observer that makes it impossible to know the forest, but the presence of the *forest* that does so. The heat that is measured heats up the thermometer; the rain destroys the pluviometer; the humidity wrecks the hygrometer. Most of these errors can be understood to be the relations the raw data has to the bits of the world that the researchers cannot control - such as lightning striking the tower, or unexpected power shortages. It is the isolation of the research sites that makes the raw data both valuable and unique, and ambiguous and uncertain. Thus the raw data is a scientific object that displays a high (and undesired) connectivity to the world, and therefore a semiotic multiplicity. It is not clear what the raw data contains within it when it is collected.

This description of raw data turns out to raise some important questions regarding actor-network theory (ANT). I suggest that the raw data is an entity that is in a certain sense 'pre-network'. It is defined by connections that expressly have to be removed in order for its meaning to stabilize - that is, for it to emerge as certified data. This implies that network building is not merely a case of accumulating associations, as ANT seems to suggest, but substituting one sort of association for another.

This chapter also proposes that as a 'pre-network' entity, raw data demonstrates some remarkable qualities. The ambiguous nature of the raw data brings into relief that way that the LBA researchers spend a lot of time dealing not with established and certain facts, but with uncertain objects. The raw data is interesting exactly because its ambiguity lies in the fact that it is neither error, nor data. It is not data, but not 'not data'. The chapter ends on a discussion of this liminal state.

'Uniquity', not Ubiquity; or, the Effort behind Uniqueness

This section will describe in more detail what it takes to collect raw data from the LBA towers in the Amazon, and what that effort results in. In order for any raw data to be collected at all, there is great deal of work that has to be done. Funds have to be obtained from foreign institutions or from the Brazilian government; collaborative links need to be established between institutions and sometimes governments, and cemented in the form of signed and certified documents. Work or research visas sometimes need to be negotiated that involve international laws, plane tickets, discussions. Data policies have to be established. A research site must be chosen, over a series of expeditions to different possible sites in the forest. For one project[63], for instance, one of the collaborators informed me that he and several other researchers had had to trek into the forest on several different occasions, in very hot and uncomfortable conditions. Once a research site has been chosen, the next stage is to ensure that a tower can indeed be built there, which means a team of engineers must be hired to 'probe' the area (*fazer sondagem*) to make sure the land is safe for building a tower. If the tower is being built in an area near to an indigenous reserve, as was the case with the tower at São Gabriel (SGC), the LBA should send someone to ensure the indigenous people, in the case of SGC represented by the Foundation of Indigenous Organizations of the Rio Negro (FOIRN), know and accept what they are doing. In the case of another project, there was a period of negotiation between the LBA and its collaborators, and the NGO that works with the local people on the Reserve the project site was to be constructed, as well as the local community leaders of those who live there, and the Park Ranger who is the governmental official representative. This resulted in the stipulation that for the towers to be built in the Reserve, the LBA must employ a certain number of people from the local communities to help construct them.

Once the research site is settled upon, the tower must arrive there safely. Different towers have different origins - the very first that were erected were made in Europe and shipped to Brazil, as is also the case for some of the towers being built for current projects. More recently, construction companies in São Paulo have been commissioned to make the towers according to specifications provided by the companies in Europe. The logistics team of the LBA often has to prepare several options for the towers, considering cost and the rapidity with which they can be ready. Once the tower is chosen, it must travel up the Amazon to Manaus, and then be transported to the research site. This depends a great deal on the time of year it is, as

[63] This project was a collaboration between LBA/INPA, UEA and the Max Planck Institute, and involved the construction of a 300 m tower in a remote biological reserve 330 km north of Manaus, and several smaller towers around it. The description here is drawn from several conversations I had with the members of the LBA, UEA and Max Planck who are or were involved in the different projects mentioned.

the river can either be too swollen from rains for a safe passage to be possible, or too low for any boats to pass at all. International and domestic transportation laws must also be negotiated here, and the tower can be held for several weeks at customs by the federal police. Once the tower has arrived at the research site, which involves sometimes months of transportation as well as time spent lugging it, part by part, into the forest, it then has to be built. This involves recruiting a team of people to work shifts in the forest over an extended period of time. Holes must be bored deep enough for the cement foundation of the tower to be laid, which involves carrying all the building materials necessary for this to the research site. The tower must be built, metre by metre, often between massive rainstorms that make it too dangerous to construct the upper levels for fear of lightning strikes.

The difficulty of access means that transporting the high-precision instruments that will be installed on the towers is a fraught process, and during the time I was conducting fieldwork, there was a great deal of discussion as to whether it was safe or not to transport the instruments to the one particularly isolated research site. The trail leading to the this research site became a sea of mud after the rain, with puddles of slick red slime reaching sometimes up to one's knees. The quad bikes the LBA technicians used could only just manage to keep upright on these trails, and any cargo being carried had to be securely lashed on to the back. A researcher from the Max Planck informed me that he had investigated the possibility of using a helicopter to transport the tower and the instruments, but it was going to be immensely expensive and it was impossible for the helicopter to land anywhere easily without clearing an area that might then affect the measurements being made. Every so often, the people who had been working on the construction of the project would arrive back at the LBA in Manaus on large trucks. They would be covered in mud and exhausted, and more often than not the truck would also deposit a broken quad bike or another piece of broken equipment that had not survived the trip.

Because of the length of time it takes to build a tower in the forest, a lodging or camp must be established near to the research site for those who are involved in constructing the tower, or subsequently, in maintaining the instruments or collecting data. These camps often need to have generators in order to charge GPS devices and radios, and allow for light during the night time. I visited several different camps at different research sites. The camp at ZF2, that was once merely a container, has now become a well-established lodging: a proper wooden building, with tables and chairs, bunk beds and access to energy and intermittent internet. The container is still there, but has become just one room within the building. On the other hand, there is no camp at the tower at São Gabriel, and when I accompanied an expedition there we had to construct a makeshift camp for the time that we stayed. At another site, a lean-to was erected, made of tree trunks lashed together with rope, on the site where a more permanent lodging would later be built. The initial lean-to was merely a structure to offer protection during a rainstorm and somewhere for around ten people to hang their hammocks. There was also a kitchen area, where someone was hired to take care of and prepare the food, which by the end of the week I stayed there became slightly rancid. There was also a satellite dish so that those that spent several months in the forest at a time could watch the television. However, it only received some channels if a potato was carefully balanced on the connection between it and the television.

What this all demonstrates is that building a tower is a particularly arduous task when the area that you want to build it in is desirable exactly because it is very hard to get to. Collecting the data that the instruments on the towers produce is also difficult for the same reasons. For every LBA tower, the process of collecting the raw data is different. Each tower is the responsibility of one member of the team. R looks after the tower K34. In R's case some of the data – the "low frequency" data of the meteorological variables - is transmitted via a radio link directly to the LBA receiver, and so is received in "real time", that is, the data that the datalogger stores is sent continuously to the LBA in Manaus, where it is stored on the micrometeorology system and displayed on a screen in the micrometeorology office. R downloads this raw data in order to process it. However, the data collectors still collect data from K34 and B34, for reasons of "back-up" I was told, and also because often, due to rain or lightening strikes, the real-time transmission can fail. The carbon flux data also still has to be manually collected. The real-time link is seen by everyone in the Micro group as a significant advance in terms of being able to do their job. Before the radio transmission, weekly visits to the tower to download the data from the dataloggers could show up an error that had occurred only slightly after the previous visit, and if it was a serious error, such as a problem with the battery on the tower or a machine losing its calibration or having mechanical problems, this would mean that all the data from the time of the last visit up until the present visit would have to be discarded, or simply did not exist. This would lead to gaps in the data, which were to be avoided if at all possible. With the live transmission, errors can be seen as soon as they appear, and someone can be sent to K34 to try to fix the problem, although this does depend on the link between the tower and the LBA in Manaus functioning. As a result, automated transmission actually means that the data collectors in Micro team spend more time in the field rather than less. Even though K34 was the most accessible of the LBA towers I visited, it took a lot of care and attention to keep it running.

In the case of the tower at ZF3, the raw data is stored on Compact Flash cards in the dataloggers, which requires a visit every two weeks, in order to swap the old data-full card for the new one. When I met him, the person responsible, T, was taking over from someone else who had been employed briefly at the LBA but had had to leave abruptly. T was learning how to download the data correctly and how to process it. The trail to ZF3 is very precarious, through now dis-used farmland, including a number of steep climbs which the 4x4s had a hard time managing in the rain. The car broke its axel once when T was on his way to ZF3 with a colleague, and it is only because they are both automobile fanatics that they knew how to patch it up enough to get to the motorway, where they were picked up and driven back to Manaus.

Raw data from the tower in São Gabriel, managed by J, used to arrive on a CD, by aeroplane. The tower there was very isolated, several kilometres along a dirt road which frequently floods, and then about 2km or so along a particularly precarious trail that has been cut through the forest. When I accompanied an expedition to the tower at São Gabriel in 2007, there were over 20 bridges made of just a few thin slippery logs tied together to wobble across in order to get to the tower. In 2007, there was a team in the city of São Gabriel consisting of 6 people who were employed to maintain the tower and collect and send the raw data to the LBA in Manaus. By 2010, there was only one, who collected data from the tower, and maintained the instruments in good working condition. He went to the tower whenever he could, and downloaded the data onto an interface, and then uploaded it to a CD, before putting it on the next flight to Manaus. This was easy for him as he also worked at São Gabriel's tiny airport. This CD was then collected from the airport by someone in the LBA Manaus logistics team, and delivered to J. Unfortunately, the tower at São Gabriel fell down in

June 2010, so J has not received any data since then. However, she is still working through the backlog of data that she inherited when she assumed the role of caring for the São Gabriel tower, cleaning and certifying it, and organizing meetings with the logistics team about re-erecting the tower somewhere else.

There is therefore an enormous amount of effort expended in order to collect the raw data from the forest. This is because this raw data, in the scientific community of which the LBA is a part, is immensely valuable. This is not only the case of the LBA tower data, but most data that the LBA community produces. On several occasions, I heard researchers remark that "this has never been done before", or that "this would be the first time this variable was measured in the Amazon". On one occasion, a German researcher working with one project told me that she considered it probable that the data produced in the Amazon would be published even if it were not up the same 'quality' as data from elsewhere, simply because there was no other data like it. This is why researchers will go to such lengths to obtain it. The possibility of obtaining raw data on something that has never been measured before means that many different researchers are drawn to the Amazon forest, and to the LBA.

Mark is one such researcher. Based at Harvard University in the USA, Mark came to Manaus to install a new piece of equipment on the tower at K34. His instrument is known as a "Picarro", after the manufacturers, and could measure not only CO_2 and H_2O, but also CO (carbon monoxide). Mark is a specialist in laser engineering, instrumentation, and trace gas measurements. He originally came to Brazil to work with a well-known Brazilian LBA researcher in atmospheric chemistry, using the LIDAR, another instrument that uses a laser to estimate the amount of particles that are in the atmosphere at certain heights. However, the LIDAR turned out to have been seriously damaged whilst in transit. So Mark had come up to the LBA in Manaus, and the head of the Micro team had suggested that they install the Picarro on the tower. Mark was very excited, because as he explained, this would be the first time this sort of data had been obtained in "real time" – transmitted and available immediately and even, Mark hoped, directly transmitted to his desk in Harvard. Furthermore, measuring methane, CH_4, $CO_{2,}$ and now CO, all at the same time would provide a unique picture of the chemical relations in the air. As Mark explained to me, the relation of these gases to each other is very important in global atmospheric chemistry directed at understanding climate change. Because, for example, CO combines with O_3 (ozone), to give CO_2 and O_2 (oxygen), the global CO_2 estimate in part depends upon the presence of CO in the atmosphere. This in turn depends upon knowing how CO behaves in tropical biomes - of which the Amazon forest is the largest in the world. Furthermore, just measuring concentrations, Mark told me, isn't chemistry. You need simultaneous measurements of all of the gases to be able to see the way in which they are interacting. This sort of chemical measurement had never been done in the Amazon before, which is why he was so excited.

Mark was therefore willing to overcome all sorts of obstacles in order to collect this data. One key consideration concerned the energy the Picarro would need. The energy for the instruments on the tower comes from solar-charged batteries, so their available voltage depends on the weather – and Mark was not sure if that would work for the Picarro or not until it was tested. He also had to decide how to house the delicate instrument. This is far from inconsequential. The Picarro had to be protected from the rain, and more importantly, from the bees, that might cover it in wax. But at the same time, it could not get too hot. Its box therefore needed to be sealed to stop water and insects getting in, but this caused a problem in terms of heat getting out. Mark spent some time toying with ways to outsmart the bees. He considered installing a little fan in the box. However, whilst the fan would keep the bees from being able to fly in, if the fan stopped working due to a power cut, he then would have given the bees easy access to the Picarro. In the end Mark decided that he would put the fan in and risk the bees, but that he would sit with the instrument whilst it was running to see how hot it got. The Picarro has an internal fan, but with all the padding that needed to go round it in the box in order to protect it on the bumpy trail, and the ambient temperature at about 30^{o} C, it was extremely likely that it would overheat so would need to be constantly monitored. When packed in the box, the instrument weighed 60-70 kilos, and it took two people to hoist it up the tower using a rope-and-pulley system, and another two people guiding it so that on its way up it did not bang into the tower, or any other equipment. It would have been impossible to do it alone. Once up there, Mark managed to get it working to his satisfaction, but given the capricious weather conditions, he was worried. He was so worried about the instrument overheating that he left it running only when he was there - a total of about 5 hours, during which time he sat with the instrument in the tower himself. Food was brought out to him whenever possible. Yet he seemed unfazed by the ordeal. "That'll be some good data", he told me, looking pleased.

There is no doubt that Mark would not have been able to collect any data had it not been for the actions of others, and Mark himself was as aware of that as anyone, as he impressed upon me several times. It requires an enormous amount of work to install an instrument such as the Picarro in the middle of the Amazon forest. It is in fact so difficult that it has never been done before, and has not been done since. What this means is that the raw data that Mark collected is the only data of its kind. There is no other data like it that exists in the world.

Dirt Kills the Curve

There is another dimension, however, that runs parallel to raw data's uniqueness. The difficulty of the terrain and the isolation of the research site, which is exactly what makes the raw data unique, is *also* what makes it uncertain and ambiguous.[64] As it is expressly not being collected in a laboratory environment where conditions can be controlled, the raw data is full of errors, and uncertainty. This is the case for all of the raw data collected in the forest, but especially so for the data that the LBA produces continuously - the data that is collected from the towers over the long term. Whereas Mark was conducting a short pilot experiment and therefore could sit in the tower with the Picarro, and E (who we met in the first chapter) was collecting data over a limited period for her master's project and so also accompanied the data collection process, the bulk of the data that is from the towers is a result of automatic data collection and storage by the instruments. Whether the raw data is sent in real time to the LBA in Manaus or collected weekly or twice a month by a technician, it still means that the instruments are left out in the forest for long stretches of time, and the dataloggers store the raw data automatically. What this in turn implies is that the raw data from the towers is potentially a lot more uncertain. I have already introduced the idea of uncertainty in the previous chapter. Here, I will describe the different sources

[64] In a subsequent telephone conversation, Mark told me had not used that Picarro data in a publication, in fact, despite his excitement at collecting it. Some of it was really good data, but there had been too many errors and there had "not been enough to draw trends".

of uncertainty and error that enter into the raw data in the specific case of the LBA tower data. It is important to note that alongside the sources of uncertainty that were introduced in the second chapter, some of which I argued are constitutive to making any knowledge through measurement, there are sources of error and uncertainty that are specific to the Amazon forest.

Some uncertainty in the raw data in inevitable, (as I have discussed in the last chapter). This sort of uncertainty comes from the fact that, as the metrological technicians at the Laboratory of Meteorological Instrumentation (LIM) told me, "no sensor is perfect". That is, no sensor can measure without uncertainty. One cause of this is the undesired effect of measuring *in situ* – for example, air temperature sensors heat up exactly by being "in" the heat, so the heat they measure is partly their own, rather than that of the air. Alongside this, there is the uncertainty due to the loss of information in the conversion of analog signals into digital, a conversion almost all data is required to undergo. No less important is the fact that all instruments are surrounded by a larger or smaller uncertainty factor. The manufacturers provide the instrument with a value of uncertainty – for example, 10%. This means that if that instrument measures 25° C, this value actually could be anything in between 22.5° C and 27.5° C. This number is a result of the metrological testing and calibration undertaken by the manufacturers, and is unavoidable wherever measurements are being made.

There are also uncertainties that are inherent to more complicated measuring methods. Consider the technique for measuring carbon flux, called the eddy covariance method. The carbon flux system installed on the tower K34 comprises of a sonic anemometer and a Li-7500 open-path IRGA. The sonic anemometer measures wind direction in three axes using sound waves that are pulsed between three triangulated detectors, corresponding to three dimensions, vertical horizontal and lateral. These pulses are interrupted by the turbulent vortices and eddies created by the interaction of the wind with the top surface of the forest canopy, and from this interference the sonic anemometer infers wind turbulence direction and speed. The IRGA measures the concentration of CO_2 and water vapour in the air at very high frequencies; both the anemometer and the Li-7500 record measurements at 10 times a second (10Hz). Correlating the concentration of CO_2 and eddy direction and speed, the eddy covariance method can give a measurement of carbon flux, that is, how much carbon is moving per area per unit squared. A document provided by Licor (the manufacturers of the Li-7500) which a member of the Micro team passed on to me, suggests that a useful analogy is with a measurement of the number and speed of birds flying through a window. Birds are one thing, but for molecules this is rather complicated to calculate, partly because of the rapidity with which molecules move about, and partly because a lot of different factors have to be taken into account. In fact, due to the complexity of the calculation, some factors that are known to influence flux simply cannot be taken into account, such as mixing ratio and air density – and this is despite the method being state-of-the-art. This omission imparts a certain level of uncertainty that accompanies the eddy covariance method irrespective of where it is applied. When I asked about this, I was told by one researcher that the scientific community knew about these issues, of course, but at the moment the eddy covariance method was considered to be the best way available to measure flux, so they kept using it. That is, just because the method is uncertain does not mean it is redundant.

However, in the case of the LBA raw data from the towers, specific uncertainties that in other places are minimal are magnified considerably. One of these sources is the footprint of the tower. The footprint of the tower – how far the IRGA and the anemometer can "see" – is calculated assuming that the terrain in question is flat, as it is in pastureland, for example. As several researchers impressed upon me, the towers are therefore conventionally understood to measure *vertical* carbon flux only. However, the Amazon forest is far from flat. What this means is that there is horizontal or lateral carbon and energy flux as well - the CO_2 "rolls" down hills and collects in basins and valleys. As a result, as one LBA researcher , whose PhD was on horizontal carbon flux at ZF2 and in Santarém told me, the raw data from the towers also included the effect of this horizontal carbon flux caused by the topography of the land. This poses a serious problem to the representative power of the eddy covariance method. The researchers had assumed it to be total when in fact it was partial. A further problem for the eddy covariance method is that at night, the eddies and vortices are much smaller than during the day, so the specific time-frame used for measurements can affect whether the effects of small-scale vortices are included in calculations or not. This problem is exacerbated by the fact that whereas in flat pastureland the topography is uniform and so the vortices do not vary in size so much, the "roughness" (*rugosidade*) of the undulating tree-tops creates all different sizes of vortices. All of these complications make the raw data collected in the Amazon forest particularly error-prone, and increase the uncertainty around it substantially.

On top of these issues are the problems of installing and keeping electrical instruments well-maintained in the Amazon forest. The forest, or the “world out there”, interferes unremittingly. The struts of the anemometer themselves can often very slightly slip – as the tower moves about an alarming amount in a rainstorm – which will affect the calculations of the directionality of the wind eddies. Other researchers who use the tower to attach their own instruments temporarily, or groups of students who are taken up the tower, can unwittingly interfere with the precise positioning of the IRGA and the sonic anemometer, causing the same problem. There were a number of occasions at the more remote tower at SGC when the technician would arrive to find the instruments disconnected, or turned upside-down, as if the result of a curious passer-by (because of this and other strange occurrences, the Micro team would sometimes joke about there being spirits of the forest intent on mischief). Bees colonize the boxes that house the datalogger and slowly turn it into a hive, coating it in wax. When it rains, the open pathway LICOR-7500 cannot take measurements because there are rain drops on the lens that do not evaporate due to the intense humidity in the forest. In the manual for the instrument, it only warns users about the effects of snow. Heavy rainfall can mean the pluviometers are so inundated that they do not measure correctly.

But the most substantial problem is also the most simple: dirt. It piles up on the instruments, coats the thermohygrometers, covers the IRGAs, spreads itself over the radiometers. It requires someone to constantly remove it, and the instruments should all be taken down from the tower and thoroughly cleaned as often as possible. The very fact of the instruments being out in the world, in the forest, affects the data that is being produced. Maintenance of the equipment is therefore one of the main jobs of the Micro team in Manaus, but despite their best efforts, raw data is always a measurement *plus something else*. As the head of the Micro group told me "dirt kills the curve" (*a sujeira mata a curva*); that is, the dirt on the instrument will affect the shape of the graphed line of the data - the dirt is included in the measurement, as it were. The raw data that R, J and T receive therefore contains potentially all of this uncertainty. Raw data thus could be said to contain *too much* of the world. Not only does it contain the measurement of the variable in question, one hopes, but it also contains what could also be considered to be measurements - digital traces, at least – of potentially all sorts of other entities as well, including insects, unwanted meteorological events, and even unknown future variables. This is what its ambiguity consists in – its plurality here is given by this ultra-connectivity to the world it is in.

The obstacles and difficulties that that Amazon forest presents to the collection of the raw data are exactly what make it unique, at the same time as they ensure its "rawness", and plurality. Its uniqueness resides in the particularity of its ambiguity. Though some uncertainty factors would apply wherever the instrument was measuring, these other “bits of the world” are specifically Amazonian. Even with cutting edge instruments, even with meticulously planned methodology, and access to a team of willing people who know the forest and can help you, it is very hard to produce Amazonian data that is singular in meaning. What is produced, in fact, is very raw data that is often, as it stands, semiotically uncertain.

Isabelle Stengers argues in *The Invention of Modern Science* that the power of modern science lies in its ability to “invent possibilities of representing, of constituting a statement that nothing a priori distinguishes from a fiction, as the legitimate representation of a phenomenon "(2000 [1993]: 87). According to Stengers and also Karen Barad (2007), the power of scientific apparatus is to "singularize" (Stengers 2000 [1993]: 85) the fiction that is being legitimized in such a way that it can be both repeated and undisputed. But the instruments that are installed on the towers do not, in fact, manage to do that. As Stengers herself points out, when not in the laboratory, this act of singularization is impossible. The field sciences do not, then, perform the same sort of singularization as the "lab sciences".

However, the raw data suggests that the act itself of "singularizing" can be accomplished in other ways. When one follows the world outside instead of making the world inside a laboratory, one of the bastions of representativeness that disappears is repetition. The experimental set up cannot be repeated, because the experimental set up was simply not set up in the first place. One cannot simulate the Amazon in different places in order to check the measurements.[65] Furthermore, going out into this Amazonian setting is very difficult, and requires an infrastructure of which there are very few. So, the raw data that emerges from the LBA is singular not because of the instruments, which as we have seen often have their authority compromised by the world, but because it is the only data that exists about the Amazon. Mark was excited exactly for this reason - studying the atmospheric chemistry of the carbon cycle in that way in the Amazon had never been done before, because it is so difficult to do. Even though he only collected data for a few hours, it was enough. Compared to nothing, a little is everything.[66]

I do not mean to give the impression that the data that is finally processed is necessarily singular in meaning because it is singular in source, although this is sometimes the case.[67] I am suggesting, instead, that the LBA raw data implies that there might be ways for the power of the singular to be co-extensive with semiotic multiplicity.[68] There has been a lot of emphasis placed, in STS and the anthropology of science, on the processes by which the 'universal' is constructed (for example, Latour 1983, Haraway 1997, Tsing 2005). Numerous scholars have traced the ways in which a fact only becomes singular in meaning and universal as a result of substantial effort, and by being reproduced everywhere. Donna Haraway remarks that "to be

---

[65] Although, in fact, this is exactly what the engineers and scientists behind the Biosphere 2 ended up doing. Biosphere 2 is an extensive indoor ecosystem constructed in the 1980s in the Arizona desert, originally to investigate self-sustaining space-colonization technology. It subsequently has been used for scientific investigations of the effect of carbon dioxide on plants. A researcher from the University of Arizona I met at the LBA told me that it was sometimes used to train students in fieldwork before they were then sent "to the real jungle". See http://www.b2science.org/ Accessed September 2012.

[66] Even if the data collected turned out not to be as high quality as Mark had hoped, his excitement was still tangible – as, when collecting it, it *could* have been. This is another version of raw data's ambiguity. (This footnote was added after the defense of this thesis).

[67] Certainly, there have been disputed results using data from the Amazon about carbon, such as the data that the LBA provides. The most notorious was the carbon sink controversy, over whether the Amazon is a sink or a source of carbon (see Lahsen 2009). As one LBA researcher informed me, it was the mismatch between how much carbon was calculated as being taken in by the forest using the eddy covariance method, and how much was estimated to be in the forest using biometric methods, that led to the discovery that horizontal carbon flux was as an important factor to take into account. This discovery challenged existing data sets on carbon flux (cf Mahli and Grace 2000).

[68] This recapitulates the importance, as I pointed to in the first chapter, of a capacity to singularize; the threat is the excess relationality in the world. In this chapter, one might say that we encounter a re-emergence of this potential excess of relations in the raw data, and that again it is the task of the data cleaners to cut this away in the next stage of the process of the production of certified data, as described in Chapter 4.

meaningful, the universal must be built out of humans and non-humans (Haraway 1997:68) and, as Anna Tsing reminds us, this means that universality "can only be charged and enacted in the sticky materiality of practical encounters" (Tsing 2005:1). What the raw data draws attention to is the observation that, for the LBA at least, a lot of time is spent not on constructing, maintaining and negotiating universal truths, but on extracting *unique ambiguities*.[69]

ANT and Accumulation

The particular configuration of the singular and the multiple that the LBA raw data exhibits can therefore be brought to bear on some of the dynamics by which scientific facts are understood to be stabilized. I have specifically in mind the dynamic that characterizes ANT, whereby the accumulation of associations in a network is what successively establishes the fact in question as a fact (Callon 1999 [1986]; Latour 2005; Law 2003). The emphasis in such descriptions is on the way that facts, or more radically truth, are a result of the collective and contingent process of what Bruno Latour calls "articulation". This is one of the terms Latour employs (idiosyncratically) to refer to the work that goes into creating the relations between the entities in a network. These associations include the persuasion of colleagues, enrolling of funders, cajoling of objects, objection of objects, and the creation of an audience, in order to achieve "*well-articulated actors*, associations of humans and non-humans"(Latour 2004a: 86; cf Bowker and Star 2000; Gerson and Star 1986).

I will concentrate on ANT here, but it is of note that a great deal of STS scholarship has contributed towards the exploration of science in these broad terms, often emphasizing the importance of factors conventionally deemed to be 'non-scientific' (for example, Derksen 2000). However, the scope of the study can expand or contract to include more or fewer actors and entities. In a classical ANT study, the emphasis of such a description is to "follow" (Latour 2005) all the multiple sorts of entities that

[69] Relevant here is Anne-Marie Mol and John Law's discussion of "fluid spaces" that "are defined by liquid continuity" (1994: 659) as an alternative way to think about social topology than that offered by ANT networks; although they admit that such a notion still does not "'really' get at the chaos" (ibid: 663). The raw data is not a fluid space, but it does exert a semiotic pressure that counters the network in the same way perhaps as fluids do in Mol and Law’s description.

need to coalesce around a task, in different ways, in order for scientific knowledge to be produced. In ANT, Latour takes this to one of its logical conclusions and suggests that agency is therefore distributed between all those engaged in the fact-building (including those entities and factors normally ignored, the "non-humans" and the discursive, for example). Latour's examination of Pasteur, therefore, demonstrates the articulation and "translation" of the interests (Latour 1999a: 311; Callon 1999 [1986]: 81) of not only the most obvious human (Pasteur) and non-human (the anthrax) in the story, but also French farmers, statisticians, public health workers, Pasteur's colleagues and the French public; as well as a whole host of sundry non-humans (Latour 1988; cf Latour 1999a: 113-173).

In more focused studies, often based in laboratories (for example, Knorr Cetina 1981, Knorr Cetina and Amman 1990), the description usually centres on how alternative possible interpretations for the "inscriptions" (Latour 1987: 79) or results produced during an experiment are progressively discarded or refused by the scientists in favour of a single, true one. These studies sometimes revolve around a controversy, such as the discovery of a new kind of particle or a new kind of physical or biological entity. This process of reduction in meaning occurs through discussion, negotiation, calibration and other material and semiotic means specific to the discipline in question (Collins 1985). This has been variously described as "black boxing" (Latour 1987: 131), "fixation of evidence" (Amman and Knorr Cetina 1990: 88), "controlling interpretative freedom" (Collins 1985: 106), amongst other denominations. Although most STS studies in this vein engage with different local practices, scientific disciplines and ethnographic settings, the transformation of ambiguous results into singular meaning via mediums considered non-scientific is frequently highlighted as the governing, (and often problematic - see Latour 1999b) dynamic, whatever the science.

There is however an interesting double quality that characterizes the reduction in meaning that ANT descriptions catalogue. The fact-to-be (or “matter of concern”) assumes a singular meaning (becomes a “matter of fact”) paradoxically exactly by being increasingly “articulated” into networks of different sorts of entities. The *more* connections and articulations it has, the more stable, indisputable and therefore *singular* it becomes: as more and more humans and non-humans are convinced of and enrolled into its existence, the fact becomes more stable and real. As Latour writes, "reality grows precisely to the same extent as the work done to be sensitive to differences. The more instruments proliferate, the more the arrangement is artificial, the more capable we become of registering worlds" (2004a: 85). By implication, whereas a paucity of connections surrounds a proposition that is ambiguous and plural in meaning, as the number of connections and associations and actors increases - as the network grows - this ambiguity decreases. In this regard, there seems to be a particular sort of connection in a successful network – enrolment – implying a specific sort of relation - one of consent, or assent. Those who are enrolled into the fact's existence assent to and thus contribute to its singular reality.[70]

There is another slant to this generalized process as described in ANT studies. Those associations conventionally seen as 'non-scientific' become, by dint of their inclusion in a process they are normally excluded from, some of the most noteworthy members of any network that produces scientific knowledge. Likewise, the inclusion of objects as actors in their own definitions is another notable addition. Thus the network "is indeed a melting pot, but it does not fold in together objects of nature made of matters of fact and subjects endowed with rights; it mixes together actants defined by lists of actions that are never complete" (Latour 2004a: 80). The well-known take-home message is that if the work of building a tower is as crucial to the reality of a scientific fact as a well-performing instrument or a scientist, then there is very little basis for suggesting that there is a stable definition of what is 'inside' and what is 'outside' science - the scientific and the non-scientific are defined themselves by the network. Further, if a well-performing instrument is as crucial as a scientist, there is also from this perspective no basis for a stable definition of 'objects' and 'subjects', or the social and the natural - these too are defined by the network. In order to make this point, certain connections therefore are brought into the foreground that otherwise might have been back-grounded - the "non-scientific". It is these sorts of connections that

[70] As an aside here, it is worth noting again that the entities that Latour (and others) discuss are often entities that are discovered/fabricated in a laboratory. It is therefore easy to summon the image of isolated ambiguity: the lone researcher at his desk, with suspicions that he might have 'discovered' something of note and the subsequent attempts to convince his colleagues that he indeed he has, plays into an idea of the increasing connectivity and stability (and therefore the "constructed" nature) of his discovery in the world (what Haraway calls the "heroic" depiction of science, 1997: 32). This is perhaps more convincing when the location of discovery is a laboratory, and what is being discovered is a new entity. Neither of these conditions generally apply to the LBA researchers.

contribute towards the reality of the entity and its status as a fact. To achieve this status necessitates the eradication of ambiguity and the stabilization of meaning, through accumulating all sorts of associations and actors - including those that conventionally lie outside the realm of the scientific and whose consenting or "enrolled" presence attests to and thereby constructs the fact. This is a successful network.

Given this overall dynamic of accumulation, it might be expected that the raw data should be unconnected to the world at large, as it is exactly an entity that is awaiting semiotic stabilization. However, from the perspective of the LBA researchers and technicians, the 'problem' of raw data seems to be exactly how overly *connected* it is to all sorts of entities in the world out there already. But these are the *wrong* sorts of connections. The extraneous relations it has with the forest, such as bees and lightning and dirt, obscure the singular meaning presumed to be lurking somewhere within it. Although there is no question that the raw data only emerges because of the work of a team of people and objects that must be constantly enrolled, cajoled or otherwise convinced to contribute to its production, this does not exhaust the connective repertoire of the raw data.[71] What makes the raw data "raw" - what makes it what it is - is precisely the tension between the work of those actors in the network that toil towards eliciting a stabilized set of certified data, and the connections and relations that the raw data has that lie, in a certain sense, *outside* this network altogether, anchoring the raw data in the liminal state between meaning and lack of meaning. Concerning what might exist before the network, Latour explains that a "vague, cloudy grey substance" (1999a: 45) turned into the "microbes" that Pasteur famously 'discovered' only through successive articulations, associations, enrollments and substantial effort by Pasteur in the "retroproduction of history" (ibid.: 169). Pasteur "happened" to the microbes, as much as the microbes "happened" to him (ibid.: 146). But ANT narratives must begin somewhere, even if the networks they trace are infinite. What is that cloudy grey substance, then, before Pasteur entered into contract with it? What can we learn by paying attention to these 'pre-network' entities?[72]

[71] See also feminist critique of ANT here, for example Star 1991.
[72] Nick Lee and Steve Brown ask a similar question in wondering where the "Other" might have a place in ANT descriptions. By including everything in the network and treating everything symmetrically, ANT implies that there is no Other. Lee and Brown conversely

I say 'pre-network' with a pinch of salt, because of course for the raw data to emerge, it needs to already be part of a network of people and things. But the raw data is "raw" because of the particular connections that it has to the Amazon forest - not only the ones that the researchers are interested in, but also ones that would be considered to be "error". These connections have to be removed in order for the certified data to stabilize as such, and in order for the data to convince others of its meaning. The first thing to be learnt then is that in some cases, scientific work consists not in accumulating but in *substituting* one sort of relation for another. The relations that are removed are the wrong relations to the rest of the world, known as "error". This substitution is what the transformation of raw data (ambiguous) into certified data (ostensibly singular) involves. Thus it seems to me that the defining distinction that emerges from examining the beginning of this particular network is not that between social and natural, or scientific and non-scientific, but between data and error.

Not Data, Not Not Data

Instead of concentrating solely on what the raw data will become, this chapter has tried to concentrate on what the raw data is already. I now want to explore the idea of what I have called a 'pre-network' entity. The raw data is understood, by those I worked with, to be awaiting transformation; it is unfinished. Thus in paying attention to it, one finds one's attention constantly directed elsewhere - to what it will become, rather than what it is. But it is exactly this transformative potential that makes it so valuable. The raw data is ambiguous and uncertain, but it is still highly-prized by the researchers. As we have seen they go to great lengths to obtain it, and several researchers told me that they will share their data only after they have processed it. It is in some sense, therefore, more valuable in its private raw state than in its public processed state - even though in its raw state, it is considered to be potentially non-

---

urge for a "form of scrutiny that is prepared to allow for the deterritorializing, rhizomatic movement of irrevocably splintered entities in their half-realized fractal strategies" (Lee and Brown 1994: 787), that lie between the purified and the mediated. Although this is not exactly how I would describe the raw data, there are obvious resonances. Latour also approaches this question through the notion of "plasma", which is "that which is not yet formatted, not yet measured, not yet socialized, not yet engaged in metrological chains" (Latour 2005:244). In Law's terms, this is “what is absent...a set of potential patterns that buzzes and dazzles and dances, that is too complicated to condense, to make present” (2004: 117). However, I think it is as important to pay attention to the temporal, and not simply the spatial, dimensions of these entities.

sensical. I want to draw attention to how much scientific information passes its time, in the LBA, in a quite normal fashion in this way: un-interpreted potential non-sense. When I say quite normal fashion, I mean in an uncontested fashion - the raw data that sits in a database waiting to be cleaned by the micrometeorology team at the LBA, for example, does not earn its ambiguous status because there is a controversy surrounding what it means. It is *expected* to be ambiguous. It is data that probably needs a lot of work, that is awaiting the transformation into certified data. Its ambiguity is inherent to the object itself. The ambiguous state I am referring to does not mean that the *content* of the data is necessarily ambiguous; the question is not, at this stage, what the data means, but if it is data even at all. If it is, then it will *always* have been data. If it is not, then it will *never* have been data. It becomes in a certain sense what it already is (cf Goldman 2009: 115).

This state of affairs, in which the relation of data to error as yet unfixed, defines all data that is unable to lose its rawness sufficiently. One student I spoke to, who was studying the flux of energy in the basin at ZF2 using data from the tower B34, had found on analyzing the data that the water vapour flux (which is a correlate of the energy exchange) was "very strange". The balance did not "add up" - the terms were unequal on either side of the equation. "In fact, I thought they [the data] were wrong. But then, when one of the reviewers of the committee read it, he told us that it was something he had noticed about the river himself...". According to this committee member, it was not necessarily an error. He had seen the same anomaly in another part of Brazil. "It's a question of the lateral transport of energy. It's like what [another LBA researcher] saw, only [he] saw it for CO_2, and mine is humidity (...) So, it isn’t an error, but to prove this, I need more data (...) I think it is the same thing as he [the committee member] observed there. We're not certain, because we need a longer series of data." Why not just accept it as an error? I asked. "The first time you see it, you think it's error because the difference is really big. So, it could be an error associated with, let's say, the calibration of an instrument. But it is actually the same sensor that measures these two. So...I thought there was a possibility, as I put there [in the thesis], that there might be a calibration error...or it might be the lateral transport of heat. Of vapour in fact, of humidity, right? When the reviewer read my thesis, he was like, "Ah, I saw the same thing. And I think the problem you have is nothing to do with calibration, it's not an error, it's lateral transport (...) I know it wasn't an error...you are never certain, though, right? You imagine that it's an error. But now, after [the committee members'] result, I know that it could be...*could be*, I am also not certain that it is lateral transport. That's why I need more data." As her supervisor told me, it's a question of "distrust" (*desconfiança*), because they "don't know exactly what it is". They were uncertain about whether the data was in fact data, or if it was error, because without enough of it, this distinction is difficult to make. The distinction is therefore to a certain extent constrained by patterns that other data, that has already been distinguished, displays.

This tension is also captured in Susan Leigh Star and Elihu M. Gerson's work on anomalies in science, especially what they call "artifacts" (1987). Star and Gerson are dissatisfied with what they see as a tautology in Thomas Kuhn's description of anomalies, that suggests that an anomaly is only an anomaly because it does not fit with the current paradigm; that is, something is an error if it is wrong. They suggest that this can be resolved by looking at the history of an anomaly, and how it flickers between being an anomaly and being a discovery, in their case-study of neurophysiological research into the localization of functioning of the brain. They write: "The anomaly went from mistake to artifact to discovery - even through accusations of fraud - until it was made tractable enough to be absorbed into an ongoing enterprise." (bid: 160). The point they make concerning artifacts is that the artifact can change status very rapidly. This is because it contains the potential within it to be *either*. Although Star and Gerson are writing with a very different approach in mind to mine,[73] the observation resonates. The raw data exhibits this sort of potential, but in inchoate form: it is neither phenomenon nor artifact, neither data nor error, but the *tension or relation between the two*. No-one knows, when it is collected, which one of the two it is.

This also means that what is eventually error to some may be data to others. David, who taught the thermodynamics class on the LBA postgraduate course, wanted to

[73] They continue: "Its history cannot be explained simply in reference to its logical place in a problem structure since it moved from contradiction to discovery as the problem structure was adapted to work conditions, nor as simply a matter of political convenience since it was taken seriously and changed the shape of findings. Rather, it shows the importance of independent bases, multiple lines of work intersecting, and visibility interacting with logical significance to produce changes in acceptability and resources deployed to control the anomaly" (Star and Gerson 1987: 160).

study the self-organization of deep convection in the tropics. In order to do so, he needed to collect data on how much water vapour is in the atmosphere. So he engineered a collaboration with some geodesists he know in S*ã*o Paulo. In the discipline of geodesy scientists are interested in collecting data on the miniscule movements of the Earth's crust, often using GPS instruments. However, in order to achieve such precision, they have to get rid of the interference caused by the effects of the atmosphere, mostly of water vapour, in between the satellite receiving the data signal, and the GPS on the ground. This can only be done afterwards, through a mathematical filtering process. The "interference" that they remove, was in fact the "data" that David needed. Thus a symbiotic collaboration resulted, with David agreeing to install and maintain the GPS systems, and his collaborators providing him with the instruments.

Thus what I am describing lies somewhere just before, (and therefore in a sense in between) an ANT analysis, in which the focus is on the construction of truth, and Gaston Bachelard's description of science as constantly finding itself to be in error. As Hans Jörg-Rheinberger tells us, "[O]ne of Bachelard's central claims is that "the scientific spirit is essentially a rectification of knowledge…it judges its historic past by rejecting it. Its structure is its awareness of its historical errors. Scientifically, one thinks truth as the historical rectification of a persistent error, and experiments as correctives for an initial, common illusion" (Bachelard 1987 [1928]: 297 in Rheinberger 2012: 29). Rheinberger explains therefore that, in Bachelard's view, "a scientific truth of the present must always be ready to find itself in turn as an error of the past. Therein lies the essence of the historicity of the sciences; it is this that stamps them as special cultures of veridiction and verdict" (Rheinberger ibid: 29). What the raw data offers instead of any one of these two positions is rather the moment before either of those positions are even ascertainable. The technicians and researchers do not know the relation of their raw data to the world when they collect it - if it is error or data. Both these positions are contained within it, because what will end up being data will have *always been* data, and what will end up being error will have *always been* error. What ends up as error is in fact, following my analysis, simply the relations and connections that the researchers are not interested in. These are the associations that need to be removed. However, until it is analysed in this way, the raw data is thus not data, but not exactly *not* data. It exists, I argue, as a moment (sometimes protracted) between being data and being not data. If that is so, the raw data is the moment before data and error exist. It is "vague" and "cloudy", perhaps, but as such it demonstrates the extent to which science, at least for those in the LBA, is populated by uncertain, vague and ambiguous entities, and not just reified, certain and stable ones, nor definitively incorrect ones. All the effort of the LBA generates in the first instance neither data nor error, but the potential for the relation between them.

Conclusion

In this chapter, I have explored the raw data. I have documented how it seems to be unique and simultaneously ambiguous. I have suggested that this state needs due attention paid to it, because it lends the raw data some particular properties. By being the only data of its kind in the world, it is highly sought-after; by containing too much of the world within it, it is also ambiguous and uncertain. This combination of semiotic plurality and uniqueness provides a perspective from which to re-assess certain ANT claims of how network-building progresses and what it results in. First, it suggests that associations and articulations are not necessarily only accumulated, but in fact, substituted. The raw data's ambiguity for the researchers resides in the fact that it manifests the (excessive) wrong relations to the world.

But the raw data also suggests that the perspective from which it offers such a vantage point might be one that, in an important sense, 'pre-exists' the network. Because ANT descriptions of network-building, in science at least, are so geared towards following the processes of stabilization and singularization of meaning, the phase of "non-meaning" or "non-sense" is often passed over as merely a means to an end. This chapter has instead tried to concentrate on what this state of non-meaning might consist in. This could be seen as a partial answer to Nick Lee and Steve Brown's (1994) question: where is the Other in ANT? However, for the purposes of this thesis, it also channels attention towards the crucial distinction that the raw data expressly does and does *not* evidence, namely, what is data and what is error. The way that this distinction is elicited and what that in turn means is the subject of the next chapter.

Chapter 4: Cleaning the Data

Introduction

This chapter deals with the process that my informants called "cleaning the data" (*limpar os dados*). Often known as "data quality control", this sort of process is ubiquitous in observational science (Zimmerman 2008; Edwards et al. 2011; Norton and Suppe 2001; Edwards 1999, 2006), but the form it takes varies. Even within the LBA, the details of the cleaning process vary tremendously depending on the conditions of the data's collection and the use to which it will be put. In this chapter I am specifically concerned with the cleaning of the long-term data from the towers, rather than the data from single collaborative campaigns or individual's master's or PhD projects.[74] However, I suspect that the more general trends and characteristics that I point to as pivotal to my argument hold for those other cases as well, at least for those data collected in the Amazon forest.

Cleaning the raw data, I will argue, involves a series of actions that, if successful, allow the certified data (*dados certificados*) to carry within it its own referent, rather than refer back to the world in a conventional relation of correspondence. The certified data becomes in this way its own reality. This transformation involves the decomposition of the raw data into error, which is discarded; and data, which is kept. Although the process is made up of several stages, which I describe in some detail, the crucial action is referring the data back to itself through a self-scaling dynamic, so that it comes to define its own limits. I employ the notion of "compaction", drawn from the work of Donna Haraway and Marilyn Strathern, to describe this occurrence. In light of the creative potential that this particular form of recursivity is seen to have in the LBA data cleaning, the chapter ends with a reappraisal of the debate around

[74] In the case of postgraduate projects, the exact form that processing the data takes is up to the individual, although guided by disciplinary norms, institutional protocols, and supervisor's templates. I was unable to follow a student's data cleaning process very closely as this did not necessarily happen in the LBA – it could happen wherever the student happened to be. As there were also no collaborative data collection campaigns occurring during my fieldwork apart from in the final few weeks, I have concentrated my argument on what I did observe and accompany which is the cleaning process of the long-term data from the towers.

circular reasoning in scientific practice, exemplified by the "experimenter's regress" (Collins 1985: 84).

The Protocol: Initial Domestication

R and J, the two data cleaners whom I accompanied most closely during my fieldwork, perform the data cleaning in a large, relatively bare air-conditioned room lined with computers. People are constantly coming and going, and on a big screen in the middle of the room the raw data of meteorological variables (the "low frequency" data) from the tower K34 is displayed in real time as a collection of graphs. These graphs change continuously as the raw data is received. The contrast between the Micro office and outside is very stark; often I would have to wear several layers of clothes due to the artificially icy conditions, despite it being 35^{o} C outside. The data cleaning consists in a lot of work on a computer, so initially I required R and J to explain to me painstakingly what they were doing, before I could try to follow without explanation – which I almost invariably failed to do even after several months as they worked incredibly fast. J was a great deal less averse to taking on the role of teacher than R, and as a result I have more detailed information about and examples of the tower at SGC that J is responsible for. However, the process is basically very similar for data from all the towers.

In 2009, the LBA instigated the use of a generic data quality control protocol, which one of the Micro team members adapted from a protocol that was in use at a university in Holland. This protocol – the "protocol for the treatment and certification of data from experimental sites in the Amazon" - is a document that forms the organizational basis for the process of cleaning data. It explains how the data should be treated, structured and stored. It was presented to all the groups involved in collecting and treating data from the LBA towers in the Amazon during a meeting that was held in 2009. Subsequently a guide was produced with each step in the process carefully explained.

The structure that the protocol describes takes the form of a series of folders nested inside one another on the computer. The first folder is identified by the location of the experimental site (for example, 'Amazonas'), the name of the experimental site (for example, ZF2), and the name of the tower (K34). Within this folder are further folders for each of the different possible data collection platform – either "AWS" (automatic weather station), "Fluxes" (*Fluxos*), or "Soil" (*Solo)*. Within each of these this folder are four more folders, pertaining to four different "phases" of the process- "raw data" (*dados brutos)* "pre-analysis" (*pré-análise),* "post-analysis" (*pós-análise),* "certified" (*certificados).* The "fluxes" folder has an extra phase called "Edd", which is where data for the programme Alteddy, which calculates fluxes of carbon dioxide and water, is stored. Within each of these "phases" folders, there are folders for every year for which data is held, and within each of those, a folder for each of the months of the year. Opening these nested folders consecutively, it is only within the “month” folders that the data in various states is to be found.

Due to the different formats that the raw data comes in (depending on which of the data collection platforms it originates from), the first step that the data cleaners take is to rename the raw data files. The new file name takes a very precise form. It is composed of the name of the research site, the data collection platform, the year, the day and the time of the last data collection, and the date and time of the present data collection. The day and month are written not in calendar form, but as DOY, "Day of Year" (or DDA, *dia do ano*), which treats the year as a continuous period of time. This means only one number is used: '35' would be the 4th of February, for example. The data collected therefore lies in between the two collection dates – so in this example, K34_aws_08_008(1500)_015(0930), the data covers the period from 3 pm on the 8th of January 2008 up to 9.30 am on the 15th January, which is when it was collected.

The protocol states, regarding the re-naming of raw data, that as only the name of the data is new, and the content has not been altered, it is "maintained integrally as "raw data"". It was impressed upon me several times how vital it was to have this untouched raw data set. "Never delete the raw data!" R would warn a student who had come to the LBA in Manaus to learn about data processing so that he could assume responsibility for an LBA tower in Brasília that was due to be reactivated. This raw data, therefore, was always maintained, even if no-one ever looked at it anymore, in case it became necessary to refer to it again or to re-do the process. Kept untouched in this way, the raw data, as discussed in the previous chapter, always retains its potential to be either error or data.

The raw data therefore arrives as an undifferentiated mass of unmarked columns of numbers. Once it has been renamed, this enormous amorphous body of numbers (as it appears to the untrained eye at least) is copied and pasted, as demanded by the protocol, into the first spreadsheet table ('raw data pre-analysis' or *bruto_pre*). This spreadsheet has sequential time running down the side, and the different variables, such as wind speed, rainfall and so on, running along the top. The spreadsheet therefore separates the raw data out into columns of different variables, and times. In order to ensure that the raw data has been copied across into the spreadsheet correctly, it is compared to the first column of the spreadsheet which is the 'control' column, containing the expected time of measurements by day, month, year, hour and minute. This is filled out before any data has been copied across. As the protocol stipulates, it is crucial for the "perfect synchronicity and filling out of the folder, as in situations such as interruptions in data recording, missing lines of data in the raw data file or the appearance of an error code by the storage device in the field, this control column does not allow these intervals to be omitted". Often, there *was* a mismatch – for example, the data extended below the edge of the table in the spreadsheet, or the times would not quite match up. J has what I nicknamed a "museum of errors", a document displaying examples of all the strange permutations that the raw data can arrive in from the datalogger – with missing sections, repeats, and numbers strangely bunched together, or strangely spaced out.

Although it is apparently unremarkable, banal, and executed extremely quickly, this initial transfer of the raw data into the first spreadsheet is very important. It is in fact the first moment in which the raw data's presumed ambiguity is revealed to contain meaning and error. There is a clear sense in which the raw data is being transferred *into* a pre-existing meaning, in the form of the temporally organized spreadsheets. Their structure gives a meaning to the mass of numbers by differentiating them as measurements of specific variables taken at specific times. It also provides the first clue as to whether something is wrong with the raw data set – which it often is. All deviations from the standard rows and columns have to be corrected, because otherwise, the data will be in the wrong place, and will not correctly reveal the meaning that is simultaneously inherent in it, and imparted to it by the spreadsheet. I suggest that the simple act of copying the raw data into the first spreadsheet is therefore the beginning of the process of semiotic domestication of the raw data, or its 'socialization'.[75] Being able to copy the raw data is crucial to this, and is possible only because the raw data is digital.

The Data Range as Self-Scaling Device

There are several other spreadsheets that are linked via the programme to the '*bruto_pre*' spreadsheet where the raw data now sits – not completely "raw" anymore, but far from "clean" either. These other spreadsheets contain the results of different functions performed on the raw data, in order to test its quality. There are several of these functions. The "good by sensor" spreadsheet shows the performance of the sensor according to how many measurements it took.[76] "Consolidated by day" is a spreadsheet showing a daily average value of each sensor, which gives a rough overview of the performance of the instrument over the day. The most important spreadsheet, however, is the "consolidated by sample", because it indicates any suspicious data in the *bruto_pre* folder, according to the "limits" that are in the "Range" spreadsheet. The Range spreadsheet contains the maximum and minimum values for each variable in the *bruto_pre* spreadsheet.

---

[75] I am borrowing this idea of "socializing" information from Alberto Corsin Jiménez (2005). Writing about the move towards increasing "transparency" in the knowledge economy, he laments that "[L]ike the pure, free gift, which makes no friends (Laidlaw 2000), transparent information and real time knowledge have no social life" (2005: 19). I would suggest that likewise, raw data has no social life. The section of the thesis that follows this chapter examines how through being 'cleaned' the raw data goes about, if not 'making friends', then making scientists and scientific sociality. As Cyrus Mody has observed, the relation between cleanliness and pollution or contamination is one of the "driving forces in the social life of a wide range of sciences" (2001: 8). In his study of materials scientists, he draws on the work of anthropologist Mary Douglas, and particularly her description of pollution as "matter out of place", (as it is exactly the science of material scientists to investigate what makes matter "in place") to investigate the creative properties of dirt in laboratory sciences.

[76] As the number of measurements – via their periodicity – is programmed into the datalogger, the total number per day that the sensor should have made is known. In the case of the AWS data, each sensor providing AWS data should make 144 measurements a day – if there are fewer than that, then something went wrong and needs to be verified.

The "consolidated by sample" spreadsheet flags any measurements that fall outside the maximum (MAX) and under the minimum (MIN) given in the Range spreadsheet. It also highlights instances where the difference between two consecutive measurements is too great (MAX. DIF), or if there is no data registered at all (SEM DADOS). This action depends on a logic that is written into the spreadsheet in a cascade of 'if…then' statements. If there is nothing next to the data value, then it is inside the range, and unsuspicious. The result is a sheet full of coloured, marked 'flags' indicating which values in the data set have transgressed the boundaries that the Range permits.

The range therefore indicates the limits of what is considered possible for each variable being measured. It is considered to be impossible, for example, for there to be more radiation leaving the surface of the earth than arriving – that is just a "physical law", I was told, so the "reflected radiation" could never exceed the "incoming radiation" (although as these are of different wavelengths the situation is a little more complicated than that). Other limits could be given by meteorological patterns, the most basic being that there cannot be radiation after a certain time of day (night). The composition of these ranges is thus an intriguing statement on the limits and boundaries - the outline, as it were - of the weather and fluxes being measured.

But what caught my attention in the construction and use of these ranges was the apparent circularity in this process – in as much as the range *itself* comes from the data set it is meant to be delimiting. R and J both impressed upon me that these ranges are different for different sites. Comparing the ranges for SGC and for K34, what is possible and impossible in these two locations differs considerably. In fact, J says, "the range belongs to the place" (*o range pertence ao lugar*); and deciding on its limits then is in fact "a really particular/specific (*particular*) thing (…) Each person has to adapt it themselves", as R told me. As a result, "ranges are not fixed. The ranges here [at the LBA] have changed several times". Progressive measurements have been added to the data sets for the different sites, and patterns slowly extracted. As R told me "you use the data to create limits (…) there's a physical part, a geographic part (…) you take the past record, four, five, seven years of data here, and then you look at it all together. You use the data to generate limits, and you use also a bibliography (…) They've already studied this, there is already an average curve". The head of the Micro group tells me that you need thirty years of (certified) data officially according to the World Meteorological Organization's standards to make a range, although there are people who use just two years. "You try to be generous", he explained. If you are not very generous with the variation, you will end up cutting a lot of data – and that is to be avoided if possible. "You only delete a number when you really understand, when you really know", as J and R said often and emphatically.

The first range used, I was told, was in fact a range worked out for K34. This was then modified for SGC and for the other towers. In order to construct this range, the members of the team compared the first data to other data sets for the region (for example from PhD theses or publications), and "personal experience of the fieldsite" in order to initially shape the data. You can go to the Brazilian National Institute of Meteorology (INMET), the head of the Micro group tells me, which has 100 years of meteorological data. You can look at published papers. You can look at what models predict. In the 1990s, people had much less of an idea what this range might be, and he is proud to tell me that the LBA was instrumental in decreasing this doubt. "Experience, literature, meteorological stations – you learn as you go along".

But the meteorological patterns that these ranges make manifest all come from data of one sort or another, even the most basic patterns and "laws". INMET archives, published papers, and models are all different forms of data. The range for the Brasília tower will be a modified version of the SGC range. J gave this range to the student who would be responsible for the tower, and he was adapting it to fit better the data from Brasília that he was cleaning. For example, the soil temperature can get much higher in the dry central plateau of Brasília than it can in the dappled and humid Amazon forest. But these specific limits themselves come from the same data that the student was meant to be creating the range in order to clean. He had to try to work out the limits from the data itself: with the radiation data, he had seen values up to 61° C, but has decided to put the limit at 55° - "I didn't have the courage to put 61°!" he told me.

The newly-arrived raw data set is not therefore compared to the 'world outside' in order to judge its veracity - that would be impossible, given that the phenomena of which it is a measurement is irretrievable. It is in fact compared to its own range, which is constituted by other data of which it is destined to be a part. The data is in this way applied back on itself, in the form of the range, as a sort of self-scaling device. The data itself is used to generate its own limits, and its own possibilities. This is a crucial aspect of the cleaning process, and I shall return to. However, it is complemented by another process: that of making the data relate to itself.

Making the Data Relate to Itself

The identification of potential errors as those that fall outside the range is the first stage of the data cleaning proper. As the protocol states, these suspicious values need then to be analysed "using as a base the knowledge of the researcher, and the knowledge of the characteristics of each research site, so that a consensus can be reached concerning whether to maintain or discard the measurement".[77] J referred to this as "hunting out the problems" (*caçando os problemas*): "I'm a-hunting!" (*Tou na caça!*) she would say to me on some mornings. J and R both spent an enormous amount of their time working out what was wrong with each data set, and then explaining in a document why they had decided to maintain the value as data, or remove it as error. Formally, this document is known as the "post-analysis" and as J also said, "post-analysis is our cross to carry…this 'why' is a lot of work!" They both found this work tedious sometimes, and difficult, but also, J told me, there was a great deal to learn from doing it. As R tells me, in the post-analysis, "it's the analyst who does the calculations. I take the 'consolidated by sample' sheet, and I look at the conditions, at the graphs, to see if it was really "minimum" or not. I'm not going to delete it, because it's real. Everything I do has to be justified...and the decisions are difficult!" One member of the Micro group, who was learning to clean the data from another tower at ZF3, was told only half-jokingly by the head of the Micro team that he had to be prepared to "woo" the data (*tem que namorar os dados*)-

[77] *Com base nos conhecimentos do pesquisador e nas características de cada sítio experimental, para que se chegue a um consenso entre manter ou descartar a medida*. Taken from the LBA protocol document.

The post-analysis documents often included graphs, photos, other numeric data, written explanations, and references to scientific texts. Each post-analysis document was different. These documents served as a sort of guarantee that there were defensible reasons for the decisions that had been made. This became important if those researchers subsequently using the certified data called them to complain or to ask them for explanations of gaps in the data (where errors had been removed) or strange values. "Sometimes," R told me, "there is data that we don't delete, and someone receives it and says it's strange (…) So we have to explain – why did this happen? Because it rained a lot then, the soil lost a lot of heat – we argue that it is right, that it's within the possible (*dentro do possível*).” The reasons could be diverse – a frog falling into the pluviometer; the instrument returning to normal functioning once it has been cleaned, indicating dirt had affected the previous measurements; an incorrect calibration constant written into the programme of the datalogger. The post-analysis document keeps a record of all this.

Ascertaining whether the data values are "within the possible" proceeds in several ways. There are values that can be quickly dismissed as a *falha*, a gap or an error, because the datalogger itself has a range in its programme that denotes values deemed to be impossible as "666" or "999" or some equally non-sensical number in this context. Certain of these nonsense numbers – that are then denoted NAN ('Not A Number') - indicate a loss in power to the instrument, for example. However, there are some cases that are harder to work out – the "grey areas" (*áreas cinzas*) of possibility, as one researcher put it. For example, there might be a sudden drop in a radiation value, and then a subsequent increase back to what was expected for that time of day in that season. The first way to investigate this occurrence would be to "check the relations between the measurements". In this case, could this low radiation measurement be explained by a sudden rain shower, causing there to be a cloud over the tower briefly blocking the sun's rays? If so, there should be a spike in the rainfall data, and an increase in the wind speed data, and a change in the pressure data, all at the same time. All these data values would then be compared, and if they demonstrated the right relations to each other, then there was good reason to suspect that the value was in fact "possible". The same could be done with carbon flux measurements and rainfall, in order to spot an error, or with radiation and soil temperature – and so on. In this way, through a progressive reconstitution of its internal relationality, the data is increasingly made to relate to itself. "We relate everything", J tells me.

The most patent manifestation of this was the means by which this reconstitution was performed – by graphs. "There are problems that only emerge when you plot graphs, others that you see immediately" J tells me. "If you plot this, you can see it's a problem with the time" J would show me; or "here, on the graph these two should show a relation and they don't". If they had a suspicious data value over a period of time, R and J would graph it against other relevant variables, and "look for the relation". "You go along, amassing reasons to believe, relating rainfall with radiation, for example", J told me. The latent heat had "passed the minimum" several times on one occasion, and so "when I saw the rainfall spreadhseet, I looked at the same times; when it rains, the temperature drops, and it's cloudy. For several hours, I saw it was very linked to the rain, through a graph that I plotted." She shows me a graph on which she has plotted all the variables. "On this day at this hour, there was a drop in radiation, of short wave, and that's cloud, and then rain – and then is when the sensible heat values went under the minimum – it's because of the sensor, do you see?" Graphs also allowed her to visualize the patterns in the data enough to investigate a suspicious value that has appeared several times. "We make a graph of the years, and compare them. We know that it rains a lot in São Gabriel (…) when we look at the graph, and see that this (the rain data) is really low, if it is sustained then there's a problem."

A graph, I suggest, is the revealed form of a data relation - the visible manifestation of the relations that constitute the data. By plotting one variable against another, the data cleaners temporarily reveal the internal relationality within the data, in order to ensure that it is there. If it is not - if there is too much noise - then there is a problem. In a certain sense, then, removing errors (data that cannot be related in this way), makes the data set *more* relational; it increasingly starts to consist of relations and patterns, rather than dissociations and "noise". But at the same time, this internal comparison again seems circular - it is the data that starts to affirm or disaffirm itself. Why trust the rain data, and not the radiation data? Even if it is in the range, it still might be incorrect - as R demonstrated to me one day by looking perplexedly at the graph of the CO_2 profile on the screen and telling me the "pattern was wrong". Even though it was inside the range, it was not exhibiting the right sort of relation to the time of day, and was therefore suspicious.

A "Feeling" for the Data

Knowing which variable to trust more is part of a wider "feel" for the data that J and R both had, and were developing. Going through raw data that had amassed from the now deactivated tower in Brasília, J would say "that one has the look of pressure data" or "that one has the look of the battery data" (*esse tem a cara de pressão*). On another occasion, R was going through some old data, and told me "this one has the look of a test, it doesn't have the look of data", using the same word in Portuguese – "cara" (literally "face"). This "feeling" for what the data might be also could suggest that certain data values were awry because of specific problems, such as a particular sort of error in the datalogger programme. The previous head of the Micro team who I also talked with frequently, would also impress upon me the importance of getting a "feeling" (he used the English word) for the research site and how the data corresponds to what you know can happen. He told me how he had acquired a feeling for the data from the Brasília tower: "for example, it's hard to get 100% rainfall there. Here, there's humidity - there, there's dust" and, talking of potential infestations, "it's not the bees there, it's the ants!" However, although J and R would sometimes go to the field (although J would very rarely go to São Gabriel, they would both sometimes accompany the other members of the team on a field trip to ZF2), they would mostly sit in the office in front of a computer. That is to say, what they were getting a "feeling" for, I suggest, was not exactly how the world works, but rather how the world in the data works, and how to make the world work in the data.

One way for them to check "outside" of the data is via the reports, *relatórios*, that accompany each set of raw data that arrives. As I have mentioned, personal experience of the fieldsite is considered crucial in trying to work out whether the raw data is data or not. The reports contains the time of day the data was collected, what the weather was like that day, sometimes what the weather had been like more or less for that week, the different tasks that had been carried out, if there was anything that looked especially dirty or had to be cleaned, or if there seemed to be anything wrong with the batteries, or the datalogger, or the pump, or anything else mechanical. As J was unable to get to the tower at São Gabriel, she had a technician who goes every week and collects the data for her. In J's words, "[he] is my eyes – he is very important…his report is very important. He describes the day, more or less…he isn't a meterologists, but he does it his own way (*ele tenta fazer no jeito dele*). It helps a lot." The technician's report, therefore, accompanies the data as its context – the world from which the data was extracted. I am using the notion of context simply to emphasize that without the *relatório*, it is sometimes very hard to make any sense of the raw data. Data that had been collected before the protocol was introduced, and before J and R arrived to work at the LBA is very hard to clean, they both told me, because they do not have these reports to guide them. J told me of one time that she had seen that the radiation data had been very low, suspiciously so. However, she was able to check the *relatório* the technician had sent, and there she read that the datalogger box had been cleaned which had caused a faulty connection, which the technician had subsequently fixed, so by now the instrument would have returned to normal. Without the *relatório*, it would have been impossible to work out that that was the reason that the data looked strange, or what to do about it (in this case nothing, as the problem had been fixed). The *relatório* therefore travels with the raw data as its context in the sense of being its 'background', or its environment. However, the *relatórios* are only relevant for the cleaning of the data; once the data is certified, the *relatório* is filed and stored.

The role of the outside world is therefore very important for the cleaning process, and on several occasions I was told that experience of the field site, how instruments work, the biological dynamic of carbon flux – as the previous head of Micro said, the ability "to see nature" (*enxergar a natureza*) - were crucial in being able to make sense of unruly raw data sets. The previous head of Micro was particularly emphatic about this aspect of data cleaning. In a series of didactic conversation he had with the student from Brasília about the tower, he stressed the importance of knowing the area that you work in – the biomass of the vegetation, the decomposition rates - "don't trust in the numbers alone" *(não confia nos números só*) he warned. He meant that when one is trying to process flux data, for example, the data cleaner has to understand how the ecosystem works as a whole. "You have to use the biology, the ecology to confirm your data", he explains, "don't just use mathematics" (*nao fica só na matemática)*.

The previous head of Micro had in fact been responsible for the tower in Brasília before it had been deactivated, and his knowledge of the site was impressive. He could remember the exact positions of the trees, and therefore the best places to install the sensors; and he had many stories to tell of things going wrong. A knowledge of the biological composition can reveal reasons for strange events in the flux data. There could be a massive spike in the photosynthesis the week before, and no alteration in your flux data – what does this mean? You have to ask yourself why you don't see the effect of one in the other. "Biological composition is indispensable" (*composição biológica é indispensável)* he insisted. This researcher is by training a biologist. He specialized in plant physiology, and is adamant that interdisciplinarity is the only way to go about understanding "nature". However, in order to do as he suggests, it would be necessary to also take measurements of the photosynthetic activity of the plants in the area around the tower – which, in fact, was the subject of his PhD, but would be very difficult to do long-term.

What is intriguing about this researcher's attitude is that, in fact, it is not a capacity to see 'nature' *juxtaposed* with the dataset in comparative relation that he is pinpointing. What he is asking the student he is teaching to develop is the ability to see nature *with* the data set; to somehow recreate nature out of numerical (graphed) relations in the data. Nature itself does not serve as here as an external referent, but an *internal* one. The "outside world" in the form of the *relatórios* is only employed during the cleaning process, but once the data is certified, these reports disappear altogether. They are not sent out with the certified data when it is requested, because if the cleaning has been successful, the certified data already has the world recreated within it.[78]

Experience and Trusting Instruments

[78] Bruno Latour mentions this in the chapter in *Pandora's Hope* that I have already discussed in the second chapter: "[K]nowledge does not reflect a real external world that it resembles via mimesis, but rather a real interior world, the coherence and continuity of which it helps to ensure" (1999a: 58). But he does not fully engage with this idea, instead reverting to the notion that there must be some sort of reference back to the truth. My proposal is that if the data is cleaned correctly, this in fact never has to happen.

In order to produce certified data, however, it is not enough to recompose "nature out there" in the sense of the biological environment that the previous head of Micro was implying. It is also essential to be able to see the errors with the data. This ability was often glossed as "experience", an accumulated knowledge of the possible reasons for the errors to appear the way they did, that could be drawn upon when confronted with a confusing set of raw data.[79] It could be that the calculation was being performed incorrectly by the spreadsheet. There may have been damage to the sensor, or dirt on the lens. Although they very rarely went into the forest, all the data cleaners needed to know what sort of relation the instruments could have with the raw data they were trying to clean. J explained to T one day, who was having problems with the humidity data on ZF3, how she had at one point stopped "trusting" her rainfall data from SGC. "If you know the sensor is having trouble, you start doubting the data", she consoled. She suggested trying to check the instrument - does it register an increase in humidity even with no rain? In her case, she said, she had to "throw the data away. It wasn't real".

Ascertaining what had caused the error in the raw data on your screen might also mean you could send someone to correct the problem, so that the next set of raw data did not suffer in the same way. Understanding the relation between the functioning of the instruments and the raw data it produces involves studying the manuals of the instruments in some depth. This is especially difficult for the carbon flux instruments because the method employed to make sense of their data is in fact encoded in a programme called Alteddy. J told me that she was shocked at the extent to which she needed to know about instrumentation. "I'm a meteorologist, I'm not going to mess about with wires!" she exclaimed. But little by little, she recognized that "it doesn't work like that. You need at least an idea (*noçãozinho*) beforehand. I find it absurd, how many times I have to read the Licor manual", in order to work out how you

[79] "The intricacy of this game of the possible is reflected in the personal experience of thse who do research...To feel one's way requires *Erfahrenheit* on the part of the experimenter. "Being experienced" as Fleck uses the expression, is not simply "experience". Experience enables us to judge a particular piece of work or a particular situation. Being experienced allows us to literally embody the judgement in the process of making new experiences ... Experience is an intellectual quality. *Erfahrenheit*, that is, acquired intuition, is a form of life... *Erfahrenheit has* to be learned, and it transcends what *can* be learned...It lines up with other attempts to do justtice to the "intimacy" of scientific work, to the exuberance of science in action, to what is beyond methodological axiomatization." (Rheinberger 1997: 77)

calculate CO_2 measurements. Each tower is a "parallel universe", each instrument a "world". "There are scientists who dedicate themselves just to understanding the instruments. Alteddy already changed many times – now it's on version 3.5 – and already encompasses many of the corrections that these scientists discovered need to be applied to compensate for the defects in the instruments. There are scientists who discover errors in the sensors and publish corrections – those guys involved right from the beginning, the 'tops' (*os "tops"*)". As I mentioned in the second chapter, understanding how the instruments worked was deemed so imperative by the head of the group that he organized a four day instrumentation course, inviting technicians from Brazil's Space Institute (INPE) who specialize in meteorological instrumentation to Manaus to teach the members of the Micro team - in fact, anyone in the LBA who dealt with data or towers - about the instruments.

Calibration factors are a potential source of instrumental error– and can create a situation where, for example, the "pattern" of the data is correct, but the values are outside the range. If this happens, the variable, for example CO_2, can vary in the expected pattern over the course of the day but at much higher levels than are possible, as the instrument is consistently overestimating it. This is the inverse problem to the one I mentioned before in which R spotted that the pattern was wrong, although the data was inside the range. The previous head of Micro told me of a situation where two different manufacturers provided different calibration factors for instruments that look exactly the same – so swapping one for the other without changing the calibration factor can produce inexplicable data. Another important numerical parameter that is written into the datalogger programme for some instruments is the "offset". J drew a graph for me to demonstrate how a particular sort of "jump" in the data normally means a problem with this offset value. The *relatórios* also include information of this sort: calibration factors, when instruments were cleaned (which might equally cause or resolve problems), what instruments had written into their datalogger programmes at different times. Thus one of the reasons why processing data from before 2008 is so hard is that there is no record of what instruments were being used, or what the calibration factors were in the programmes. Those who were in charge of the data cleaning at the time did not keep such records. As R said, that information left with them. Knowing about the instruments is vital to being able to recognize error in the datasets.

Sometimes, it was impossible to work out what was wrong with the data – there seemed to be an error, but its source remained elusive. This happened when I was there with the data for the profile of CO_2 and water vapour concentrations. R tried various solutions to see if the data would return to within the relevant ranges. The instrument, called the Li-820, was re-calibrated and cleaned. The datalogger was checked for errors. The batteries were checked, and a solar panel fixed. The Li-820 would sometimes start working again, and then return to producing impossible data. L, the electronics technician in charge of the LBA instrumentation lab, complained that every time he went to ZF2 to implement another solution they had devised, the instrument would mysteriously start working again - only to stop as soon as they had returned to the LBA. The whole team were used to this sort of inexplicable behaviour – "it's just the way it is", I was told. Once, J told me, the road that led to the tower at SGC was so bad due to the rains, that even though the data-storing computer on the tower stopped, the technician could not get to the tower to re-start it. So there was a long time with no data. Then, suddenly, it started working again. The technician had said that when he had finally managed to get there, by riding 15 km on a bicycle, he had cleaned the lens of the Licor, but without any type of calibration. Nothing had changed, and J was mystified – "how on earth can it suddenly start working again?"

During the time I spent with J, a particularly long-lasting problem with the rainfall data from SGC gave me the opportunity to see a number of different aspects of cleaning the data combined. J had noticed that the pluviometer data was lower than expected, and had been so consistently for some time. This prompted her to review the range. This is a difficult process, because trying to ascertain if the range needs to be adjusted a little – whether MIN is in fact the normal – requires one to be able to trust the rainfall data; and she didn't. In fact, she had been involved in a long-term battle with the pluviometer on the tower at SGC, in an attempt to work out whether she was looking at error, or data. "You can't just accept any old thing (*não dá para aceitar qualquer coisa*)" she said, "but you can't throw away data. The question is up until when is it error. Experience is very important, you go along (...) relating rainfall with radiation, for example". She checked the SGC rainfall against the INMET data for rainfall, but even though there were clear discrepancies, this was not conclusive as the INMET weatherstation is 60 km away from the tower. The story she told me demonstrates a very typical weaving together of the various aspects of the process I have talked about:

"The measurements were very low, for a region like São Gabriel where it rains a lot, it was really below the expected. It was as if it was a drought, really dry. So we made some graphs, and it was like this, the average was here [drawing a line to show me] and it [the SGC data] was going along beside it, when suddenly it would do this [indicating a little drop with her finger] and then from then onwards, with time it got worse, and we started to mistrust it. But we took some time to mistrust it, because during this period, coincidentally, it *was* raining less - so we thought it was normal. But it continued, and we started to suspect something. That's when I got the pluviometer that was here [in the Micro instrument lab] and sent it to LIM, and they calibrated it and that one from São Gabriel was sent here, and [a technician] did a test with multimeter and it looked like something had short-circuited. So he was going to see if the part that is needed is available in Manaus, but I don't think it is, it's pretty specific (...) We also saw there was an error that had been there probably since the beginning, when it had been installed on the tower - someone put the wrong calibration constant. [The previous head of Micro] already knew how to identify this because he had already experienced it, using this model of pluviometer. And one day I just mentioned [the problem] to him, and he told me. And I discovered that that really was the case!"

Alongside show-casing the various aspects of cleaning the data already described, what I would like to highlight from this tale of the pluviometer is the extreme uncertainty that J experienced, when she could not tell if the data was data or error. The raw data, pre-cleaning, contains exactly this ambiguity, and one of the ways it manifests is as a lack of *trust*; one of the most important results of successful data cleaning is that the certified data can then be trusted.

Making the Data Real

Thus far, I have explored in some detail the various processes that constitute "cleaning the data". The first was the naming and organization of the raw data as directed by the protocol. This, I argued, is the initial action that decomposes the raw data's ambiguity into either data *or* error. I then described how the range turns the data into a self-scaling device, defining what is "within the possible" through a process in which the data is brought to bear continuously on itself in order to determine its own limits. Alongside this process, the data is also made to relate to itself internally. It is by making sure that it its constituent parts all relate correctly to each other that noise and errors can be localized and removed. The removal of these errors makes the data more relational, and more internally coherent. The success of this activity depends on the data cleaners developing a "feeling" for the data, in order to be able to see nature *in* the data; and to gather experience, especially of the instruments, that allows them to see not only the world, but errors and noise in the data as well. Thus the way that the LBA data cleaners ascertain whether their raw data is "data" or "error" is by comparing it to other data; there is no reference "back" to the world from which it came, but there is reference internally to itself, both through the range and through the other data cleaning practices. The *relatório* that refers to the world that the raw data was extracted from is important in this process, but it becomes less and less important until it is eventually discarded altogether. The sources of error can sometimes be established through applying external knowledge of how instruments work or are failing to, but the data cleaners accept that there are many mysteries and inexplicable occurrences that no-one is able to explain, and take this as par for the course. The certified data that emerges from this is (or should be) data that can be "trusted".

I propose that what is happening in semiotic terms during the LBA data cleaning is that the certified data is being made to be its own referent. The raw data's ambiguity is split into "error" and "data". The "error" joins the rest of the world - that which can be discarded - and the "data" itself starts to carry *internal* to it the relational position of "referent/real". It does not refer "outwards", but "inwards", and the work involved in data cleaning is the work it takes to make this possible. The lack of external reference or comparison 'backwards' to the world is from this perspective not so much a flaw in the process, but integral to effecting this crucial transformation. I have suggested that as part of this, what might be called "the outside world" is transformed into a mobile "context" that can accompany the raw data in the form of the *relatório*. But if the data cleaning is successful, this outside world then disappears altogether. This outside world, as I hope to have demonstrated in the first and second chapters, is not itself the "reality" that interests the data cleaners, exactly because there is no question as to its status as real. The world just *is*. The question as to what is reality and what is not can only be asked once that world has been transformed into a discrete version of itself through the act of relating to it - through measuring it, as discussed in the second chapter. It is at this point that the potential for error or data comes into being in the form of the raw data investigated in the third chapter. However, the raw data itself is only in fact the possibility of the relation between data and error. It is the ambiguous substrate from which this relation will be elicited through the cleaning process - through self-scaling, which is here self-referentiality and self-relationality. [80]

Despite this circularity, it would be incorrect to state that in the process of being made to be its own referent, the certified data therefore becomes, in the eyes of the researchers, fictional. On the contrary, it becomes more real than it was as raw data, even though it contains "fewer" (of the wrong) connections to the world. I suggest that rather than implying a lack of reality, comparing the data inwards is exactly the process that (re-)creates the world *inside* the data set, and it is this which assures its reality. It is as if the relation of 'correspondence' is turned inwards. The certified data carries with it its own reality, not because it corresponds to the world but because it does not need to refer back to the world. Once it is certified data, it is real. There is of course a grey area as far as the data cleaners are concerned, when they are not sure whether a certain data point is data or error, but as far as the certified data goes, this grey area does not appear; the only options are for it to be data or error. The post-analysis document that contains the reasons for discarding or maintaining the raw data as error or data does make it possible to summon these grey areas, if needed. But, ideally, that should not be necessary at all.

Annelise Riles has described what I believe is (up to a certain point) a similar situation in the judicial world for "legal fictions", such as the fictions that the adoptive parent is the biological parent of an adopted child, or in the financial world, that of collateral (Riles 2011: 173). The question she asks is why do people 'believe' in such

[80] Here an observation made by Strathern is apt: "what validates one fact are other facts - always provided the connections can be made to hold" (Strathern 2003: 174, paraphrasing Hume).

fictions? Riles' answer is that it is not so much a question of belief, as of these fictions coming "to take on a cultural reality of their own", becoming "practical, technical scripts for the management of the parties' relationship" (ibid.: 174), irrespective of whether they correspond or not with what could be considered "reality". That is, they are "non-representational" (ibid.:173). Riles suggests these fictions are "techniques, tools, means to an end" (ibid.:173) because they are simply placeholders, in order for negotiations to go ahead. They have a consequence, but not a representational meaning, because they do not correspond to a reality. There is certainly a clear resonance here with my description of the process of cleaning the data. Clean data does not refer back to the outside world in order to gain its status as trustworthy or useful in the production of scientific knowledge. But at the same time, the data is expressly not a fiction in the sense Riles describes, because the representational relation is still very present. Reality does indeed matter for the data to be considered good data. The certified data definitely does have a meaning that corresponds to a reality. However, what the cleaning process does is to separate 'reality' from the 'world', and then recompose that reality within the dataset.

Bruno Latour has also provided a fruitful comparison of scientific and legal practice (2004b). One of the relevant distinctions he makes between legal and scientific practice concerns the different ways that "the facts" figure in the two. In the case of the law, Latour argues, lawyers try to get rid of the facts in order to move towards the "particular point of law that is of interest"; in the laboratory, on the other hand, the fact "occupies two somewhat contradictory positions: it is simultaneously that which is spoken of, and that which will determine the truth of what is being said about it" (2004b: 89). In as much as the "particular point of law that is of interest" becomes in this way a judicial tool, Latour's description chimes to a certain extent with Riles' description. But I suggest that the two "contradictory" positions held by the fact in the laboratory are not contradictory at all. The fact - like the certified data - but an example of the fact - like the certified data - is both referent to and representation of itself. There only seems to be a contradiction if one imagines that scientific practice holds these two apart all the time.[81] The non-representationality of certified data, therefore, should not be envisaged in the terms so often suggested by social scientists,

[81] Or, as Martin Holbraad poses, it may present "a properly ethnographic paradox - a condition with which [informants] themselves must reckon" (Holbraad 2005: 243).

in which "any referent, upon closer inspection and as soon as we try to get hold of it, is turned itself into a representation" (Rheinberger 1995: 51). As this cleaning process in fact starts with what is conventionally understood as a representation of the world (the raw data), it might be more appropriate to suggest that, in the case of the LBA at least, scientists are involved in grabbing hold of what appears to be a representation, only in order to turn it into a referent. In the case of the LBA data, 'realities' are not turned into mere 'signs'; on the contrary, great effort is expended to ensure that the signs can carry with them their own realities.

Compaction

The preceding sections have traced the processes by which the relation between real and representation comes to be contained within the data, so that certified data is finally neither real nor representation -it is rather what I will call the "compaction" of the two. Donna Haraway illustrates this notion of compaction evocatively when talking about the gene. She writes that "within both the organic and the synthetic databases that are the flesh of life itself, genes are not really parts at all. They are another kind of thing, a thing-in-itself where no trope can be admitted....to be outside the economy of troping is to be outside finitude, morality and difference, to be in the realm of pure being, to be One, where the word is itself" (1997: 134). The term itself however comes from Marilyn Strathern, who picks up Haraway's powerful imagery in a discussion of the history of another analogic dynamic - not exactly between information and nature, but between knowledge and kinship. Strathern suggests that these two have, since early modern times "provided figurative ammunition for each other" (2003: 184). Tracing the way in which these domains have thus drawn upon each other in British sociality, Strathern draws particular attention to the times when knowledge and kinship cannot be told apart, when "it is as though analogies between [them] were compacted into one another" (ibid: 185). The work done by the single term "relation" in each domain captures this notion of compaction, as does, in Haraway's formulation, the idea that the genome *is* a code. Relations are related, and genetics is an idiom in which nature is information. There is no issue of correspondence between the two because there is no "co". However, conversely, this does not mean that this co-relation cannot be elicited subsequently. Haraway points out that this compaction means that "no trope can be admitted", it is the "god's eye view" (1997: 134), a dangerous perspective to assume. Although she sees this collapse of the potential for troping - and the "fetishistic" attachment to a singular reality that accompanies it - as alarming, I suggest (more in line with Strathern's rendering) that, at least in the case of the LBA data, this compaction is merely one phase in a series of compactions and separations that constitute the production of scientific knowledge.

I propose that what we are looking at, when we look at the 'representations' of the LBA, are a series of compacted entities that only mean anything because they give rise to the potential for their own separation. Rheinberger writes that "representations are generators of bifurcation" (1995:83), meaning that the production of scientific representations results in the production of deviant or surprising results. Although the bifurcation between data and error is one such bifurcation, I would suggest that more generally these forms generate *their own* bifurcation internally. Taking a look at the whole process I have described illustrates this. The ‘world’ as an undifferentiated substrate is separated ("discretized") through measurement to give uncertainty (the gap) and raw data (the value); raw data, as an ambiguous material, is then separated through data cleaning to give error and certified data; certified data is itself a compacted form of the relation between reality and representation. Undifferentiation, ambiguity and compaction emerge in between different moments of separation. It is not that any of these stages - world, raw data, certified data - are either real or representations, but that they are a related series of different separations and compactions. Where they are collapsed, they are often taken as "reality", although this is *not* to say that they assume the position of the world. The purpose of compaction is to leave the world behind and allow the certified data to travel away from its site of collection.

The dynamic I am sketching out resonates with Bruno Latour’s insistence that we recognize that in science (and the world in general) processes of purification leading to pure forms and processes of hybridization leading to miscegenated ones, engender each other (Latour 1993 [1991]) [82]. However, there is an important difference to note. Compaction is not hybridization or "translation" in a Latourian sense, that "creates

[82] See also here Roy Wagner's dynamic of invention, convention and obviation, which indirectly influences this argument; eg Wagner 1981 [1975].

mixtures between entirely new types of beings, hybrids of nature and culture" (Latour 1993 [1991]: 10). Rather, compaction is the collapse of one side of the relation into the other. The data is certified when it is neither more nor less than the world, when it is life-size, and when there is nothing behind it, and nothing beyond it. When all the noise has been removed, and when it does not have to correspond to the world because to all intents and purposes, it *is reality*. It should be characterized not as containing a mixture of nature and culture, but as having reached a state in which nature and culture are indistinguishable. Separation is the subsequent drawing apart of the terms of the relation, and is thus the realization of the potential for meaning that is inherent in the collapse. As I was told on several occasions, the certified data is real. But, at the same time, the data is not the world. And neither is the world exhausted by data - at the very least, the world always escapes it to a certain extent. There is always error and uncertainty - the constitutive mismatch between reality and the world, as it were.

There is another result of the cleaning process, related to the 'trustworthiness' of the data. Data is "clean" data is when it demonstrates all the relationality it can, to itself - not to the world outside, but as a world *inside*. This relational coherence could also be seen as ensuring that the data in a certain sense 'agrees with itself'. When the data agrees with itself fully, this means that it can be inspected from every angle and it will look the same – it becomes a singularized object. It also, however, becomes a *singularizing* object. There is certified data that no-one will question or complain about when they come to use it, because it is coherent. What everyone agrees with is in fact the data's internal agreement. That is, in this case the certified data becomes an object that does not *allow* for any change in perspective - it enforces acquiescence. Its compaction has been successful when there is no need to dispute - that is, no reason to elicit or create - its relation to reality, as I have argued, and if there is no dispute then there is singular agreement. It is this singularizing capacity that allows the certified data to subsequently travel to other scientists and into other comparisons. Like Latour's notion of immutable mobiles (1987: 227), the certified data is mobile precisely because it is fixed in meaning. This is what being "trustworthy" means in this context. However, the certified data only moves in order to be separated out again, this time as it becomes entangled in different models of property and ownership. This separation is the topic for the next chapter. Before moving on, however, I would like to pause to tease out some further consequences of this process of "cleaning the data" within STS theory, especially as concerns the idea of circular reasoning or "the regress" in science.

Reassessing the Regress

The foregoing analysis of "cleaning the data" can be employed to reassess some aspects of constructivist STS theory, as there is certainly a kind of "creativity" involved in cleaning data. This is distinct also from the prospect that anyone might be "cooking" or "rigging" the data (the general occurrence of which in science my informants were aware of, of course). Cleaning data is a process that involves a substantial amount of work and requires that the data cleaners' knowledge be brought to bear on the data. This does not imply that the data, in their eyes or in those of other researchers, has been manipulated. Certifying the data certainly does not mean adding to the world.[83]

There are different notions of creation, or construction, to be explored here. Appropriating Michel Foucault's term "care of the self", STS scholar Mike Fortun employs the term "the care of the data" to refer to the "creative and tacit epistemic virtues" (2011: 19) that characterize the relation that geneticists have to the "avalanche" of data that is now the norm in contemporary biotechnology. Fortun intends "care of the data" to signify both an ethical engagement with the data as an "Other", "respecting its alterity and mutability" (ibid: 18); and as an attempt to turn away from a reductionist deconstruction of scientific practice. This move is one in fact, as Fortun tells us, that acknowledges the *affinities* between science and deconstructionsim. Thus a relationship of "care" is one that allows the scientist to grasp the data "in its entirety in order to shake it, querying and rattling the whole fragile yet sturdy structure to find its play, its surprises, it excesses" (ibid: 19), as an algorithm that analyzes a data set may produce more questions than it answers. Fortun thus emphasizes not the consructivism but the *creativity* of genetics, and locates this in what he terms "biosemiotization", the coaxing out of new pathways for signals to

[83] As James Leach writes, "through their labour, artists add to culture", whereas scientists find connections in the already-given (2011: 156). One might rather say that through their labour, scientists do *not* add to nature - this is the task, to *not* add.

cross interfaces that span the material and the semiotic. This focus leads him to draw out an image of the generative but interminable aporia of the defining boundary of genetics, that between "the gene" and "the environment". He writes: "There is no getting beyond the gene-environment impossibility, but only a careful taking of the next step in the eternally recurring play of the scientific game of creation" (ibid.: 47). As Rheinberger would have it, this implies not a tendency towards equilibrium, but a "never-ending ramification" (Rheinbrger 1995: 50); science is nothing if not a "generator of surprises" (Rheinberger 1997). This then is an image of creative excess that is always lacking, and yet always sufficient, because there is no totality to aspire to; an open-ended process, leading "to unprecedented excesses which cannot be anticipated" (Rheinberger 1995:88). The notion of "representation" does not capture this constant displacement and generation, because "the activity of scientific representation is to be conceived of as a process without "referent" and without "origins"" (ibid.: 51).

However, the same critique can be leveled at both Rheinberger's and Fortun's portrayals of science: if there is "no complete account" possible" (Fortun 2011: 47), and no referent, then by what criteria is knowledge produced? Scientific practice and theory are left without a bedrock of 'reality' underneath them. This is the issue that Harry Collins presents through the "experimenter's regress", in which

"the problem is that, since experimentation is a matter of skilful practice, it can never be clear whether a second experiment has been done sufficiently well to count as a check on the result of the first. Some further test is needed to test the quality of the experiment - and so forth...the failure of these 'tests of tests' to resolve the difficulty demonstrates the need for further 'tests of tests of tests' and so on - a true regress" (1985: 2).

This, Collins proposes, demonstrates that "'epistemological criteria' alone cannot resolve disputes in science" (Collins 2002: 186) and a resolution must be sought outside of the experiment in question, in the form of "various non-experimental and 'non-scientific' activities" (Collins 1985: 100) that can be employed to break the circle. We are here a long way from the creativity with which we started; in fact, this might even be "Science's nightmare: the example of a mode of unfettered arbitrariness in which a closed assembly decides, without reference to any external arbiter, with no tools other than words, and by simple consensus, what should be held as the truth" (Latour 2004b: 38). Observing a similar dynamic but pointing to its stabilizing tendencies, Ian Hacking points to the "self-vindication" of the laboratory sciences, that is a result of the fact that "as a laboratory science matures, it develops a body of types of theory and types of apparatus and types of analysis that are mutually adjusted to each other" (Hacking 1992: 30) creating a "curious tailor-made fit between our ideas, our apparatus, and our observations."(ibid: 50). Established laboratory sciences are therefore closed systems, with each component in alignment, and it is this, and not the cohering power of some underlying reality, that accounts for the surprising accumulation of stable knowledge that we see in the sciences [84].

In both of these characterizations of scientific practice, what is brought into relief is the circularity of the logic by which it operates. The description and analysis of the data cleaning process that I have presented permits an alternative interpretation of the circularity I encountered in the field, that requires neither its resolution nor privileges its stability. The interpretation I wish to put forward instead emphasizes its creativity in the mode of Fortun and Rheinberger, but with a distinctly different inflection. In Collins' depiction of the regress, circularity is presumed to be non-generative: it is only by breaking it that knowledge can be stabilized. I have argued on the other hand that it is exactly the regression that appears in the cleaning process - that data "cleans" data - that creates its potential to be *real*. Bringing outside knowledge to bear on the issue, and developing a "feeling" for the data -in short, the negotiations that surround the data cleaning process - are all with a view to recomposing "nature" within the data set itself. Although I do not doubt that those I worked with are aware of the potential for the "non-scientific" to influence the production of scientific knowledge, I would argue that Collins does not need to "look outside" of science in that way in order to find a way to explain scientific practice. The circularity is itself an integral and instrumental aspect to the creation of this entity called certified data. This also then is a more radical version of Hacking's understanding of the role of circularity, or closed systems, in scientific practice. Whereas he proposes that it is this stability that takes the place of reality, I argue that for the data at least, the circularity *generates* reality in

[84] "A coherence theory of truth? No, a coherence theory of thought, action, materials and marks." (Hacking 1992: 57)

the dataset. It is not the stability exactly that interests me, as it does Hacking, but rather how the data comes to take the place of the world altogether. It only does this because it becomes real, *whilst leaving the world as world*. Thus my description also departs from Rheinberger's suggestion that science has no referent. I have argued that what the circularity, and the self-referential and self-scaling dynamic, does is to compact the representational relation such that it is understood to be reality itself. It is this compaction of the data, that is, its capacity to be its own referent, which allows it to travel and gives it a relational capacity that it did not have before. The cleaning is not only a transformative act of creativity, but that which it creates is also itself given the capacity to be creative.

I have described on the one hand the separation of data from error, of true from false, of world from data; on the other, the successive compactions of the two terms into each other. I am talking about a transformative chain, nonetheless. I suggest that this sort of transformation is akin to those documented in other situations, often in laboratories, through which "animals" become "scientific objects" (Lynch 1988; cf Stengers 2005, Despret 2004). This ontological alchemy is I think common to both field and laboratory science, although it takes different forms. Thus Michael Lynch describes the process by which "laboratory animals are progressively transformed from holistic 'naturalistic' creatures into 'analytic' objects of technical investigation" (Lynch 1988: 226). This is a transformation by which it is established that these animals do not have minds, and therefore cannot share that which makes us humans with us - a move that might be glossed as the "objectification" of a sentient being. At the same time, one might argue that this process requires what Stengers calls the "anaesthetization" (2005: 997) of the scientist, in order to "protect" him or herself from the recognition of the suffering he or she is afflicting. I would as counterpoint like to end this chapter by offering the process I am describing as an inversion of this. In the case of the LBA, and perhaps other data-driven field sciences, the transformation is not one of "objectification", but "socialization". The data is being socialized, in as much as it is being transformed from an entity that has the wrong sorts of relations and no singular meaning, into an entity that can relate in prescribed ways to other entities - subjects and objects, scientists and other data sets. But in contrast to the depersonalizing actions of the scientists who must numb or protect themselves from their own actions by de-subjectifying their subjects of study, here the symmetrical counter move to this socialization is that it bestows upon the *data* -

the socialized object - the capacity to create different sorts of scientific subjects. Whereas objectifying test animals seems to require scientists to objectify themselves (although see Candea (A) and (B) forthcoming), socializing data bestows it with a capacity to generate specific scientific subjectivities [85]. This, it seems to me, is a particularly interesting form of creativity, that I shall discuss at length in succeeding chapters.

Conclusion

In this chapter, I have presented the process by which the LBA data is cleaned. I have suggested that this cleaning is constituted by a process of self-referentiality and circularity that is fundamental to the production of certified data. This is because it creates reality, or reference, inside the certified dataset. Outlining a dynamic of collapse and separation, I suggested that the dataset comes to carry within it then a compacted (collapsed) relation between representation and reality. This simultaneously marks the data off as "real", but not "world". If this cleaning is done successfully, then the certified data *is* reality, and this property allows the datset to be trusted, and therefore to travel to other settings. Based on this analysis, I suggested that regression or circularity in scientific practice be recognized for the creative effects, rather than the constructivist implications, that it has. One of these creative effects is the capacities it endows the certified data with, and it is this that the next chapter describes.

[85] This analysis also offers the opportunity for an interesting comparison with non-western techniques or practices of what I am calling here "ontological alchemy", or the transformation of what the West might think of as subjects into objects or vice versa. I have in mind the literature on the Amerindian production of "kinship" in contradistinction to "affinity" (Viveiros de Castro 2001, Vilaça 2002) and Amerindian shamanic practices that ensure that one is not eating one's kin when one eats certain game animals (Viveiros de Castro 1998), as well as the practices around funerary cannibalism in the region (Vilaça 2000).

PART II - THE SOCIAL LIFE OF DATA

Chapter 5: Doing Difference with Data

Introduction:

The preceding chapters of this thesis have described the production, processing and certification of data within the LBA in a more or less localized and linear fashion. This is not for heuristic but for empirical reasons - the movement from collection to cleaning is along a well-trodden, if sometimes obscure, path into and out of the forest. At this point, however, the story becomes more difficult to follow in a sequential fashion. After the tower data is certified, the next task of the Micro group is not to analyze it or publish with it, but to store it and make it available to other researchers.

This dissemination, however, does not depend upon the data cleaners alone; the data also needs to be solicited by others in order to travel in this way. As it is copied and sent, the potential appears for the certified data to multiply exponentially. Although it is the cleaning that permits the certified data to move in these ways, as it only becomes trustworthy enough to travel once it has been certified, the timing and nature of its movement now involve those who had previously been external to the process. Tracing the outward movement of the certified data is therefore methodologically challenging, and for reasons that are ethnographically salient: suddenly, a multitude of potential "others" appear, in many guises (even if they were known to be there all along).

This chapter investigates this dissemination of the data from the LBA. Two models of data exchange are identified, one that is premised on a logic of “flow”, and another on a logic of “exchange”. I argue that these are not two versions - one ideal and one misfiring - of the same system of exchange, but two separate systems, that create two very different scientific communities. This realization shows sociological descriptions of scientific exchange systems as analogous to investment markets to be limited in their applicability. It also sheds new light on the frictions that emerge during collaborations in which these two models, or others similar to them, operate side-by-side. What is in question, I argue, is not only the emergence of separate scientific communities, but also different scientific objectivities and subjectivities.

I turn here to the literature on scientific property, which also draws on notions of creativity and subjectivity. I argue that rather than simply being "truth" and therefore an inappropriate item of ownership, the data has certain properties that refigure such notions as creativity and ownership altogether. Through being exchanged, the data becomes not only what it is but also the potential to be transformed into something else - knowledge. It also gains the power to transform those who are themselves working this transformation.

The CPTEC Database: "Flow"

The LBA, as I was told by the LBA Data Information System manager in Manaus, has two official data repositories. One is in the Centre for Weather Forecasting and Climate Studies (CPTEC), which is part of INPE, the Brazilian Institute of Space Research. The other is in the USA, at the Oak Ride National Laboratory Distributed Active Archive Centre (ORNL DAAC).[86] The same data is held in both databases, and both are governed by the same data policy. At the end of my fieldwork in Manaus, I spent one month at the first of these locations, in CPTEC. During this time, I had the opportunity to talk to Luiz M. Horta, responsible for the LBA CPTEC database. Luiz also gave me a brief tour of the storage facilities. A lot of the older LBA data is archived on tapes, rather than on hard drive. These tapes are held in a large case in an enormous, freezing, humming room in CPTEC, which also houses several supercomputers that run climate and weather models. The data that is held in this database - many terabytes of it - is freely available to anyone who wants to download it. When the older data is solicited, a robotic arm suddenly

[86] This is one of the National Aeronautics and Space Administration (NASA) Earth Observing System Data and Information System (EOSDIS) data centers managed by the Earth Science Data and Information System (ESDIS) Project" (http://daac.ornl.gov/, accessed 6th July 2012, link provided by Luiz M. Horta).

appears and whizzes alarmingly around the tapes, selecting the necessary ones to copy the data from. As one of the documents that Luiz provided me with explained, "[T]he LBA Data and Information System is a data management system acting as a repository for all the LBA data. The data is quality checked, rendered to a and made available to the LBA community as rapidly as possible and transferred to a permanent archive. To facilitate use by non-LBA investigators, each data set is carefully documented and linked in the orderly framework, so that it remains useful after the project has been completed."[87]

When I asked if Luiz had any idea as to who uses the data, he showed me the LBA data management site, again freely accessible online. This site shows every single download made of the LBA data from the CPTEC database, the IP address of the downloader, and the country that the IP is in.[88] It also shows the percentage of total use assigned to each country. A cursory check for March 2011, for example, shows the predominant user to bethe United States (79.68%), but there is also Brazil (12.47%), Australia (2.07%), and several countries with under 1%, such as Korea (Republic of), France, Germany, the UK, and Portugal. The immensity of the network constituted by the circulation of the LBA data is evident. It travels all over the world, and has been used in hundreds of publications, as I realized when the secretary of the LBA in Manaus presented me on request with a list of publications that stretched for several pages.

"From the moment that the data leaves here, leaves the database at CPTEC, we don't have control" Luiz told me. "But the data policy of the LBA specifies that the data is public. The Brazilian government makes the data available, without any cost, free. And no questions asked." The formal LBA data policy "establishes procedures for data sharing, citation of data from other investigators, access from investigators to restricted data and promote [sic] the exchange of quality controlled / quality assured data".[89] Drawn up at the inception of the LBA by the Steering Committee, it has several stipulations, so "that cooperation and synergism should be maximized in all LBA activities" [90].The data can

[87] Taken from a document provided by Luiz M. Horta, originally intended for LBA data users.
[88] http://lba.cptec.inpe.br/lba/lbadis/FTP_Usage/usodosdados.htm Accessed September 2012
[89] ftp://lba.cptec.inpe.br/LEIAME-README.txt link provided by Luiz M. Horta. Accessed September 2012.
[90] See footnote 3.

only be in possession of its creator for a maximum of two years (although normally only one) before it has to be made available, and this is strcitly policed. Luiz told me that "if someone is part of the LBA and doesn't make their data available, the next time they apply for money for their next project, the NSF [US National Science Foundation] consult me - is this person's data up to date *(de acordo)*? And several times, I have said no, they haven't sent the data. They have already published papers, but haven't sent any data at all. And so this person is not considered for more funding. This has already happened with two people." I asked if this was the same in Brazil. "Not with the same intensity," Luiz replies, "but it's a question of time until we change our ideas here in Brazil. Because there is still a lot of that thing here of researchers not allowing their data to leave their computers. It's insecurity, this is just a question of scientific maturity, one day it will happen."

Although the LBA data policy to which Luiz is referring stipulates that the original data set must be kept in Brazil because mistakes can be generated by the copying process, [91] there is also no apparent distinction made between insides and outsides as to accessing this data: "outside investigators may be given access to this data as soon as the data has been submitted to LBA DIS, after a short period for quality control".[92] Furthermore, "there will be no periods of exclusive rights to publish LBA results", [93] with the exception of PhD theses where publication is prohibited prior to acceptance. Openness is also evident in the recommendations concerning co-authorship. Data must be analyzed cooperatively, and publications resulting from LBA research should be co-authored by all scientists who have "substantially participated" in the work. [94] In one of the documents that Luiz passed on to me to read about the LBA-DIS database, I found the following diagram:

[91] Although when talking to the head of the ORNL DAAC LBA database, he told me that it was a legal vestige from a pre-digital time that demanded that the "original" be left in Brazil; and in fact, the original digital data was seldom left, but a "copy" was.
[92] See footnote 3.
[93] See footnote 3.
[94] "Where data is used for modeling or integrating studies, the scientist collecting the data will be credited accordingly, either by co-authorship or by citation. Researchers using data provided by another investigator as a considerable component of a paper should offer the originating investigator co-authorship. In cases where data from other investigators represents a secondary contribution to a paper, the data should be referenced by a citation. Users of the data should always have to state the source of the data." LBA data policy, provided by Luiz M. Horta.

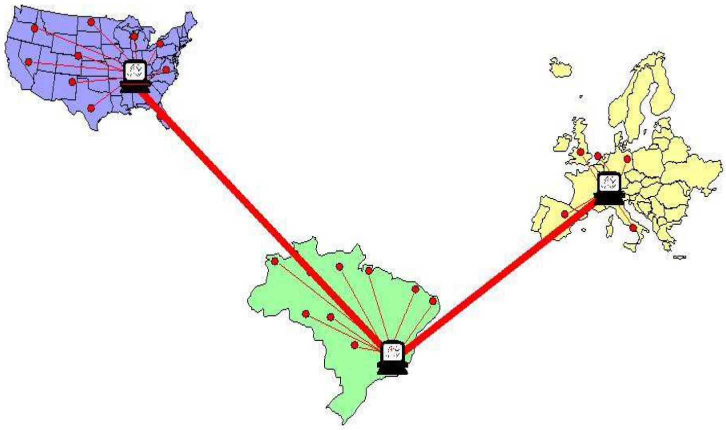

Figure 2: Diagram taken from the LBA-DIS documentation provided by Luiz M. Horta.

What my conversations with Luiz portrayed, backed up by the substantial documentation that he provided me with, is perhaps an archetypal notion of free data sharing and scientific collaboration. Such promises of the digital information age have been intensely critiqued and scrutinized by social scientists (Hilgartner 2012; Edwards 2010; Zimmerman 2003; cf Castells 2007). Despite this, I would like to sustain them here as integral to the way that the CPTEC Database operates. Data is accessible through international centralized nodes; access is free; and the data is in the public domain. Everyone and anyone can own the data, and everyone and anyone can publish using it, with due reference to the data's source. After one year, the data, to all intents and purposes, no longer has an owner, only an origin. As the diagram indicates, this sort of system does not respect boundaries between countries. The CPTEC database both proposes and embodies a generalized sense of community, collaboration and informational exchange governed by open access. Luiz even told me a story to emphasize the point. Aptly, he had himself been told it by a Russian who he worked with on another project. The crux of the story was that during the Cold War, when relations between the USA and Russia were at their worst, the scientists of the two countries still shared meteorological data. Luiz was struck by the capacity of science to overcome boundaries "even in war time!"

The scientific community that the CPTEC database brings to mind is an international one that has equal access, and equal rights. Ownership is limited within the parameters of the system, and is purely functional, in as much as it does not endow you with more rights of use than a non-owner. A corollary of this data exchange system is that it proved almost impossible for me to actually follow the data as it was being used. The network was so vast and the form of solicitation was automated, so there was no hope of my being able to have a sense of who used what data and why, beyond what was provided by the automatically-updating site that displayed nationalities, and bytes of information - and no more. The documents that Luiz provided me with stated how it should operate, and he told me that sometimes these imperatives were transgressed. These transgressions, however, do not effect the logic of the system itself, based on the free flow of data between countries and researchers.

The LBA community is here in a sense *automatically* generated by the database and what it embodies. Like other databases, the one at CPTEC is a "condensed site for contestations over technoscientific versions of democracy and freedom" (Haraway 1997: 130), but in the model as presented to me by Luiz the area of contestation seems remarkably small. Although Luiz mentions that Brazil does not quite manage to live up to the model's criteria at present, it is only a matter of time. Instead of dismissing this as merely a form of idealism on the part of the data manager, I would like to sustain the CPTEC database as something to think with. What seem to disappear in this model are its own limits; it makes a point of dissolving its own "outside" and "inside". The only people who are excluded are those who exclude themselves, by not sharing their data. This makes it appear limitless. As the data is always being transformed [95] into multiple forms (publications, conference presentations, comparative studies), the point, one might say, of this system is 'flow' itself. This is a logic that suggests, as several scholars have discussed (for example Biagioli et al. 2011: 7), that what scientists produce is always *already* in the public domain, and the point is merely to make it as accessible as possible. The penalties for refusing to take part in this system are high but, according to the logic it embodies, justified.

[95] This transformation of the data is facilitated by another aspect of its digital nature that I have not explored here, which is the speed with which it can be copied.

The LBA Tower Data: "Exchange"

I have argued that the CPTEC database embodies one model of scientific community. Importantly, however, all the data in it is data that came out of the first phase of the LBA, a series of intensive collaborative campaigns spread out over the different research areas. All of these projects were European or North American collaborations with Brazilian scientists. They had already finished well before I arrived at the LBA - although some of these links lived on in the form of student exchanges, and specific further projects. The current LBA tower data is not stored at CPTEC, but instead at the LBA in Manaus. It provides an alternative model of a scientific community.

The certified tower data that J and R have cleaned is saved in several different places. They keep copies on their own computers and on the "micro-cerebro" computer in the Micrometeorology office. The Micro team also has a hard drive, kept in the LBA Database and IT department (DIS), which all the data is copied onto, and they also store a copy on Mo-Porã, the data platform that the LBA DIS team constructed for data management purposes. The storage of data was, however, a slightly haphazard affair. Sometimes, R told me, there might be 2 weeks when she did not send the data to the LBA DIS team, other times she might send the data twice in one week; it depended on how busy she was. During the time I was conducting fieldwork, there was some talk of writing a programme that would automatically "harvest" the data from each member's computer. A modeler with an interest in programming who used to work with the LBA appeared every so often in order to have discussions with the head of Micro about it. However, to date nothing concrete has been installed.

Prior to the implementation of the protocol, described in the fourth chapter, that specified and standardized the storage of the data (see Chapter 4), the data was generally copied onto CDs and stored in a cupboard at the back of the Micro office. This cupboard is now full of boxes and packets of CDs heaped chaotically on top of each other, as well as files of papers and books. It was where the student who came from Brasília was first sent to try to find any old data from the Brasília tower. When I asked about this cupboard, I was told that its lack of organization was a result of the constantly shifting membership of the group. As people came into and left the Micro group often, the lack of stability had meant in the past that there was never a standardized format for saving and storing data - people just did it "their own way". There was a similar sort of storage system for the old instruments at the tower, in a shed at the back of lodging at ZF2. Piles of old instruments, canisters of gas, and other sundry objects were stacked up willy-nilly gathering dust, having been abandoned after projects or because they broke and the funds to fix them were lacking. Like these instruments, the data in the cupboard at the back of the Micro office was very rarely used.

Once stored on the various hard drives, the tower data then awaits solicitation (*solicitação*). These solicitations most commonly arrive via email. Although both R and J receive requests, J receives very few because her data is not as usable as R's. It is not such a long series of data and, for the reasons discussed in chapters three and four, it has a large number of gaps and *falhas*. On one of the infrequent occasions that J received a request for data during my time there, it was from a researcher in ethno-astronomy who worked at an institution in Manaus, who was interested in comparing the SGC tower data with the indigenous understandings of weather patterns in the Alto Rio Negro. This was an extremely unusual request. She told him that her data was not really in a correct form to be used, and passed his request on to R. In this case, unless the data is considered to be in a particular state, it will not be passed on. I would hazard, however, that if the person making the request was someone who J knew, someone who was already acquainted with the tower at SGC and the details of the data collection from it, she would have perhaps passed on some data. For example, when the head of the Micro group had asked for some data whilst he was in Colorado at a workshop in order to run it through a model, J had supplied the data.

Unlike J, R receives what she described as "a lot" of requests for data. The data from K34 is more attractive to other researchers because it is long-term - collected since 1999, it consists in 11 years of data - and coherent enough that the relations it contains can be taken to be "patterns" in the variables, rather than errors.[96] As R put it, "K34 is much older, it is a longer series...the advantage is that you can see many more variations, patterns, changes in patterns, extreme events". At my request, R and I go through some of the emails she had received from people requesting data. There was one from 2009, from

[96] On top of that, R tells me, compared to SGC the researcher can visit ZF2 relatively easily, which helps.

a PhD student of a professor at the LBA; I also saw one request from a student I knew at the LBA who studies VOCs, and another from a Swedish student in 2008, who also wanted to know the height of the tower and the sensors. David, an American researcher who teaches thermodynamics in the LBA postgraduate course, and who I am friendly with, also requested data. I saw emails from the researcher who used to be the head of micro, and from one person who used to work at Micro but now works at another University, and from another researcher, who had moved to CPTEC in the Southeast of Brazil. There are also requests from other researchers and students at INPA. Other than sending an email, an alternative way to get the data is to come and introduce yourself. One ecology student from INPA arrived at the Micro office one day because he had taken part in an interdisciplinary and international data collecting workshop that had been run at ZF2 and had been interested by the towers. He was hoping the tower would provide him with data that he could "relate with forest dynamic". The head of the Micro group told him that he would have to check with the executive manager.

"The majority of people who ask are from or came from here" R told me. "And I already know the ones from outside (*as de fora*), foreigners, collaborators...some have spent a long time supervising someone doing a PhD or a Master's. They have already been here, they know here". As well as being acquainted with those requesting, R also has to know what the data will be used for. Sometimes, she has to exchange three or four emails with the person in order to understand what they need and why. I ask her why it matters what it is used for. "You have to say" she said adamantly, "whether it's a personal interest, if you are going to use it for something". Often, she will pass on the request to the head of Micro or the executive manager if she is not sure. "We have to keep control, we have to know, because it is our job" she told me when I pushed the issue. "You have to obey the rules. When things are all loose it doesn’t work (*as coisas muitas soltas não dá).* It's like when someone wants to go to the tower, you have to ask, "Who are you? What are you going to do there? We need a document for the visit, what if there was an accident?" Each group in the LBA has rules, she told me, and it is important to follow them - "that's organization (*isto é organização*)". On one occasion, someone asked for every single year of all the data. Even though he was working under a very well-known professor in the LBA, R told me that after talking with the head of Micro, they decided he simply could not have it all. Not only would it have taken R a long time to get all that data together, but they felt that it would be impossible for him to work with that amount of data in any case.

R asked him to be more specific. On one point, R is emphatic - "to get the data, you need to have contact with people here (*para conseguir os dados você tem que ter contato com pessoal aqui*)".

In the course of this conversation, I noticed that although showing me, R also seemed a little reluctant to let me read the emails of *solicitações*, and I struggled to glimpse the names on the list. It appeared as if she did not want me to see too much, clicking through the list rapidly and singling out only a few for me to look at. Although R was often unable to talk to me because she was so busy, there was something about her reticence that piqued my interest on this occasion, simply because in this case she had agreed to spend some time with me. I propose that R's behavior is part of a wider phenomenon that is manifest as controlling the flow of data. When I asked her if she could pass on the statistics of requests to me, she informed me that she had to ask the head of Micro first. When asked, he wanted to know if I was going to publish them. I assured him I would not, but I was struck at his assumption that I would want to. What I failed to understand, and am now in a better position to appreciate, is that as far as those in the Micro group are concerned the movement of this tower data is charged with meaning - it has effects and it makes differences (cf Hilgartner 2012). More specifically, it makes what I have come to gloss as "insides" and "outsides".

This was made apparent to me through several stories of disputes over access to data that I encountered throughout my fieldwork. Generally, the people involved in these disputes were reluctant to allow me to use their stories, because they were worried that they would be "cut off" from the LBA's data, or that the LBA would "turn its back" on them (*virar as costas*) if they made their stories public. Whether their fears would have been justfied is of course unclear. Nevertheless, their reluctance in itself is a form of ethnographic evidence. Their need for data is much more important than the desire to air any slights they might have felt they had received at the hands of the LBA. It was partly because they were afraid they could be shut out that they felt they belonged to a community that had access to LBA data. This possibility of exclusion, here, is not an infrequently occurring transgression of a free-flowing data exchange system. Rather, it is constitutive of the community itself. As one person told me of their research, "there's no other place to study this. You always end up with the LBA", so it would be foolish to jeopardize that access, as "anyone who leaves the LBA knows this - it's closed doors (*portas fechadas*)".

I am not, I would like to emphasize, suggesting that the LBA wantonly denies data access. The point I am making is that people feel that exclusion is a possibility within the logic of this system. Those who work in the LBA would not deny that sometimes, they might not give data immediately to people on the grounds that, for example, they did not know who they were. The person soliciting data does not have a relation with the LBA, and so that relation has to be forged. This can occur in different ways: an email is enough, because even that email is itself a sign of the relation that is being created. Similarly, if someone leaves the LBA, then likewise, they cease to have a relation with the LBA and that relation has to be re-forged. This community is not automatically generated, but very actively and intentionally generated. The way this might be experienced by the data solicitors in question varies widely - some may feel unfairly treated, whilst others accept it as par for the course. I am not interested in adjudicating these disputes, but in tracing the logic of the system that produces them.[97]

This LBA community is constantly being made and re-made, but not only through the exchange of data. Instruments, names, money, favours and people all circulated, although not necessarily in connection with each other. What was notable about this circulation, however, is that it was constituted by specific and often temporally-bound exchanges. It was not simply a free-flowing stream. Each exchange had its own particular contours. Money and data were often exchanged in different ways at different times. A German researcher who came over to perform a pilot campaign with an LBA project told me that even though the data he had managed to collect was not nearly as much as he had hoped for, it was enough for him to take back to his funders to request more funding. Instruments and names were other frequent exchange items (cf Biagioli 2008). One

[97] Subsequent correspondence with several informants on this matter has revealed that an automatic system is still planned in the future for the LBA tower-data dissemination, which may affect this scenario substantially. At the same time, when I was directed to websites where the LBA tower data from K34 was said to be available, I was unable to locate that particular data – only LBA data from previous projects. When I enquired if the K34 data was being uploaded onto these sites, I was informed that the process at the moment "remains the same" (i.e concerning solicitations of data) although research was being done into a software tool that would disseminate the data. This remains to be further unpacked. However, again I would like to underline that I am not suggesting that the LBA withholds data - the LBA tower data is certainly available - but that there are different data practices to take account of here. Concomitantly, my description is not about what is denied through these specific data practices, but what is creatively generated, as I observed and experienced it during the period of my fieldwork. (This footnote was added after the successful defense of this thesis).

researcher associated with the LBA told me that in order for the head of his group at INPA to get funding for his projects, or borrow instruments, he often offered to put the funder as co-author. Another LBA researcher told me that he had swapped laboratory expertise and analysis for co-authorship on his paper. And this was not always a question of specialist equipment - in fact, any researcher who wanted to collect data on trace gas exchange in the Amazon often had to use the LBA infrastructure in order to do so. what was exchanged on one side was the use of the lodging, the access, the towers, even the instruments in some cases, and on the other either co-authorship or acknowledgement (sometimes not forthcoming) and future collaboration. The executive manager's name also therefore appeared on a large number of publications that were based on data collected at ZF2, because, I was told, ZF2 was under his jurisdiction. [98]

These exchanges were always specific to each project. Mark, who we met in Chapter Three, had come to the LBA to collect data with the Picarro, an instrument that was owned by another LBA researcher in São Paulo. In return for using the tower that the Micro team manage to attach the Picarro to, and the use of the Picarro, he would share both his expertise in instrumentation, and the data that he collected - this was conceived of as a fair exchange. In fact, he told me, as far as he was concerned the data was free, once he was sure it was in the right (clean) state. Another example was David's research. David, as mentioned previously, was a researcher from the USA who taught on the LBA postgraduate course. His research was into the self-organization of deep convection in the tropics. He needed data on water vapour in the atmosphere. In order to get this data, he organized a rather clever exchange. Geodesic researchers interested in the movements of the Earth's crust collect data using GPS instruments that measure the tiniest of tectonic shifts. In order to achieve such precision, they must get rid of the interference caused by the effects of the atmosphere, mostly water vapour, in between the satellite receiving the data signal, and the GPS on the ground. This can only be done after the data is collected, through a mathematical filtering process. The "interference" that they removed from their data was in fact the "data" that David needed. So David contacted them and suggested that he could help them collect data in the Amazon, if they provided the expensive GPS

[98] The executive manager’s name would also appear because he supervised a lot of students; here, there is an exchange of work on the part of the students for the "use" of the name of the more experienced and knowledgeable researcher. "Reputation" here is part of the exchange (cf McSherry 2003: 239).

instruments. What this motley collection of different sorts of exchanges between different people at different times and for different motives demonstrates is the material and specific means by which the LBA community is generated. Although each of these exchanges had its own specific duration and details, collectively they contributed to a sense of a community. When I asked the technician responsible for the LBA's meteorological instruments about the reasons for lending the instruments so readily, he told me that it was to "keep up good relations" (*boa vizinhanca,* lit.: good neighbourhood). I would also like to emphasize that almost all of the exchanges were aimed towards eventually obtaining data.

The contrast between the LBA tower data dissemination and exchange network, and the CPTEC data model is clear. The CPTEC database is part of and constitutes what Paul Edwards has referred to as "knowledge infrastructures", which are "robust networks of people, artifacts, and institutions that generate, share and maintain specific knowledge about the human and natural worlds" (2010: 17). These networks are international, and part of the work of the systems within them is to "transform data (among other things) into information and knowledge" (ibid: 84). This is a transformation however that is beset by what Edwards has called "data friction". Friction reduces the amount of knowledge one can get from a "given input" (ibid: 84). More specifically, "data friction" is what impedes data's movement, between places and people. As the point of these networks is flow itself, many people, such as Luiz, spend a lot of time trying to reduce the friction that accompanies such flow, by building faster networks and removing people who impede that flow. The notion of friction as an impediment belongs to a certain ideal of flow that the CPTEC database embodies.

Reducing friction, on the other hand, does not seem to capture the activity that surrounds the LBA tower data. But I would be reluctant to therefore conclude that the LBA tower data system is merely a malfunctioning version of technoscientific democracy. It does not evidence the 'opposite' of flow, but it does suggest a very different logic. The LBA tower data does not take part in the CPTEC database. Luiz tells me he could include it if "the owner of the data" (*o dono dos dados*) gave him permission - but so far, it is not available through the CPTEC database, principally, Luiz thinks, because it is too "raw". The reference to limited access through ownership, a notion deliberately eclipsed by the CPTEC database logic, is crucial, and I shall return to it later in the chapter, I would argue, however, that in the case of the LBA tower data it is exactly the control of the release of the data that creates the LBA "knowledge infrastructure". It is precisely through the "friction" that the LBA community is constituted. In stark contrast to the automated and publicly-accessible CPTEC database, R tells me that she does not share the data with people who are not recognized members of the LBA community. The movement of the LBA tower data therefore reconfirms old relations, but it also elicits, and often manifests, new ones, causing the limits of the LBA to constantly shift. In this model, insides and outsides are very prevalent, and they matter.

The existence of the people who asked me not to share their stories openly demonstrates the importance of data to the limits ascribed around the LBA community. Despite their obvious involvement in the LBA, and with the LBA data, the refusal to grant access can still ensure that they are, to all intents and purposes, excluded. The data from the LBA tower therefore can be seen to be part of a dynamic and productive system of exchanges and negotiations. What this makes clear is that different forms of data practice can create not only the limits of a scientific community, but different forms of scientific community altogether. If it would be wrong to see the LBA system as an example of misfiring open-access, it would be equally incorrect to describe the CPTEC system as a misguided fantasy. I would like to sustain the notion that they are different forms of scientific community - but not opposite. In the CPTEC model, we are presented with one form, in which data is open-access and free, the community is limitless and data flow is the most highly-prized dynamic. With the LBA tower data model, the community is constantly being made and remade through quotidian data sharing, generating potent but fraught limits and boundaries.

Scientific Economies

In order to make sense of how scientific communities operate, sociologists of science have often turned to economic activity as a productive analogue. My description of different communities generated by different models of exchange therefore shares something with sociological examinations of the self-organization of science; but it also suggests some alternative avenues for enquiry. I would argue that the emphasis placed on economic analogies runs the risk of missing the fact that the act of exchange itself can generate the limits within which it functions. Once this is recognized, it becomes necessary to reassess what exactly is being exchanged, and by whom.

Karin Knorr-Cetina succinctly summarizes early sociological ideas as to how the "scientific community" self-organized: "the early postulation of relatively isolated economic mechanisms (such as competition) was replaced by the assumption of a pre-capitalist economy, which was then succeeded by strictly capitalist versions of an economy of scientific production" (1982: 104). Thus Robert K. Merton's "quasi-economic competition" (ibid.:104) was followed by Warren Hagstrom's suggestion that gift-giving was the main organizing principle in science. Hagstrom posits that "the organization of science consists of an exchange of social recognition for information" (Hagstrom 1982: 22; see also Hagstrom 1965), where the initial "information", in the form of a scientific publications, is reciprocated with recognition by their peers on publication. As Knorr-Cetina notes, the next influential model of scientific exchange, one that distanced itself considerably from Hagstrom's conclusions, was Latour and Woolgar's "cycles of credit" (Latour and Woolgar 1979: 194). In moving away from the sociological insistence on norms that Merton and Hagstrom had suggested in different ways, they instead provide a model based on "credibility" and "investment", still in an economic idiom. They propose that, based on their empirical evidence, it is incorrect to suggest that scientists are motivated solely by the "receipt of a reward" (ibid.: 197). In fact, they say, there is "no ultimate objective to scientific investment other than the continual redeployment of accumulated resources. It is in this sense that we liken scientists' credibility to a cycle of capital investment" (ibid.: 198).

Unlike a reward, Latour and Woolgar define credibility as the ability to "actually do science" (1979: 198), and accumulating this credibility allows scientists to invest it in different entities and projects that will in turn, accrue them more credibility that they can then invest, and so on. "Doing science" is as much about being able to deal with "external factors, such as money and institutions" (ibid.:198) as it is about producing facts; it necessitates the successful conversion of one sort of "capital" into another (ibid.: 201). Like in investment banking, the ability to do science increases one's ability to do science, including finding funds, sourcing instruments, currying favours and making facts. Thus it is through their amassed credibility that scientists move between the very different spheres they inhabit - epistemological, disciplinary, or institutional.

As Knorr-Cetina (1982) remarks, the power of Latour and Woolgar's analysis is to demystify science by making it analogous to another, non-scientific domain - the economy.[99] Their claim is not so much that science in fact relies on money, as that science is no different from what we already understand as neo-liberal economics. Further, as credibility is composed as much of scientific as non-scientific capacities, their theory collapses (in a move now familiar to students of STS) what is internal and what is external to science. On these grounds, I would argue that despite their subtle differentiation between credit as reward and credit as credibility, (the former being just a sub-section of the latter), and their insistence that credibility is capital and not currency, Latour and Woolgar still demonstrate that the logic underlying such cycles of credit is monetary in nature. That is, credibility can be amassed, invested and 'cashed in'. Exploring this analogy in a little more detail, it might be said that, because of what might be called its transcendental character, credibility in fact resembles money more than it does capital. Latour and Woolgar are keen to emphasize that credibility collapses the difference between the entities involved in the transaction, acting as a sort of universal commensurating medium by which 'science gets done'. It is what allows the conversion of one form of doing science into another. This holds as much for scientists' motives as to the different domains they traverse in order to exercise them. It is this transcendental nature that gives the notion of credibility its power. As Latour and Woolgar write, "[S]ince the credibility cycle is one single circle through which one form of credit can be converted into another, it makes no difference whether scientists variously insist on the primacy of credible data, credentials, or funding as their prime motivating influence" (1979: 208).

[99] Although it is true that the questions that social scientists were asking changed during this time. The 80s and 90s saw constructivism on the rise in STS, and more traditional sociological questions concerning the nature of the units of scientific action and the mechanisms of scientific integration were eclipsed by investigations to establish the ontological effects of that action. The formal language of economies seemed somehow insufficient to the task of describing the complexity apparent in the new appreciation of contextual interrelationships: Knorr-Cetina (1982) suggests that quasi-economic models in science, for example, fail to capture the realities of actual scientific work because they portray scientific work as isolated and self-sufficient, when it is in fact constituted by a context of interrelationships that spread far beyond the laboratory. Thus under this new appreciation for the construction of scientific knowledge, "cycles of credit" perhaps became "actor-networks".

Over and against this, I would maintain that it does indeed matter in some cases what the specific exchanges in question are; or rather, that it is important to pay attention not only to specific exchanges, but to exchange itself as specific. As Martin Holbraad points out, one of the defining features of money's quantitative nature is that this transcendental character is necessarily temporally bound, because "the moment I actually decide to spend my money, the 'as if' scenarios necessarily recede...my pound is now important not because it could buy anything, but because it will buy me something in particular" (2005: 244). Although Holbraad's insight comes in the course of a discussion of the divinatory practices of the Afro-Cuban religious tradition of Ifá, it is relevant here. As my description of the circulation of the LBA tower data illustrates, it is the moment of exchange, or the lack of it, that stimulates the production of exactly those boundaries that the transcendental characteristic of credibility is meant to eclipse - between scientists' motives, between scientist and claim, between what is appropriate for exchange and what is not. If credibility's transcendentalizing or abstract aspect has often commanded the attention of sociologists, here I would rather direct focus on to the moments of specifying materialization.

Latour and Woolgar's analysis - alongside others that take the economy as analogy for community - concentrates on the behaviour of individual scientists, as investment bankers of credibility. The form of community that cycles of credibility create are therefore "markets" (Latour and Woolgar 1979: 206). These are all terms that are implicated in one another: scientists as investors, credibility as capital, and community as market. In a community as market, Latour and Woolgar argue, the success of each transaction is determined by how quickly the scientist can proceed through the cycle. The point of the market is conversion and accumulation, and researchers therefore concentrate on how much the friction within the cycle can be reduced in order to speed it up as a whole. Latour and Woolgar tell us "[T]he relationship between scientists is more like that between small corporations than between a grocer and his customer. Corporations measure their success by looking at the growth of their operations and the intensity of the circulation of capital" (ibid.:207). But although this may arguably hold as an analogue for the CPTEC model of data where flow is indeed the most important characteristic, it seems clear that the LBA tower data exchange community is governed not by ideals of flow, but through specific and marked exchanges that spread data at the same time as create limits, boundaries and asymmetries. It is important to note here that I am not suggesting that this market analogy ignores the "points of nonfit, miscommunication, dislocation and non-portability" (Riles 2010: 799) within the 'market' as a social form. Luiz is very aware of these problems, as are the other database managers I spoke to. I am suggesting instead that using this market analogy to understand both the CPTEC database community and the LBA tower data community will necessarily imply that the latter is a malfunctioning version of the former. Whereas in fact, the lack of flow, the negotiated movement of objects and entities, and the control exerted over the movement of data are *essential* to the LBA tower data community *in contrast* to the CPTEC data model. The one is not a failed version of the other. It is a different model altogether.

The existence of "informal data economies" within science has already been noted by Stephen Hilgartner and Sherry Brandt-Rauf. They write:

"Much discussion of data access uses a model of research that emphasizes open publication and the process of peer review, replication, and reward through credit for discovery...But open publication is by no means the only way to grant access: Data in their many forms are also bartered with other research groups as part of the terms of collaboration, distributed to selected colleagues, patented, transferred by training visitors in novel techniques, provided to limited groups of recipients on a confidential basis, bought and sold, "pre -released" to corporate sponsors prior to publication, or kept in the lab pending a future decision about their disposition" (1994: 363).

Hilgartner and Brandt-Rauf point out the importance of "informal exchanges that form the underpinning of a dynamic and play an important role in such processes as the brokering of collaborations" where gaining or restricting access to data is a complicated political act. The idea of an "informal exchange", or "the underground economy of data", implies however a continuity between the two models, an implication which, in the case of the LBA, I would dispute. These are not formal and informal parts of the same system, but two different systems, one privileging flow and the other, exchange, which marked differences of all sorts between different groups of people. It is in fact crucial to recognize that they occur simultaneously and separately and that, as I shall now argue, *both* have ideal and malfunctioning versions of themselves.

Data Equivocations

Very early on in my fieldwork, I had a conversation with a well-known researcher after a master's defense. He impressed upon me in no uncertain terms that he considered the flow of data to be in one direction only – from Brazil to the rest of the world – with very little recompense in return. "You have to import scientists [into Brazil]", he tells me. "It's a statistical question. There just aren't enough people with a top education. Brazil is a data provider, what about [climate] models or development in return?" He went on, "the 'ground truth' is that there is always a flow of data in this direction', that is, out of Brazil and to the rest of the world. On the other hand, the head of the LBA database and IT department who was responsible for the LBA data management had a different idea of data flow. As the person responsible in the most explicit way for data sharing, in that he managed the LBA database, this researcher had a very keen interest in keeping data flowing or circulating. In fact, he explained to me that the point of a database is to keep the data in constant use by someone, anyone – the more used it is, the better-kept the data is and the less 'entropy' sets in. As he told me "[we make sure that] all this data is always an object of attention, and so it starts to change its tendency [towards entropy], it's not going to keel over (*cair*) like a person, like in the past. Whoever has collections now still on paper, and never passed them on to anyone...no-one uses that data anymore. This happens a lot. But not here, here you know your data will become more valuable and will be visible, well-documented and well looked after." The more flow, the better.

Both research and data management are important parts of LBA scientific practice. Both the researcher and the database manager were concerned with the movement of data. However, they saw this movement very differently. The first pointed to an unreciprocated movement, in which data flows out but nothing flows back in – without fair exchange, it is a movement that *separates* Brazil from the rest of the world. The second sees data flow as a crucial aspect of its own existence, and the free circulation of data as a *unifying* force. Again, the difference between "flow" and "exchange" emerges. This difference seemed to occur again and again throughout my fieldwork, at all levels of description, and with different torsions. Some of these were more extreme than others. At a meteorology conference, I had a conversation with a meteorologist who is involved in trying to draw up guidelines for the meteorological community concerning the freedom of data produced by Brazilian scientific projects. She told me that there were two clear camps – those who favoured data-sharing and free access to data, and those who did not. "We're up to our necks (*enterrados*) in data!" she exclaimed. "So much data, not just observational but from models as well" - but even so, sometimes "it's easier to get data from the United States than it is inside Brazil!"

I was struck by how versions of these two models of "flow" and "exchange" presented themselves to me in different guises throughout my fieldwork. They appeared most dramatically in accusations of data having been "stolen" (*roubados)*. Several members of the LBA knew someone whose data had been unfairly taken or stolen by a foreign research partner – that is, either not shared during a collaboration, or published without any forewarning or acknowledgement of the partnership. This seemed to be an on-going concern for many of the researchers I spoke to. At the same time, other researchers expressed confusion as to accusations of stealing; yet others suggested that there were only certain conditions under which stealing data was at all possible. When I asked an American researcher, for example, about the accusations of stealing data, he was a little nonplussed. "You can't really steal data like that," he said (meaning "in that way"). He explained that although he accepted that people have been known to take data without permission, as far as his personal experience goes he had never personally heard of anyone doing that, at least not with a state funded project like the LBA. This was because the data is free - it belongs to everyone, he said, and told me how surprised he had been that he had such an unexpectedly hard time getting data from certain institutions in Brazil. A Portuguese researcher told me that although he might be considered unconventional for it, he would not have any qualms using someone else's data - if it was available online, as far as he is concerned it's free, and he'll use it whether they had published or not.[100] Another researcher, however, told me "imagine, I've got some raw data, and I'm not working on it. And then someone comes along from Scotland and uses it to write something – you can't do that, that's stealing. It's very rare –people generally have a moral conscience – if someone uses your data, you can denounce them. If you like that soil data, and they haven't written anything – you can't just use their data, the

[100] This informant subsequently also expressed some irritation for not being included as co-author in a publication, for the data he had provided a collaborator. When I brought it up, that collaborator told me "he's not a meteorologist, that's not his area", and said he had written the article pretty much by himself.

scientific community will come down hard on you. If someone goes and uses this data, you think – hang on, I was there working, I don't know that guy – and he will end up being discredited. Officially, I'm the owner of the data."

These accusations of stealing are the result of radically different sets of data practices being brought together within a collaboration. These differences often seemed to lie along the fault lines of nationality.[101] However, almost every time that I heard stories of difficulties and conflicts between foreign (non-Brazilian) researchers and Brazilian ones, they turned out to centre around data theft or some similar act, or the refusal to share. The salient point of the stories from one side was always that someone from outside had come in (to the Amazon, to the LBA, to the collaboration) and inappropriately taken data, or stolen it. On the other side, when I asked people from 'the outside' about such accusations, I was told it was not really possible to steal data like that. One researcher asked me what the foreign researcher was meant to do, wait for the Brazilian to publish first? That might take years. On this side of the debate, I also encountered lots of stories about how data had been unexpectedly hard to obtain in the first place, or how questions of ownership had been unexpectedly important. Mark, for example, had come up to me perplexedly, having been approached by an LBA researcher and asked to whom the data from his Picarro experiment would belong. "As far as I am concerned, the data is free," he told me.

The crucial point is that this was not one discourse, but two going on simultaneously and exclusively, even in their interaction. To those that privilege free-flowing data streams that eclipse difference, the other side seem self-serving or naive; for those who understand data exchange as being a vehicle for marking differences, the other side are to be treated with caution, or even seen as bullies. "It's no good being innocent (*não dá para ser innocente*)", one researcher told me. I propose that they are both correct, because there was not one model of exchange, but two. It is not that Brazilians do not know how to share data, or that USA researchers are thieves - but that there are two models of data exchange running side by side that operate according to different principles. As anthropologist Roy Wagner puts it, discussing the relation he had with his Melanesian informants during fieldwork, "[t]heir misunderstanding of me was not the same as my misunderstanding of them" (Wagner

[101] One researcher from the USA even spoke to me about "scientific colonialism".

1981 [1975]: 20). The way that the different groups misunderstood each other was different; or rather, they did not agree even on the terms of the argument. Any sort of polarization of these positions along international lines is compromised by the fact that this difference was not only between Brazilians and foreigners, but also within Brazil, and between different disciplines, as well as hinted to me as between individuals. There was some discussion, for instance, between the different Brazilian institutions that had been involved in the construction of the telemetry system concerning due recognition online of the source of the LBA tower data. On another occasion I was told that collaboration is the only way to do science today, by the same person who would at another point tell me a story about how their data had been stolen by a French scientist at a conference. Therefore it is insufficient to suggest simply that one form of data practice is Brazilian and the other American. It was through such equivocations[102] (Viveiros de Castro 2004) as these that I came to understand that perhaps the only commonality across the spectrum of positions in my field was that foreign-ness and indeed any sort of "outsideness" is a way of talking about data, and data is a way of understanding foreign-ness and "outsideness".

I even experienced this myself, as I described above, when I asked to see the LBA's data-use statistics, and was asked what I intended to do with them. I had not expected any sort of resistance to my request because the receipt of such data would not mean, for me, that I was either inside or outside the LBA; whereas the exchange of it obviously implied exactly such issues for the LBA Micro team. Notions of foreigners and differences were construed *through* different data exchange models. The CPTEC database is one version of one of these models, and the LBA tower data is one version of another one. These two were the most prevalent in my fieldwork, but it is important to stress that they were not necessarily the opposite of each other, merely different. I suggest that the LBA tower dissemination system is somewhere between, for example, the two poles that the meteorologist described to me at the conference. I suspect that there are many models operating in every scientific collaboration. These models may exist side-by-side, sometimes running smoothly parallel, and sometimes

[102] I have discussed Viveiros de Castro's notion of "controlled equivocation" in the Introduction to the thesis; suffice to repeat here that in Viveiros de Castro's terms, "[E]quivocation appears here as the mode of communication par excellence between different perspectival positions, and therefore both as condition of possibility and limit of the anthropological enterprise" (2004: 3). I am extending this insight to take account of the contours of the collaborations I saw in the field.

colliding - but even in their collision, as in their collaboration, they remain separate. I suggest therefore that in order to understand how scientific communities and scientific collaborations work, it is necessary to move away from an all-encompassing market analogy, in order to examine how these communities emerge from different systems, logics or models of scientific exchange, and how they co-exist.

Owning Data and Scientific Subjects

I have made the case that all scientific communities are not necessarily akin to economic markets; I would like now to turn to the other implicated categories in the economic analogy, that of the "capital" and the "investor". I have already argued that it is important to countenance the specificities of the exchanges that occur. Elaborating on this, my argument will now be that if what is being exchanged is not abstract credibility-cum-capital (or indeed publication-cum-gift), but specifically certified scientific data, this has certain repercussions for those involved in these exchanges. One of the most pronounced differences between the models of "flow" and "exchange" is in their embodiment of very different notions of when and how data can be or become property. In the model of "flow", ownership is eclipsed as quickly as possible; in the model of "exchange", whose data it is becomes very important in determining how it will move. I will also pick up my suggestion, in the previous chapter, that the certified data that circulates in these networks has certain creative properties. These creative properties, I will propose, lie precisely in the possibility of ownership.

Questions of ownership appeared frequently throughout my fieldwork - as the accusations of stealing clearly demonstrate. This might be considered surprising. Several scholars of scientific property have made a case for the separation of the scientific exchange system from the liberal market economy, on the grounds that what circulates in scientific systems by definition cannot be owned. As Corynne McSherry writes: "perhaps the most vexing feature of authorship in academic science is its ability to instantiate and traverse two visions of scholarly exchange. According to one vision, scientific authors participate in a gift economy, a system of exchange premised on reciprocity, reputation and responsibility in which the commodification of scholarly work is immoral" (2003: 226). On the other hand, she says, citing Bourdieu

(1988) and Latour and Woolgar (1979) [103], there is a system of "capital accumulation and investment" which defines a market of exchange, accumulation and return. In following the instantiation and dissolution of the boundaries between gift and market economies (thus defined) in the lengthy legal resolution of a controversy about rights of use of a science course syllabus, McSherry points to the uncomfortable cohabitation of these two domains when they are brought into relief in public disputes. This tension has also been pointed out by Mario Biagioli, who draws our attention to the "distinct yet complementary" (1998:1) development of the liberal economy that privileges original expression, and the reward system of science that privileges truth - systems between which the idea of scientific authorship has been uncomfortably sandwiched. These studies, and others like them, [104] point to the murky areas and unstable legal definitions that materialize when decisions have to be made as to who owns and can use scientific products, and what happens when scientists "have to confront, usually with unease, the tensions between the traditional ethos of academic authorship and the logic of property and the market" (Biagioli and Galison 2003: 3; see also Biagioli et al. 2011; Mcsherry 2003).

In Latour and Woolgar's analogy credibility takes the place of and in effect acts like money (rather than capital, as I have argued). In these studies, by contrast, emphasis is placed on the ways in which it is inappropriate to treat scientific knowledge as capital - something that can be exchanged for money. Although funding is an acceptable form of "money" within these systems, the problem here is rather with the power of money as the index of the market economy, regarding its power to create the wrong relationship between objects and subjects. As Biagioli has extensively explored, the problem is that market monetary transactions imply relations of private property because "money is the unit of measurement of the value of that form of property" (Biagioli 1998: 2). Scientific knowledge, however, cannot be private property, and therefore cannot be exchanged for money. This is because scientists, Biagioli tells us, produce truth, and - as a representation of nature - truth, by definition, cannot be owned (Biagioli 1998, 2003, 2008). Even if the quantitative aspects of knowledge are in some cases sought after in

[103] As Knorr-Cetina did in 1982, pointing to the salience of these models even 20 years later.
[104] For example, in her investigation of the oncomouse patent Fiona Murray (2011) suggests that paying attention to the interactions and intersections of the separate cycles of "patents and papers" allows her to move beyond "a simple portrayal of patents and the commercial economy as either totally irrelevant to or destroying the traditional academic economy" to explore how patents also become means by which scientists can negotiate their own value.

order to be able to commensurate its value,[105] exchanging knowledge explicitly *for* money is a very different matter because it would imply that legally the scientist privately owns that knowledge - and no-one can own truth, as it is in the public domain. The evaluation of knowledge in monetary terms is conceptually impossible, because it refers to what is literally "priceless" (Biagioli 2003: 255).[106]

The only way a researcher could own 'nature' is if it is demonstrated not to be 'nature' anymore, and if this is the case, the relation is no longer considered to be scientific. Thus genetic procedures and biotechnologically engineered organic objects can in fact be patented, but this is only by "arguing that these products have been extracted from original state of nature and packaged within processes that are deemed useful" (Biagioli 2003: 257). This, however, shifts the product, and the relation that the scientist has to it, out of the realm of science and into the realm of the market, and IP. "Patents and papers" (Murray 2011), even if positively reinforcing, cannot be simply converted into each other.

This argument is doubly-insulated. On the one hand, scientists cannot own what they produce because what they produce is truth and that by definition cannot be owned; on the other hand, the *way* they produce it also prevents them from owning it: the scientist is "not represented as someone who transforms reality or produces "original expressions"" (Biagioli 2003: 257). This is not only in the eyes of IP law, as anthropologist James Leach demonstrates with an empirical example from a study of an interdisciplinary project involving artists and scientists. During this project, he noted that "the scientists were represented as not being creative in a subjective sense" because their work consisted in "revealing relations that already exist independent of any human subject" (Leach 2011: 145). Indeed, the scientist cannot be seen to create in any way, as this would disqualify his or her product from being "true". Creativity is therefore bound up, in the eyes of IP law and perhaps beyond, with subjectivity, and both are bound up with the terms of ownership. As scientists do not create in this way, they cannot own what they create, and as they cannot own it, they cannot exchange it for money. Exchanging their truth-bound

[105] For example with the Bibliometric index, (cf Jensen 2011), that nevertheless meets much resistance.

[106] In fact, neither nature nor people as such are appropriate property in Euro-American convention, which is why questions of disputed ownership often come down to drawing the line between what counts as nature, and what does not - or what does not count as property and what does, respectively (cf Strathern 1996).

knowledge for money would not only create a transgressive relation of ownership towards it, it would also imply that they had created it. The elision of scientists' knowledge and truth functions therefore as a means of maintaining the scientist as a sort of uncreative subject.

The idea of ownership that is in play here, in the form of IP law, is one in which subjects own (some) objects, and by so doing, have the "right" to transact them. That is, it is one's position as an individual subject in a market economy that bestows upon the object the capacity of being transactable and exchangeable. You have the right to sell your possessions, or what it is you have created. "Copyright" is exactly the manifestation of this belief. However, in the case of the discussions around scientific ownership this relation seems to run the other way. That is, it is because the object ("truth") in question is inherently 'unownable' that scientists cannot be owners. It is the *object's* inherent characteristics that make the scientist into a certain sort of subject - a non-creative one without property rights. The scientist is seen as a simply a "laborer", making "(new) connections between existent things" (Leach 2011: 155)). But what this then implies is that if the object in question were not in fact truth, the sort of subjects in question might end up looking very different.

Whereas Hagstrom’s "gifts" are publications, and publications (and therefore authorship) are also the centrepiece of Biagioli’s argument, the LBA produces an enormous amount of data, and not just publications. Data is not "donated" (Hagstrom 1982) to scientific journals, and nor is data a claim on truth. Despite this, it is still one of the most crucial items in these exchanges, because "everybody wants the data", as one researcher told me, though "no-one really wants to come here and get it". It is the possibility of obtaining data that draws researchers in from all over the world; as David told me, the reason that he has leverage with those he wants to collaborate with is because he is in the Amazon in Brazil, and therefore has the possibility of collecting data that they do not have. Amazonian data is precious because of its scarcity, which means it is highly valued. And it definitely, at least in my field, elicits relations of ownership.

Previous chapters of this thesis have shown that, as far as the researchers I worked with are concerned, nature cannot be asked whether it is real or not. It just is. Nature and truth might be understood, in the same way, not to be a particular person's property - the question does not make any sense, because they are everyone's property, and no-one's. Biagioli argues that whereas the starting point for scientific products is "generic" nature, the result is a "*specific* item of true knowledge about nature" (1998: 5). He suggests that although there is the transformation of something unspecific into something specific in both the liberal market economy and scientific systems, in the latter these two forms are not legally different enough to be able to be "legally traced or monetarily qualified" (ibid.: 5). He thus elides not only nature and truth, but scientific claims, nature and truth.

However, my own description of the production of data inclines me to disagree that what scientists produce is best understood as essentially no different from truth/nature, even though it is certainly understood to be reality. This is because, as I have argued, it is only when nature is measured that the question of reality becomes salient. With measuring, uncertainty is introduced into the world. In order to overcome this, and create reality in the dataset - that is, to give the data the capacity to carry its own referent with it - this uncertainty must be removed, as far as possible. But it will always be there. As one researcher told me "the data, it's always wrong (...) but it is trustworthy" (*os dados estão sempre errados (...) mas são confiáveis*). Eliding data and truth would eclipse the uncertainty that is an inevitable feature of knowledge in observational science. Furthermore, in this process of making the data refer to itself, nature itself recedes into the background, in order to allow reality to emerge. Data is not truth, and it is not nature - but it can be more or less real. It can also be stolen, according to my informants, which means it can be owned. So, given these ethnographic particularities that seem to distance my field from Biagioli's conclusions, the question here is, what is it specifically about the LBA certified data that means that it can be considered as property, and what different notions of property are in play here? And what does this in turn mean for our understanding of scientific subjectivities involved and their potential for creativity?

Means and Ends

I would like to propose that ownership can be understood as an axis that runs between data and knowledge, where knowledge is understood to be a claim about the world. This claim might be in the form of a publication, but it could also be in the form of a presentation or any form of substantial analysis of the data.

To demonstrate this, I will return briefly here to two ethnographic instances that I have already introduced. I asked the researcher who had impressed upon me that data only flows in one direction, out from Brazil without due recompense, if he thought that was because data is only "discovered", compared to knowledge, which is considered to be in a sense "made". He disagreed wholeheartedly, telling me, "[Y]es, some people say "oh, the data in itself doesn't have any value. It needs a certain amount of intellectual work." But data has a value. There's an enormous amount of work done on the data already – collection, treatment and so on – each part is not a single thing, one thing (...) The data can be very different in a complex environment [like the Amazon]. To install the instrument is already a lot of work, as much intellectual as any other."

On the other hand, the database manager that I have also already introduced had a very different idea of this relation between data and knowledge. He told me, "Data, it generates knowledge (…) It's extracted, by the sensor, or observations (…) but looking at this data, it doesn’t make much sense (…) you have 40 MB of data from an experiment but it won’t give you any scientific answer. What is there in this data?’ The difference he was pointing to, he told me, was between "data and knowledge. Data doesn’t tell you anything. I can get the data and analyse it, but the number isn't going to tell me anything. The data is just a number, it doesn't have a nature (*natureza*). It can be in various scales. The number is not knowledge. It's essential to generate knowledge – I can't affirm anything without a number." Although both researcher and database manager agreed that work needed to be done in order to make data meaningful or valuable, they differed concerning the point at which this happens. In the former case, data already *was* valuable, because it had already acquired meaning through the intellectual and physical work required for collecting and processing it. In the latter case, data "doesn’t tell you anything", and had to be turned into knowledge first before it could be thought of as meaningful and valuable.

If, as the first researcher argues, it is because data already has a meaning and value that its circulation should not be free, then meaning and value are tied to the possibility of ownership, through the work that is done on the data. In this model, what the data gains through being worked on and made into certified data is the potential to be transformed into *something else*, and this potential has a value. It is around this transformation that ownership coalesces.

The transformation of data into knowledge, mediated by work, is therefore crucial to understanding data as property, and particularly to understanding the different models of data exchange that I have already described. A key feature differentiating the CPTEC data from the LBA tower data is that the researchers in LBA-ECO, the NASA led section of the LBA that contributes data to CPTEC database, have all collected and cleaned their own data (or had their students do it for them). Furthermore, those who contribute data to the CPTEC database have all had the chance to publish the data - at least formally, within the logic of the data policy. That is, they have all had the chance to "own" the *process* of the transformation of the data into a claim about the world, or into knowledge. They have all had the chance to realize this potential that is understood to be in the data, thanks to the process of cleaning and certification. In the case of the LBA tower data, on the other hand, the data cleaners are excluded from this transformation. That is, they exchange their work on the data for money, so that *others* may use the data - as several people told me "it's just their job". By this reasoning, because they do not publish with the data, and therefore maintain a certain distance from it, they do not own it. As R told me,

"Look, I...no, I'm very clear. I don't have any link to this data. Whenever [the head of Micro] or anyone says that I do I say no, that this doesn't belong to me. This is not my property. Because...I didn't see...right, I didn't see how it was there in the field. I wasn't there. I don't have a close record, see? So I can say "I use this record, it's very close, it's like I was there in the field". So, I am going to meet lots of barriers, it's going to be difficult...and this would take me an enormous amount of time, time that we generally don't have here. So, what do I do? I try to deal with it with what I do have (...) This is what I do have [pretending to speak to someone who is asking for data]. You're the researcher, you're the one who is going to work with this."

She was referring here to situations in which people would complain about the state of the data - that it had too many gaps, or why she had left in certain values (as discussed in Chapter 4). Not being the owner of the data absolves one of responsibility for it on the one hand, as R is suggesting; but on the other, it also means one cannot carry that data forward into knowledge. The situation almost seems to be at right angles to the one described in Biagioli's analysis. Here, the fact that they have exchanged the *work* they did on the data for money is what prevents the data cleaners from claiming to own the data itself. Conversely, those who work on the data themselves and who therefore 'own' it, are able to do so because they have not accepted any money for that work.

Of course, the situation is not as clear-cut as this. In J and R’s case the issue is that they *do* in fact have a claim in the data, *exactly because* they worked on it - whether they were paid to do so or not. As the researcher I cited above points out, even raw data already has had a great deal of work invested in its collection. The free-flow of data out of Brazil is inappropriate because the data is not free at all. I asked a researcher from the USA about what he thought of R and J’s position, and he told me that he imagined that in the USA, their job would be done by a post-grad student. When I asked him if that student would feel like the data was in some way theirs, he replied "hell, yeah!" Despite the fact that J and R are not post-grad students, and even though the data-cleaning being "their job", they still feel a claim on the data. This sentiment was made very apparent during one trip to the tower K34. During the preparations for the installation of some radiometers at ZF2, the head of the team told me about a campaign that had taken place in June whilst I was away from Manaus. Students from the LBA and INPA and other Brazilian institutions, and from the University of Arizona, had come to ZF2 to spend a week or so collecting data in different ways, during a didactic field trip to train them in what data collection campaigns consist of. The American professor in charge had asked for the data from the tower K34 afterwards to compare to the data the students had collected during the field trip. He was very complimentary about the coherence and usability of the data set provided to him. In the discussion that followed, several people commented: "well, they’ll write an article quickly from that data. No-one considers that someone worked on that data, or will include her" and, "the data is used so much that no-one includes who treated the data…if everyone did, R's CV would be amazing!" R's work was clearly linked here to the data, and although the professor had made it clear he would not publish anything from the data without permission, the Micro team felt that their work on it gave them certain rights to it, even as they recognize their role in disseminating it.

In this way, the data is recognized as having a past as well as a future. Almost all the activities that contribute to a data set's emergence as an object have the potential to be a claim. When I raised the subject with a researcher from the USA, he acknowledged this but wondered where you would stop, if you were to include everyone involved in producing the data as the owner of the data. Do you include the motorist who drove you to the tower, the tower builders, the cooks at the lodging? On the other hand, as one member of the Micro group remarked to me, up to their elbows in soap suds whilst scrubbing a thermohygrometer, "who wants to be here doing this? Generally, people just want the data, they don’t want to work for it". This sense of entitlement was tempered however, in many cases, by a sense, as R said, that the data also was not hers, that the work she did was just her job. One researcher also acknowledged that in publishing using the data, "the foreigners are only doing what us Brazilians are not." J and R both have master's degrees, and they both expressed an interest in doing a PhD. But there simply was not time. Both of them had been struggling during the period of my fieldwork to find the time to write an article based on the data from the towers. Their frustration was quietly held, and rarely surfaced explicitly. If I asked them about it, they avoided the question. But as one researcher told me of his decision to leave the LBA to do a PhD elsewhere: "do you want to be a technician with a master’s degree your whole life?" This was a sentiment I often heard: "at the LBA, people with master's degrees do the work of technicians." (*No LBA, quem tem mestrado faz trabalho de técnico*). It was often aired as an expression of frustration, and explained as a result of the lack of funding that the LBA was experiencing at the time. Whatever its cause, what it signifies is the difficulty in classifying or categorizing J and R, and their work, and their role as scientific subjects. Given the way in which work and ownership are mutually implicated, I would argue that their particular position can be understood as one in which they own the data, because of their work on it, but *do not own the potential the data has in it.* The data is in this way split for them in a way it is not for those partaking in the CPTEC database, for whom the data as it is now and its future transformation are merged into one singular form.[107]

[107] Returning to the differences in data practices previously examined, the data's temporality provides another perspective on the different models of data presented. As anthropologist Marilyn Strathern writes, "[O]wnership gathers things momentarily to a point…[It halts]

R and J are in an unusual and liminal position in this sense; perhaps not unlike the egg donors of anthropologist Monica Konrad's study, who give away potential without any claim to its realization (Konrad 1998, 2005). Here too, "what signifies is being the origin of a process that another carries forward" (Strathern 2003: 185). But whereas in Konrad's analysis, it is in the nameless flows of donated ova to unknown others that lie "the beginnings of a non-possessive modeling of these...anonymously pooled, body parts" (Konrad 1998: 653) lie, what J and R experience is a sense of ownership that cannot be openly expressed, but that they nevertheless feel strongly. The ova donors Konrad describes have a sense of a connection between themselves and their recipients, but they also preserve anonymity. There is a nameless relation, such that they "cannot quite put [their] finger on how to identify the 'something' that is the relationality between them, and this kind of diffuseness (...) is the form of the relatedness making up the connection" (Konrad 1998: 652). But whereas what surprises in reading about Konrad's donors is their capacity to be anonymous and non-possessive in relation to their ova, here it is the inverse: it is surprising that J and R should feel any sort of ownership towards something that is firstly simply their job, and secondly, if we recall Biagioli's and Leach's arguments, not creative in any way. In Konrad's analysis, it is exactly the anonymity that creates what she calls "transilient" kinship (ibid.:559), whereby "women themselves can make their procreative powers into an ovular economy of intersubjective (cross-corporeal) agency" (ibid.:655). Conversely but just as creatively, I argue it is exactly J and R's eclipsed possessiveness that brings forth a certain controlled economy of information, in which the relations evoked by the data's movement are of great importance. As Soumhya Venkatesan draws our attention to in a very different context, the transformations worked upon an otherwise anti-social object (or "free gift") can make it into an "enabler of relationships, not between gift giver and recipient but between hitherto unrelated people" (2011: 48). [108]

endless dissemination" (Strathern 1999:177). Turning this around, I would say *momentarily halting the flow makes an owner*. The point at which the data's past and future are decided is the act from which ownership precipitates.

[108] Another way to contrast the British ova donors with the LBA data cleaners is to suggest that whereas in the former case, it is surprising that such objects as ova can be "alienable", in the latter it is curious that such an object as scientific data can be "inalienable". What this distinction speaks to are Western notions of the relations between subjects and objects. As Christopher Gregory defines it, in a "commodity" economy such as the West, exchanging

Through the work that they expend in cleaning the data, J and R create a scientific object that can have a specific form of relationship with them - one of property. But this relationship runs both ways. The certified data also creates them as a certain sort of scientific subject. The different configurations of the transformation of data into knowledge that I encountered in my field not only contribute substantially to making "humans into particular kinds of subjects called scientists" (Haraway 1997: 142), but also to making them into particular *kinds* of scientist subjects. Fortun and Fortun write that "[S]cientists are, inevitably, coded – by the technologies with which they work, by hegemonic cultural formations, and by forceful political-economic currents" (2005: 50), but that "coding", I would argue, is a form of relating. There needs to be a space for a notion of a scientific subject based not on whether he or she adds to what he or she produces, but on the inevitable fact that what he or she produces gives them the potential to relate in certain ways. The two models of data sharing that I have discussed differ most profoundly not because one (the CPTEC database) deals with "truth" and is therefore considered to be in the public domain, whereas the other (LBA tower data) does not. The important difference between the data that goes into the CPTEC database, and the LBA tower data, is that the data that is in the CPTEC database is data from projects that are run by researchers who control the means as well as the claims to 'truth', or what am glossing as "knowledge" or "future claims". That is, the researchers in LBA-ECO, the NASA-led section of the LBA, worked on their data, and produced publications from that data. Their means and their claims were never separated.

In the case of the LBA tower data, however, the claims are separated from the means. The people who work on the data do not then go on to make any claims about the world using it. Thus the data cleaning process, and the resultant certified data, bestow upon the

alienable objects "establishes a qualitative relationship between the objects exchanged" (1982: 101). Objects come to relate through commensuration and exchange to other objects (in much the way that Latour and Woolgar would have it). The contrast to this Western model is a "gift" economy, where exchanging inalienable objects establishes "a qualitative relationship between the transactors" (1982: 100). Here, exchanging objects is a means of relating people (as perhaps is the case for Konrad's ova donors). Using this distinction very loosely, and with due regard to the work that has been done to think beyond it, and especially beyond the dichotomy between persons and things it presupposes (cf Strathern 1988), I suggest that the process I have been describing is not only about relating researchers to researchers through exchange, or relating data with data. It is about the relation that researchers have to their data; about subject-object relations.

data cleaners a relation of 'ownership' that they subsequently *cannot* turn into a relation of knowledge. What J and R do not have any stake in in this sense is the "forward future" of the data (Strathern 1996: 18). In some IP regimes "potential becomes an asset" (ibid: 17) that one can establish a claim to, in order to benefit from its future. But here, these "future uses" are exactly what turn the potential into something that is unclaimable. Publications (claims to truth, in Biagioli's terms) are different sorts of objects than data (means to truth), because different types of work have been invested in them to make them what they are. In turn, they create those who have invested the work into different sorts of scientific subjects.[109] J and R are therefore left, in a certain sense, data owners but never knowledge producers. [110] The situation with the LBA data is not, I suggest, that different subjects have different claims to different products, and that scientists simply cannot own what they produce. Instead, different objects allow different claims (as relations) to be made of them. Certainly, scientific data practice does "political work" (Hayden 2003), but I suggest it does so in a Stengerian fashion. Twisting Stengers' formulation, the production of certified data gives the data the power to turn the scientist into a certain sort of subject.[111] Thus it is not only that "objectivity here is an assumption of – and made possible by – specific processes of creation" (Leach 2011: 155). "Objectivity" is itself capable of creativity.

Conclusion

[109] I did not conduct fieldwork concentrating on the scientific publications of the LBA. But two hypotheses now present themselves, given my argument. The first is that, as the end product of this process of scientific knowledge-making, turning data into publications is a matter of moving further and further away from nature *just as* the publications become more and more real. If this is the case, then I would hesitate to elide publications with "nature/truth" so quickly, and instead examine the properties of this knowledge. The second hypothesis is that publications are in fact a re-instatement of the indifferentiation of nature. They become so real that they become nature.

[110] Another interesting capacity of publications is that they themselves close down the potential for infinite expansion that is evidenced in the certified data. The concern was where you might cut a potentially infinite series of relationships, what I have referred to as the data's past - that includes driver, tower-builder etc. The question of contribution boiled down to what *sort* of contribution, and the publication made an important cut in this way. See note 106.

[111]"This is the very meaning of the event that constitutes experimental invention: the invention of the power to confer on things the power of conferring on the experimenter the power to speak in their name" (Stengers 2000: 89).

In this chapter, I have argued that there are at least two different models of exchange operating in my field, one I have called "flow", and the other "exchange". These create two different forms of scientific community. The moments of disjunction, misunderstanding and friction that arise during international collaborations can be seen as a result of the cohabitation of these two models.

I have also suggested that one of the crucial differentiators of these two models is the varying relation between data and knowledge, and its corollary, ownership. Whether data is appropriate as property is not a question only raised in the social scientific literature. It is also an issue that my informants found themselves negotiating in the course of their scientific collaborations. One means of understanding the different proprietorial positions my informants assumed regarding data is by examining the relations that the person in question has to that data. Thus the data cleaners assume a liminal position because, although they are paid in order to relinquish claims on the data, the work that they are paid to do has the opposite effect; it ensures that they feel they have a stake in the data. They are thus in a position where they feel a sense of ownership, at the same time as being unable to convert that sensation into a fully-fledged entitlement - that is, the conversion of data into knowledge, or a claim about the world. I suggest that, given the emphasis in the literature on the elision between scientific publications and truth, it might be the case that in this conversion, entitlement is in fact eclipsed altogether, or replaced by another form of relation (see Biagioli 2008 on the role of names in scientific publications). In other words, 'data' and 'knowledge' are perhaps different relations. Examining the data cleaners' liminality more closely, it becomes apparent that it is the work that they do in creating the certified data that enables them to have a claim on it. Isabelle Stengers describes a recursive agency for scientific claims, such that science's relation to nature resides in "the invention of the power to confer on things the power of conferring on the experimenter the power to speak in their name" (Stengers 2000: 89): I would similarly suggest that the data cleaners give the data the power to make them into certain sorts of subjectivities. Instead of concentrating on the lack of creativity and subjectivity in science, I therefore draw attention to their creation.

This chapter has attempted to sketch out a relational theory to describe particular scientific communities, scientific subjectivities, and scientific objects. The relation I have identified as generative of these three identities is the one between data and those who work with it. However, there were several important members of the LBA scientific community distinguished precisely by the lack of any relation they felt to the data. It is to these, finally, that I turn in the next chapter.

Chapter 6: Doing Indifference with Movement

Introduction

This chapter will take movement as its descriptive framework. The LBA, as a working environment, is characterized by instability, and precarious immobility. Several different sorts of movements are engaged in, against this fluid background, by the researchers, students, data cleaners, and technicians. The last two chapters pointed to the potential of scientific data, not only to produce knowledge, but also in so doing to create and embody productive relations with those involved with it. In this chapter, I want to concentrate on descriptions of those who are crucial to the data's collection, but have no interest in controlling its movement or effects at all. I will refer to these people as the "forest technicians". They appear as an entity *within* the systems of exchange I have described, constantly circulating around the LBA research sites, and between projects, coming into the office every so often, only to head out to the forest again. They thus appear to be excluded from these systems - always circulating on the periphery, ignored and invisible.

The forest technician's invisibility and exclusion resonates with the conclusions reached by other studies of laboratory technicians, past and present. The endless circulation of the forest technicians between the different research sites, which seems to be a marker of their exclusion, can be contrasted with the international, volitional movements of the students and professors within the LBA. But what the forest technicians feel excluded from is not the system of exchanges that forges scientific subjects. Rather, they feel excluded from a fair salary and a secure job - from the benefits of the wider economic or employment system that the LBA is a part of. The instability or fluidity of the LBA working environment, I argue, encapsulates two very different positions: the forest technician's exclusion from economic security; and their *indifference* to the system of scientific recognition. Their incessant circulation is an indication of their exclusion from stability, but it also contains within it the motile index of their indifference.

The technicians' movement around the system of exchanges contains within it a movement in and out of the forest. From the perspective of the office, this movement from the office into the forest is a demonstration of the forest technicians' impotence, low status and vulnerability. But when in the forest, the technician's capacities and knowledge come into their own. The movement between office and forest is therefore transformed according to the perspective from which it is seen. At the same time, it is itself transformative. In conclusion, I offer the possibility of countenancing indifference, as opposed to resistance, as a way of relating to dominant discourses.

Unstable Immobility

In a very general sense, the Micro group at the LBA was characterized by a fluidity of membership and an instability that I found quite startling. This fluidity was in fact apparent to a certain and unsurprising degree at almost all levels in the LBA. However, it was most marked in the Micro group and the extended workforce of different technicians who worked for the LBA. During the time I spent with the LBA, people seemed to join and leave the Micro group very frequently. Generally, this instability co-existed with a perceived lack of movement within the group, a lack of progression up through the ranks. This sense of unstable immobility particularly characterized the way that the extended workforce of technicians spoke to me about their work.

Between my master's fieldwork in 2007 and my PhD fieldwork in 2010, several of the members of the Micro team had decided to leave the LBA and pursue a PhD abroad, or had left because they were offered more secure "concursado" jobs in other Brazilian institutions. Even J and R, who had been with the LBA for several years, recounted tales of uncertainty and trepidation. Although R was employed by the LBA when I got there in 2010, J, who I had had a lot of contact with when I was completing my fieldwork for my masters in 2007, was still a *bolsista* – she had a grant from the LBA, and therefore very little job security. R on the other hand was in a more stable position - employed, with a salary. By August 2010, J had also been given a *carteira* (a more permanent position as a government employee), but as several conversations I had with other members of the LBA demonstrated to me, it was considered long-overdue. Two members of the Micro team left during the time I was there in 2010, and one more had wanted to leave but had been convinced to stay by a slight increase in salary. The fact that J and R had such a backlog of data to go through was attributed to this impermanence. There had been so many different members of the Micro group over the years, each one with "their own way of doing things" (*seu jeito*), and many of them had left without recording important information such as the calibration constants of instruments, the dates instruments were upgraded, or the field reports (*relatórios de campo*).

One of the distinct features of this impermanence was that many of those in the Micro group had very little formal training, and rather had to learn as they went along. R and J, for example, had both been asked to join the LBA after completing their master's degree. Both of them were faced with an intimidating task. Neither had had any formal training and were faced with processing an enormous amount of data that had piled up over the years. J told me that in her first year at the LBA, when she was helping more with the data collection at ZF2, she was also expected to learn how to deal with tower data and maintenance without any specific training. She began simply by accompanying another member of the team until she learnt the ropes. But when he had unexpectedly and tragically died in an aeroplane crash, J was forced to do the job by herself, or with another newly-arrived technician. The lack of training meant that those people who left took with them a great deal of experience that they did not have the chance to share with whoever filled their shoes.

A similar story was told by other data collectors in Micro. One had been taught how to collect the hydrology data in just two weeks by the foreign Principal Investigator managing the project at the time before he returned to Holland. Another had recently been "promoted" from being a driver (*motorista*), which he had done for many years already with the LBA. Although he was highly valued as a driver due to his ability to navigate the perilous dirt tracks leading to the research sites, this was nonetheless a clear promotion. He was taking over the care of another tower, ZF3, from someone who had left inexplicably some time before I arrived. However, he was having some trouble with understanding the system at ZF3, which he said was like a "black box", because it had been installed by someone who was now working in São Paulo. R and J did not quite understand it either. He was in the process of being trained by the head of the Micro group to manage the data from ZF3, when the head left to take up a position in another institution in Belém, in a different state. Another member of the Micro team had been hired to work in the field at ZF2 around the time I started fieldwork. He was keen to start a master's, which would mean he could not spend that much time in the field. I have since been told he has started his master's at the LBA CliAmb programme, which means he will only be working part time with the group.

Underlying this fluidity in the composition of the Micro group, there was also a sensation of immobility. The result was a feeling of precariousness, exacerbated by the fact that the LBA has struggled with financial difficulties for several years, weathering state cuts to science and technology budgets and to educational funding. I recently discovered that several of the LBA's permanent members had been fired, including the chef at the lodging,[112] or have decided to leave. Others have had to go several months without being paid. Coupled with this financial insecurity is the fact that many of those who leave the LBA do so because they feel - or they know - that they certainly cannot progress any further with their work. I heard many times that after years of working as a *bolsista*, collecting data for other people, the only option to continue with one's career is to move to a University, or to another institution. As one data collector with a master's told me, she had decided to take a *bolsa* to collect data for a project that had nothing to do with her own work, in order not to "lose continuity" - but she felt depressed not to be working on her own data. This is exacerbated by the knowledge that increasingly fewer places are available every year for those looking to pursue a post-graduate or a post-doctoral position. As one technician-researcher told me, the LBA was losing recent PhD graduates because there was no "space" for them to continue working there: "this is the reality". There are only a certain number of places opened for researchers in INPA every year, and the competition for them was fierce; the same went for PhD places within the LBA. The LBA therefore sits, partly, on a shaky bedrock of potential researchers doing technical data cleaning jobs.

This sense of immobility was felt even by the more secure members of the Micro team, such as L, the electronics technologist who was hired the year before I arrived

[112] The chef, however, is looking forward to setting up his own restaurant in Manaus with his severance pay.

in 2010. As he explained to me, he saw his role in the LBA to be that of an important "tool" (*ferramenta*). No-one else had the knowledge that he did. He was in charge of the meteorological instruments, and having just finished a degree in electronics, he was invaluable in that role. He had therefore already reached the top of the LBA salary, and felt like his "voice was heard". However, he recognized that whereas this salary was enough for him to live comfortably at the moment, if he wanted to start a family or buy a house, he too would have to find another job - he would be forced to look elsewhere. L told me he had no interest in the data itself, and impressed upon me that he felt no need to gain scientific recognition - but he also recognized that he would need to move on if financially necessary.

In emphasizing this unstable immobility, I do not mean to suggest that it was a direct or intentional result of the actions of those in charge of the organization of the LBA. Like most Brazilian state-funded institutions, the LBA is part of a funding hierarchy, and cannot ultimately control how much funding it receives, nor how many places open for researchers. The executive manager of the LBA was constantly trying to ensure continued funding for the project, from all sorts of sources. However, whatever the causes, the resulting institution was organizationally unstable, and this creates a certain sort of social topology. In the previous chapter, I described the limits of the LBA community as being generated by access to data and the control of this access. I would suggest that this organizational instability is interdigitated with the expansion and contraction of the LBA community through those data practices. The two phenomena combine to create a fluid background of experiences of being excluded from different benefits or rights. It was against this background that my informants went about their work. It was always present, although some felt it only indirectly.

A Hierarchy of Movement

Against this background fluidity, L and the data cleaners and collectors in the Micro group who had contracts with the LBA felt relatively secure, even if they felt themselves to be limited or immobile in another sense. Alongside the Micro group, there were several people I will refer to as the "forest technicians", and the *mateiros,*[113] who were also very involved in the production of the LBA data, but whose work was characterized by the fact that it occurred exclusively in the forest. Whereas *mateiros* might be called upon by different groups in INPA and the LBA to help with specific work in the forest, I am using the term "forest technicians" for those technicians who collect data on a more permanent basis for different LBA groups. Their work with the LBA is more or less limited to work in the forest, and they very rarely come into the LBA office. It became clear that this fact distinguished them from the other technicians working at the LBA.

One forest technician who I spent a great deal of time with in the forest was a data collector for another LBA group. He would spend two weeks or so out in the forest at a time, leaving his family in Manaus. He had assumed responsibility for the data collection after his brother had left to work on another research site. After his initial *bolsa* had expired, he had told the LBA that he was leaving unless he got his *carta assinada* (a form of contract). The executive manager of the LBA acquiesced. This forest technician had also had to learn the ropes as he went along, and it had been hard at first, he told me. After only a few months training on-site with his brother, he sometimes had had to collect the data completely by himself due to the high turn-over of people in the group (similar to that encountered in the Micro group). This was not only dangerous, as he had to walk long distances through the forest on his own, but also very hard work. As the data collection involved several different data collection platforms spread out throughout the forest, it was complicated to learn all the different periodicities of collection for the different instruments and the different programmes on the dataloggers. If he made a mistake, he told me, it was even more difficult to explain to the head of the group. "Of course I was unprepared, I didn’t even have one class about it!" he exclaimed.

Now however, after six years, he feels comfortable in the job, although he would sometimes express his frustration at being kept in the dark concerning the future of the hydrology data platforms that he cared for. When I asked him about his prospects,

[113] "Mateiros" are literally "foresters" - normally local people who know the forest, often both in a traditional sense (they know how to live in the forest) and in a botanical sense. They are often employed when inventories of the tree species of certain sites must be taken. The word was sometimes used to refer more broadly to any person who worked in the forest.

he was not hopeful of ever becoming anything more than a data collector. I asked someone in the Micro group if they thought he could do a master's at the LBA, and was told it was highly unlikely, even though it was ideally meant to be part of the process of recruiting young local people like him. Experienced and trustworthy forest technicians were an extremely rare resource, and therefore very valuable. This particular forest technician had impressed one foreign researcher so much with his skills driving a quad bike, negotiating the difficult terrain, his knowledge of the instruments he looked after and his general trustworthiness that the researcher had told me he thought he was almost like a post-graduate student. Another informant, who had recently been "promoted" to the Micro group, told me that the head of his previous group had considered him to be "more valuable in the field". Now however, he wanted to "go up the ladder". After eight years in the previous group, he did not feel he had anything else to learn, and he had never progressed (*nunca subi*).

This last informant's case is interesting because he was extremely highly-valued in the LBA for his trustworthiness, his experience and his general capacity to turn his hand to anything. Even though he had no formal education in the Earth Sciences, he had worked for so long with data that he was no longer considered to be "just" a forest technician, which is why he was so keen to move up in the Micro group, and do more than just go out into the forest to collect the data. There was another construction technician who likewise had been so involved in the building of the LBA towers that he was considered indispensable. However, unlike the other technician, he never expressed his discontent about his position to me. These two technicians were often fought over by students and researchers who needed them for their particular projects, and were constantly criss-crossing the Amazon to join the different data collecting projects they were asked to help install or run. Their expertise made them more secure than the other technicians principally because, researchers told me, they were "trustworthy" - and this in part because they had stayed so long working in the LBA, they had amassed an enormous amount of experience. In the first case, however, my informant was tired of being stuck as a data cleaner and technician and wanted to move on, and, as it turned out, the second informant lost his job due to personal reasons.

Alongside the forest technicians, who have been more or less formally trained, and these special cases, there are several other more permanent members of the research site ZF2, who have nothing to do with data collection. Their job is to ensure the smooth functioning of the lodging. This involves looking after the trails, cleaning the lodging or cooking the food. One told me that even though he had been working there for 5 years, and had a *carta assinada*, he still did not feel like his voice was heard: "the big ones don’t listen to the little ones", he told me. Other workers (*mão de obra*) were often drafted in to help with the heavy physical labour involved in the installation of projects, and they expressed similar concerns. Their concern was in part due to a peculiarity of the Brazilian labour system, which they referred to as the *vínculo empregatício*. As a technician explained to me, what this meant in practical terms was that none of them could be used consistently as workers for too long a period, or they had the right to request a salary. This meant that they were kept on a sort of short cycle: they would be used for one project, and then another person would be contracted, and then perhaps they might be used in the next project - or not. But the effect, again, was destabilizing and this staccato working rhythm does not lend itself to a sense of progression. As one worker in this position told me, he travelled two hours by bus to get to the LBA from the outskirts of Manaus, and was told there was no work for him. But he knew there was work, it was just being given to someone else. Why did the LBA choose the other person? I asked him. "We’re all in the same boat", he told me forlornly. "Everyone needs work. If I manage to get a good job, I won’t come here anymore."

The forest technicians and the workers are in a sense excluded from the possibility of gaining scientific "credit" for their work, but in general this mattered little to them. When I asked whether they minded not being given due credit in researcher’s publications or studies, many would tell me that they did not. The important exclusion for them seemed to be from the other benefits, such as money and stability. They indicated to me in various and often indirect ways that they did not like the uncertain position they found themselves in with regards to having continuous work. They felt their exclusion from this as a form of unstable immobility, which manifested itself throughout the constantly shifting membership of the extended workforce of the LBA. This instability seemed to be felt most keenly by those who were least specialized - thus the manual workers were the most affected, considered to be more "replaceable" than the forest technicians who have spent years amassing experience. As one moves further away from the LBA office and into the forest, it seems as if the sensation of precariousness increases. Whereas those with a contract with the LBA can be assured of a salary, those who are workers find themselves constantly worrying about where the next job will come from, let alone whether they will ever be able to progress "up the ladder".

This movement in and out of the forest, I observed, is part of an already-established hierarchy of movement, whereby the "top" (*top*) researchers tend to stop going into the field, and the instruments, data, favours, recognition and technicians tend to circulate around the different research sites. As several students remarked to me, the higher up the scientific hierarchy you go, the less you go into the field. Although a scientists' name can circulate (cf Biagioli 2008), the ideal would be for the researcher to be able to remain in the office. It was always something of an event when a top researcher from the LBA went to ZF2, because it was such a rare occurrence. The movement that defined the top researchers of the LBA was, rather, international and volitional. They decided when to travel, and they often travelled to another state in Brazil, or less often to another country in order to conduct research or to engage in collaborations.[114] Many of the well-known researchers that worked full time at the LBA had postgraduate degrees from illustrious institutions abroad, such as Harvard or universities in Holland. There were several students in the LBA who had spent or planned to spend a term at a university in Europe or the USA. Several people previously associated with the LBA had left to do a PhD in Holland, the USA or the South of Brazil, in Rio Grand do Sul or São Paulo. It was clear in the way people talked of them and sought to achieve them that these foreign credentials were highly

[114] In more general terms within STS literature, the mobility of scientists or "trained personnel" (Barnes and Edge 1982: 20) has been noted in the context of the transfer of knowledge about procedures, experiments and instruments (Collins 1985; Law 1973; Barnes and Edge 1982). In the case of the LBA, when a "foreign" researcher arrived, and there would often emerge a new collaborative relational formation, usually including the exchange of knowledge for the use of the infrastructure. Several researchers from outside Brazil and even from the South of Brazil, visited during my fieldwork; of those I observed and had contact with, it seemed clear that there was always at the least an informal exchange of information, many times the sharing of expertise in an instrument or the formal sharing of data, and in one case a lecture series. This information transfer was not in order for the results of experiments to be proven via repetition (Collins 1985), but was a condition of the foreign researcher's "mobility", as manifest in their presence in the Amazon.

valued. This international movement was therefore not only volitional; it also brought with it prestige. The executive manager of the LBA was almost impossible to interview because he was always travelling. He often went to São Paulo or Brasília or even further afield - but never, to my knowledge during the year I was at the LBA, into the forest. Students conducting postgraduate degrees went into the forest only to conduct their short-term data collection experiments.

In contrast, then, to the sporadic intrusion of foreign researchers, the past academic sojourns of the "top" researchers, and the frequent prolonged absences of the students, the forest technicians travelled in and out of the forest, around different researcher's projects. Going into the forest was linked with a certain sort of activity and assigned to a certain level of employment. This level of employment was invariably seen as involving manual labour. When *any* researcher went into the forest - "to the field" (*ao campo*) - they expected to have to behave differently. They dressed differently, ate what was cooked for them, slept in a hammock and spent all day out in the forest, often carrying heavy equipment, driving a quad bike or fixing instruments. However, this period, as well as being optional and occasional, was limited normally to a day or two. The constant movement into and out of the forest by the forest technicians becomes elided with being shuttled around the system of exchanges that I discussed in the previous chapter. This movement allows the forest technicians only to circulate around the system, without being in control of their own circulation.

The Elision of Exclusion and Indifference

This hierarchy I have described generates a sense of exclusion. There is a sense in which those who do benefit from the LBA, in terms of professional recognition and funding that they accrue, do so at the expense of people who are easy to ignore but nevertheless crucial to the process. This exclusion can be coordinated analytically with by-now familiar concerns as to the exclusion of the putatively 'non-scientific' from our understanding of the production of scientific knowledge. This concern is a double one. It indicates an impoverished understanding of the process of the production of scientific knowledge on the part of social scientists, and it simultaneously indicates that technicians might not be recognized as crucial actors in this process and therefore not gain just or fair recompense by those who are in a position to give it to them.

This double exclusion has been increasingly documented in the social scientific literature since historian of science Steven Shapin drew attention to technicians' conspicuous absence from the descriptions of scientific laboratories. In a much-cited article, he suggested that the technicians of Boyle's laboratory in the 17th century were "doubly invisible" (1989: 556): firstly to Boyle and his philosopher-gentlemen peers, and secondly to contemporary historians and sociologists who study science (cf Daston 1992:611). Of course, 17th century sensibilities as to the relation between master and servant are no longer probable reason for the contemporary invisibility of technicians in laboratories - and the ethnographies about them. Shapin also maintains, however, that a major factor determining the various ways that different laboratories configure the relation between scientists and technicians today lies in the "evaluative distinction between skill and knowledgeability" (Shapin 1989: 562).

Organizational theorists Stephen Barley and Beth Bechky (1994) similarly offer this distinction between skill and knowledge (or even between experience and education), as one source of the ambivalent status enjoyed by contemporary scientific technicians working in two laboratories in a University in the USA. They also suggest it might account for the invisibility of technicians within the sociology of science literature, concerned as it is with "scientific knowledge" and not "scientific work" (ibid: 87). The importance of the distinction between what technicians know and what they do has been confirmed by many studies done with other types of technicians, ranging from emergency medical technicians to computer support technicians (Nelson and Barley 1992; Pentland 1991; Scarselleta 1992; Tansey 2008).

One of the functions of scientific technicians in these descriptions is to serve as "brokers" between "physical entities and a world of symbols that presumably represent the physical" (Barley and Bechky 1994: 89). Laboratory technicians therefore mediate between the world of instruments and procedures, and the world of ideas, data and publications - this latter being the realm of the "scientists". The technicians thus produce usable data for the scientists, and are responsible for the caretaking of the instruments that produce the data. By so doing, they serve as "buffers" between the scientist and the messy world of the laboratory (ibid.). Recognizing their interstitial role between the concrete and the abstract also brings into relief their not-quite fit into the categories of "skilled" and "knowledgeable", that manifests itself a "status ambivalence" (ibid.) in the context of a university workforce.

This literature pinpoints a lot of the characteristics of the LBA technicians that I have highlighted. Its focus on the role of the technicians as mediators between nature and the data is one I share, and the notion of "status ambivalence", resonates with the data cleaners' liminal position between researcher and technician, (as I describe it in this chapter and the fifth). Without wanting to lose the richness of these images of technicians "brokering" reality on the "empirical interface" (Barley and Bechky 1994: 98), I would like to recall some features specific to the LBA data cleaners, that I have already described in the previous chapter. It is important to reiterate my argument here because I intend subsequently to point out its flaws.

J and R both have master's degrees. Both have already worked on their "own" data. Both are keen to continue to do so, but for whatever reason are unable to. They have instead taken jobs as *bolsistas* working with other's data. These technician-researchers straddled two forms of recognition - money and scientific credit (cf Biagioli 1998). They thus felt immobile or excluded insofar as they were prevented from "carrying the data forward", without being able to, or even wanting to, fully express that. They also lacked the full job security that might come from starting a PhD. This endowed them with a specific sort of liminality.

Instead of attributing these researcher-technicians' liminal position to an institutionally-entrenched asymmetrical relation between knowledge and skill, I proposed instead the possibility that it was the product of differing relations that obtained between the data and the people in question; relations that were themselves emergent from the work being done on the data in order to carry it forward into knowledge. The data cleaners did not see themselves as just skilled workers, but as technician-researchers, and the work that they did on the data made it capable, in turn, of creating them as particular kinds of liminal scientific subject. It was precisely *because* this work was effective along the trajectory the data takes from mere "skill" to knowledge, which is also a trajectory along which ownership coalesces, I argued, that the data cleaners felt themselves to be liminal - not simply because they were considered to be liminal by the institution that they were in.

This relation between data and ownership was, I suggested, what generated the particular form of exclusion that I encountered when observing the LBA data cleaners. Whether or not it can be applied in other scientific settings can only be ascertained ethnographically, but my argument was that taking the relation that the technicians have to the product of their labour - the data - as a focal point could also be more nuanced way of accounting for the differences between different technicians in different settings. Rather than suggesting that their position was due to a singular (and erroneous) conception of the relation of skill to knowledge, this permitted the re-analysis of that very relation - which is what, in fact, these particular technicians were themselves brokering by being involved in the transformation of 'mere' data into knowledge.[115]

L's (the electronics technician) knowledge of electronics, on the other hand, did not pertain to the data at all. As he told me, he had no interest in what happened to the data after he had finished with it. His sense of immobility was only financial, and he did not expect to be able to make any claim on the data that he had helped create. Similarly, the forest technicians I have described, such as the data collectors and those who looked after the lodging at ZF2, were generally not at all worried about being excluded from gaining scientific recognition - even if, when I asked them, they said they would not mind being mentioned in publications. They were more concerned about their work being recognized and recompensed in other ways, through job stability and salary.

Within the LBA, then, different technicians have different relations to the product of their labour - the data. One interesting result of this difference was that several researchers and students I spoke to about the issue were very indignant and outspoken on the forest technicians' behalf concerning their lack of recognition in terms of credit, even though the technicians themselves in fact seemed to have very little to say

[115] Barley and Bechky (1994) suggest that comparative studies are viable given the relatively stable job title of "technician" across the US workforce, and, one might add, its unionization. I would suggest rather that it perhaps might be worth looking at the differences between them.

about it. The exception was the informant who was moving from data collector to data cleaner, who was quietly indignant that his work not be recognized in terms other than financial by those higher up in the LBA hierarchy when we talked about it; but I suspect this is because he was already "transitioning" into a data cleaner - that is, his relation to the data was, in fact, changing. But generally, I noted that the LBA forest technicians for the most part have no interest in owning, claiming or controlling the data at all.

The LBA forest technicians and other members of the Micro team who collected the data in the forest are only partly-involved in the systems of exchanges because almost invariably, they were circulating *around* these systems, being exchanged between projects and researchers who needed their help whilst in the forest, or ensuring vital maintenance work was done on the instruments on the towers. They never authored the exchanges, and were not concerned with the control of the flow of data.[116] The technicians, from this perspective, become a valuable *resource* - with all the non-agential connotations of the word - for anyone wanting to conduct research in the forest.[117] Several times during my fieldwork, there were skirmishes over who would get to have certain forest technicians for their projects - on one occasion, for example, J and E were both vying for one particular technician, who was at the time working at another research site on a different project. E needed him to put up her scaffoldings near Manaus, and J needed him to help find the new site for the SGC tower in the North of the Amazon. He did not have much of a say in which he did. What characterizes the participation of the forest technicians in these exchanges then is their exclusion from certain positions in them - they were never the ones making the exchanges, they were the ones being exchanged, as it were.

Concerning the invisibility of women in 17th century scientific practice, Donna Haraway has argued (*pace* Shapin) that a classification such as gender "is always a relationship, not a preformed category of beings or a possession that one can have"

[116] Nor, in fact, did they express any problems with foreigners - in fact, all of them told me that they had always had good relations with the foreign researchers they had worked with, and had in many cases learned a lot from them.

[117] In fact, it was a rule I heard often repeated by members of the LBA that no foreign researcher was allowed into the forest without being accompanied at least by a "mateiro", a "forester" - someone who knew the forest and could guide them.

(1997: 28). It is through their relationships - to men, to other entities - that women *emerge* as invisible, and this process is politically and intellectually charged. Gender is always "gender-in-the-making" (ibid.). The indiscernibility of women and the prominence of men in science is therefore, according to Haraway, a constantly emerging relational configuration. This configuration contributes to "the continually vexed boundary between the "inside" and the "outside" of science" (ibid.: 29). Haraway therefore explores how "most men and all women were *made* simply invisible, removed from the scene of action either below stage working the bellows that evacuated the pump or offstage entirely" (ibid.: 29 italics mine).

My description of scientific subjectivities-in-the-making chimes with Haraway's relational approach. The way that the forest technicians talked about their positions within the LBA certainly bears testament to the extent to which they feel themselves to be unconsidered or invisible. I would argue (and I have made a similar argument concerning the liminality of the data cleaners) that it was exactly through the forest technicians' relation to the data that their invisibility was constantly renewed. They become invisible because they did not have a relation to the data. They often seemed to have very little interest at all in it, or in what happens to it. Their invisibility was manifest in their incessant but arrhythmic circulation around these systems of exchange, alongside all the other entities that might be considered to be "merely logistics", and ignored in the same way.

But it is important to be careful here of the doubleness inherent in the invisibility of technicians - both within the eyes of the organization for which they work, and to the social scientist who analyzes that organization. This double invisibility relies on a singular notion of exclusion. The endless circulation of the workers and forest technicians around the system contains within it both their exclusion from the scientific credit system, and their exclusion from job security and financial stability. These, however, are not the same exclusion. Even if the forest technicians have a sense of injustice concerning their employment, they do not have this over their exclusion from intellectual credit. As far as most of them are concerned, they might be ignored intellectually by those who make use of their sweat and experience, but their battle lies elsewhere. Even if the forest technicians do not feel any relation or claim to the data, this does not make them excluded - what they are is in fact *indifferent.* In the case of the LBA forest technicians at least (and this became more evident the further from the office I went), what does make an increasing difference is financial security and stability. It is from this that they feel excluded; it is here that they feel their voices are ignored. They do not feel excluded from scientific recognition, because they do not seek it.

My model of data exchange and ownership - like most models - cannot take account of indifference to itself except by rendering those that are indifferent to it, invisible. The invisibility caused by their exclusion from what matters to them is therefore all too easily elided with the invisibility due to their indifference. Both are made manifest in the same circulating movement, and in the same status at the bottom of the LBA hierarchy.[118]

The Office and the Forest

How, then, can we take account of the forest technician's indifference, which is also what makes them unique? I propose that it is possible to subvert the two movements I have identified. Instead of seeing the movement in and out of the forest as constitutive of their circulation and therefore being an index of their exclusion, one can assume the perspective *of* that movement in and out of the forest. I argue that their movement in and out of the forest is what makes the forest technicians *specifically* what they are. This is not to say that their exclusion from international travel, and from job and financial security is in some way trumped or neutralized by this second perspective. One movement does not eclipse the other - they are in one sense, the same movement "seen twice" (Riles 2001: 69). Nor, however, does one movement exclude, explain, represent, or exhaust the other. Countenancing the doubleness of the forest technicians' exclusion does not involve merely revealing them twice over in a single fashion; it involves allowing them, in this double way, to be more than just "visible" or "invisible" from a given perspective. From the perspective of the data, the

[118] Shapin makes another very interesting observation concerning the place of the 17th century technicians as "sources of labor and muscular extensions of his master's will" (1989: 557), pointing also to the way that Boyle would rely on his technician's eyes and ears, in much the same way that J talked about the technician at SGC as her "eyes". That is to say, J does not do the same sort of work as the technician, but he acts as simply an extension of the work that she does. The question here is, what else does the technician do, apart from act as an extension of a network in which he has no apparent interest?

forest itself dwindled the further the data travelled from it. From the perspective of the forest, however, it is the data that dwindles, and then disappears, taking with it the researchers who come to collect it.

The journey into and out of the forest serves as a transformative axis. The forest technicians certainly had capacities that the researchers did not, but these were tied to the forest. I often heard a particular forest technicians being told that he was invaluable; another technician was one of the three people in the LBA who knew how to build a tower; no-one could drive a four-by-four like yet another, who managed to negotiate the most perilous tracks even in the pouring rain. Mark remarked with awe at skill with which one forest technicians drove a quad bike along the trail at ZF2, and at the dexterity with which the knots were tied that kept the precious instruments secure on the back of it. On numerous occasions, projects suffered from the lack of a forest technician to help install the equipment, or carry the heavy batteries, or erect scaffolding, or collect data. But these capacities meant little back in the office, and indeed, the forest technicians themselves seemed quiet and subdued when back in the LBA building. They appeared out of place, even, and uncomfortable. When I asked them, all of them - from data collector in the Micro group to the workers - told me that they loved going into the forest, that that was why they did the job. In fact, going into the forest *constituted* their job. Their capacities were expressed only in the forest, and this is why they were valuable but also invisible from the perspective of the system of exchanges.

The transformative potential of this movement in and out of the forest was demonstrated very forcefully to me during a trip I took near the end of my fieldwork to the newest of the LBA's research sites, known as Balbinas. This was the site where a 300 m tower (known as ATTO) was going to be built, along with several smaller towers; these smaller towers were in fact themselves a smaller project, known as CLAIRE. Both projects were the result of a collaboration between INPA/LBA, UEA, and the Max Planck Institute in Germany, and had been given a top priority in the LBA agenda. What this meant in practice was that the forest technicians were almost all working there, which caused some grumbling amongst the other members of the LBA who needed them. The site was still in the early stages of construction during my fieldwork, and severe rains had delayed the progress of the work considerably.

The research site was on a Biological Reserve called Uatumã, 330 km northeast of Manaus. The access was very difficult. A trail had to be built to even reach the chosen research site, and the river had to be full enough to carry the building material to the trail, but not so swollen it was dangerous. Every so often, an enormous truck would draw up outside the LBA building in Manaus, and deposit a group of extremely muddy workers, and several broken quad bikes, and barrels of rubbish onto the pavement.

The first towers to be built were the smaller CLAIRE towers, which were to be used to collect pilot data on trace gases. I had taken a trip up with two members of the Max Planck Institute who were in charge of the German side of the CLAIRE project, A and B. We were travelling with H, the chief technician of the LBA. H was something of a mythical figure in the LBA, having been in charge of tower construction since its very beginning. He was generally reticent and quietly spoken, but no-one in the LBA was in any doubt that he had authority. A and B wanted to conduct some pilot measurements from the tower that had already been built. In order to do so they needed a generator to provide a clean energy supply for their instruments, so we were also travelling with the electrician the LBA used, who was an old friend of H's. My informal role was as translator for A and B, so that they could communicate with H and with the team of workers who were constructing the site and the tower.

A and B were also planning to confirm the site for the next CLAIRE tower. The director of the Max Planck department had given the coordinates of the new site to them. Three or four towers were to be built, and certain scientific criteria had to be met concerning their positions in relation to each other, in relation to the prevailing wind, and in relation to the topography of the land and the homogeneity of tree-top cover. [119] Using these criteria, the coordinates had been decided on back in Germany using Google maps, and duly brought out by A, the project manager, to present to H. However, as we made our way through the forest with a GPS to the spot that had been selected – A casting her eyes around warily for snakes and already complaining of the heat – H started to shake his head. "This is a really really bad spot" he pronounced emphatically when we had arrived. It was full of dead or dying trees, which would

[119] This was due to considerations of the towers' "footprints", as discussed in previous chapters.

soon fall, not only endangering the tower, but also creating clearings in the canopy. The site of the first tower had already been changed once due to its proximity to a clearing in the canopy, which would affect the turbulence rate of the wind above the trees, an important aspect of carbon flux that needed to be kept as homogenous as possible. H wandered around with the GPS, trying to find a suitable spot, as A wilted under a tree and protested that they should try to use the coordinates that she had with her. Eventually, H found somewhere he liked the look of a little better, although he was still not terribly happy with it. A was also nervous about it – she knew that her boss had very carefully chosen the previous coordinates and it was up to her therefore to decide against that choice in her boss' absence. For everyone present, however, including her, it was clear that she had no real choice. H's experience of the forest and knowledge of building towers had to be the basis for the decision. What made her nervous was knowing that she would have to explain to her boss back in Mainz why she had made the decision to change location. That is, taken out of the forest, her decision would look much less convincing.

This event stuck in my mind because it was such a clear example of the way knowledge and skill change as one moves in and out of the forest. We were in the depths of the Amazon, with no way for A to contact her boss. In this place, H’s knowledge and skill mattered more. But she also knew that trying to convey this to her boss *when back in Mainz* was going to be difficult. Thankfully, B told me, the whole of the ATTO and CLAIRE committee had taken an expedition to the forest the year before, in the planning stages of the project. One of them had almost collapsed in the heat, and they had all got soaking, hungry and wet whilst trekking through the forest. So they all knew at least a little, B said, about what it was like to try to construct a research site in the forest. But even so, A's boss might find it hard to understand. This was not a case of H knowing about the forest, and the scientists not knowing. Both knew about the forest, but in different ways. And it was the movement into and out of the forest itself that changed the relationship between these two.

The transformative aspect of this movement was also apparent during the more quotidian trips into the forest to the research site ZF2. Driving off on one of these was always exciting. Those who usually went from the Micro group were used to making this trip together, and would always stop at the same service station on the way to pick up newspapers for those staying in the forest, and coca cola and biscuits to supplement lunch. If there was no-one more senior accompanying them, they would joke and gossip continuously. Due to the shortage of cars, we would often share space with data collectors from other LBA groups who also needed to go into the forest, giving the excursion an added hint of novelty and variety. Even if someone more senior was also on the trip, there was always a more relaxed and informal feeling than in the office. This would intensify as we went further away from the city. As we turned off the motorway and into the dirt track, people would usually start telling stories of the animals they had seen in the past, how many jaguars, snakes or packs of jungle pigs, vying in a friendly fashion for the most dramatic sighting. Pulling up to the *alojamento*, there would be affectionate shouts of greeting exchanged, and more jokes, before people would start preparing their equipment to go their separate ways into the different data collection sites in the forest – normally either to the hydrology sites along the streams (*igarapés)*, or the leaf litter collection sites for biogeochemistry, or the towers.

There was a markedly different rhythm to life in the forest. Those who came from the Micro group would usually spend a great deal of time up the tower K34, reprogramming the dataloggers, or installing new instruments, or trying to work out what was wrong with the sensors. This was often a slow process, because everything had to be first ferried up and down the tower. If they were working on instruments at the top of the tower above the canopy of the trees, the heat of the sun would sometimes impede their work. Although there was meant to be an internet connection at K34 that allowed whoever was on the tower to speak via skype chat to R or J in the LBA office, the antennae were often struck by lightning, making the connection rather unpredictable. Sometimes times it was impossible to know whether the problem had been fixed until we arrived back at the lodging and were able to radio the LBA office. Sometimes, the problem could not be solved in a day, and so we would have to spend another day or two unexpectedly at ZF2. But in the context of the forest, none of these occurrences seemed like mistakes that could have been avoided; it was just what was involved in trying to install or fix equipment in the middle of the Amazon. It took time, and it did not always work. As L explained, soldering battery connections in the lab, for example, and soldering them in the forest were two very different tasks.

The forest technician I spent most time with was the hydrology data collector. I often accompanyed him as he did his rounds of data collection from one data collecting platform (*sessão*) to the next. I was astonished at the distances he would walk (at times by himself) through the forest, sometimes carrying heavy equipment. Sometimes we passed big trenches or trees with rusty tags attached, remnants of past experiments and now grown over with vines and covered in forest debris. He would spend hours in the forest by himself, collecting data from the various different collection points, or sitting by the side of the *igarapés*, monitoring their flow rate and conductivity. He did not get bored, he told me, although sometimes he has been scared – especially when one researcher asked him to collect data at night. "You get totally paranoid about jaguars!" he said, laughing. The equipment he dealt with was sometimes only half-working, and he knew all of the foibles and flaws of each data system he worked with – "each sensor has a different problem" he told me. I asked him if doing the same thing every week got boring. "Sometimes" he said. "But I think it's important. And every week is different – even if I have already understood the logic of the work". He enjoyed having the time to think, and he enjoyed being in the forest. I ask him if the researcher for whom he had worked at night had thanked him in his thesis, and he simply shrugged, not bothered.

At lunchtime and in the evenings at ZF2, people would converge back on the lodging again for meals, which the chef masterfully served up using the food that was brought in every week from the city by the LBA logistics team. Rice, beans, and some sort of fried meat was usually accompanied by salad – at the beginning of the week at least, as by the end the fresh produce was a little wilted. If we were staying for several days, hammocks would be slung up on the second floor, and showers taken before sitting down to eat. And after dinner, we usually played a game to see who had to do all the washing up. This consisted in trying to guess how many beans each person had in their closed hands, and always ended with whoops of laughter as the two last people engaged in a dramatic duel of bluff and counter-bluff, with the pile of dirty plates and dishes that the loser would have to scrub looming ominously on the table behind them. If it was not too late, someone might challenge someone else to a game of pool on the pool table that had been set up in one of the back rooms of the building, amongst all the old gas analyzers and data logger boxes. By that time, however, most people were exhausted from the day's work, and would be in their hammocks, several snoring loudly. You had to get as much sleep as you could in case the guariba monkeys woke you with their howling in the early morning. The contrast with the air-conditioned and contained nine-to-five routine of the office could not have been more stark.

But people could also be excluded from this sense of community – not in an intentional fashion, but simply because of their temporary status. This was most apparent when foreigner researchers came. Unable to keep up with the banter in Portuguese, they would end up spending their time in the *alojamento* reading or on their laptops. Several knew at least a little Portuguese and would try to join in the conversations from time to time – but even though there were rarely any obvious limits demarcated, there was still a tangible sense in which those who worked all the time in *campo* were a tight-knit, and to a certain extent closed, group. Being someone who obviously enjoyed being in the forest granted you some access, as I found out after a while. But this sense of community that the forest technicians had was very strong and never more so than when it was threatened, as was demonstrated when a foreign researchers' data storage device went missing. This contained all the data that the researcher had collected over several weeks. After looking for it in his belongings and all around the lodging, he had to tell the head of logistics. He took it very seriously, and a sudden pall was cast over ZF2. The forest technicians and logistics team told me that they had been informed that "everyone is a suspect". They were indignant at such an accusation. They rallied together, and told me, and anyone who would listen, that it was ludicrous and offensive to think that one of them had taken it. The device turned up in one of the cars a day or so later. But the incident demonstrated how immediately the ZF2 "crew", as I think of them, came together at the sign of any threat – as a family might. As one of them told me: "we’re more like a family here". This sense of family was not restricted to the research site ZF2 – as they circulated around all the different projects, the workers generally all knew each other, and had done, in some cases, for over 10 years. At any one research site most people would know several of the other people there. There was always a sense of cohesion and belonging.

But the further one went into the forest, the more one needed knowledge that was not one's own. In the more isolated research sites, the LBA projects were dependent on local people who had very little if any interest in the research the researchers were conducting. In order to gain legal access to a protected area such as Uatumã – which would allow the researchers access to highly desirable "undisturbed" forest – the project had to employ members of the riverine communities that lived on the reserve as workers during the construction of the site.[120] During the time we spent at the emerging research site at Uatumã, we ate food that the local people had caught and fished. On one occasion, one of the local workers picked up a tortoise to take back to cook for dinner. A was horrified and, gesticulating widly, begged them to let it go. Laughing at her, they did so, but after she walked on I saw them casting about for it again – a tortoise is a particular delicacy. I asked them what they thought of the tower being built, and if anyone had explained to them what it was for. "Yes, [H] explained. It's for monitoring leaves or something" one of them told me vaguely. Utterly indifferent to the scientific possibilities of the tower, they wanted to know if I could ask H to put an antenna on the top of it so they could get a better signal on their mobile phones.

"What happens in the forest, stays in the forest" (*o que acontece no mato fica no mato*), the forest technicians would jokingly say to each other from time to time. They meant that secrets could be kept in the forest, but also that there was a safety there, a sense of belonging impervious to the outside. This very clearly, albeit unintentionally, captures the dynamic I want to draw out. What "happens" in the forest is, from one perspective, *data*, and data certainly does not stay in the forest. It has to come out in order to become certified data. But the forest technicians do not have any interest in the data. The movement in and out of the forest is transformative, in that moving in and out of the forest served as a sort of scaling device by which certain knowledge assumed more or less importance: what was important in the forest was not useful in the office, and vice versa. But this movement also mediates between two very

[120] I tried to gain access to these communities, in order to investigate how they saw the arrival of this flurry of scientific activity in their territory. I went to speak to the NGO that was in charge of the relation between the Reserva and the Government, and they seemed amenable to my plans. However, the LBA researcher in charge of the ATTO project told me that I was either part of the LBA team, or working with the communities. It would look very bad for the LBA if I did something inappropriate whilst working the communities, as the LBA were responsible for me.

different places. These are certainly not mutually defined as each other's opposites, but each appears very differently when viewed from the perspective of the other.

Indifference as Relation

If, in Chapter 5, I made a case for a certain sort of scientific sociality, there is a case here for an alternative sort of sociality, one which depends instead on work that is "not scientific". This work happens out in the forest, in the periphery, or at the margin. As far as the researchers are concerned, the part of the forest that the technicians deal with is precisely the part that must be excluded in order for *them* to make *their* sort of knowledge. They are made invisible, along with the forest itself and all sorts of other entities and relations, in order that the sorts of singularizations that are characteristic of scientific knowledge can be carried out, as discussed in Chapters 1, 3 and 4. If we are to take the two locations as different perspectives on each other (Strathern 1991), then the technicians should be seen not only from the perspective of the office, as brokers between scientific world and forest, but also from the perspective of the forest. Here, they appear as experts in the forest, or in what is "not scientific".[121] If we take both of these two locations, and the movement that mediates between them, seriously, what emerges is not simply a paradigm in which the researchers are knowledgeable whereas the technicians are skilled. Instead it becomes clear that the importance or power of various capacities, all of which can be parsed into "skill" and "knowledge" if desired, changes as you move in and out of the forest.

In this description, the incessant circulation around the systems and the movement in and out of the forest are in a sense, the *same* movement. However, they are the same movement "seen twice" (Riles 2001: 69). One can separate them, even though they occur at the same time, and are the same thing. They are the same thing differently. On the one hand movement in and out of the forest is a marker of the forest

[121] In order to complete this switch analytically, I would of course need to describe how the researchers could also become something else from the point of view of the forest. Without having a very clear idea of this, it is certainly true that the researchers are also transformed by their trips “into the field”, although this is slightly more difficult to describe as they still came into the forest for data. Nevertheless, the Wendy Houses they built to protect the instruments that were attacked by termites or colonized by bees, seemed like very clear signals of the forest’s indifference to their attempts - the forest’s lack of a relation with their data activities.

technicians' exclusion: it is contained within and constitutive of their circulation within the system. On the other hand, however – or rather, from the perspective of that very movement – other possibilities open up. The movement itself starts to "scale" the analysis, and becomes a transformative axis along which the technicians and the researchers can both travel. As Marilyn Strathern remarks, "it is possible to say of places that a place is both a point along a scale – as one travels from one to another, places seem separated by distance that can be measured – and is also the only point at which one can ever "be" – the place from which all distances are calculated" (2002: 92). In this sense, the forest and the office are both simultaneously points along a scale, and perspectives from which such points may emerge. However, this potential for figure-ground reversal also contains within it its own potential for transformation. Making the forest technicians visible does not necessitate including them in the systems from which (from the perspective of that system) they seem to be excluded. But equally, remaining sensitive to their exclusion does not mean desensitizing oneself to what escapes that exclusion. The index of what escapes exclusion is their *indifference* to at the system of data exchange or scientific recognition.

Indifference proffers itself ethnographically as a challenge to analyses that turn on notions of inclusion, exclusion, visibility and invisibility. It is because of the forest technicians' indifference to the systems that I was trying to describe that I was forced to acknowledge the potential that their movement in and out of the forest had for being something *other* than a movement of inclusion or exclusion. I suggest it provides its own system altogether, with its own choreography of elision and separation from the other logics, movements or socialities. If the technicians are excluded from one sort of movement along a socio-economic scale that would allow them to "progress", they are also at the same time performing another sort of movement, which constantly shifts them onto another scale altogether. This scale at times touches the previous one, but it is not encompassed by it.

Indifference is also at odds with stories of resistance. Resistance appears when questions of inclusion and exclusion are at stake, and when hegemonic discourses are understood to be both salient and contested. Through a sensitivity to exclusion and resistance anthropologists often try to attain a sort of double vision, in order to be able to see both the dominant and the subordinated at the same time (cf Haraway 1997: 38-39). This double vision may even reveal ways in which the two are mutually dependent upon the other (Coombe 1998),[122] so that their relation is a "continuous mutual disruption – the undoing of one term by the other" (Coombe 1998: 9). Indifference, however, is not disruption, nor is it resistance. It is not even "friction" (Tsing 2005). It is, I suggest, in analytical terms at least, the index of difference that most escapes our analyses. The transformative axis "office-forest" does not only provide a means for each pole to give a perspective on the other. It also allows for one extremity to slip out of focus altogether. The local people at Uatumã, for instance, are not determined by their relation to the LBA tower, and do not care about the data it produces. It is, however, remarkably hard to keep a grip on indifference in one's description; it too slips out of focus. The movement in and out of the forest, that takes the forest technicians away from concerns about data and ownership, itself becomes constantly re-engaged in that from which it provides an escape route. The movement in and out of the forest is at the same time the thing that separates, and the thing that connects, these two perspectives.

Conclusion

In this chapter, I have approached the subject of the forest technicians, and their exclusion from the systems of scientific claims. In a broader sense, I have explored the extent to which exclusion and inclusion are dependent upon singular perspectives. As a result, the exclusion/inclusion duality can sometimes shift in and out of focus. In this case, it was the technicians' indifference that served to effect this shift.

I have described the forest technicians' sense of instability and precariousness within the LBA organization, as well as their restriction to a certain position within these systems of data exchange, and a certain movement in and out of the forest. Although

[122] As Rosemary Coombe writes of resistances to IP regimes: "the law must be understood not simply as an institutional forum or legitimating discourse to which social groups turn to have preexisting differences recognized, but, more crucially, as a central locus for the control and dissemination of these signifying forms with which difference is made and remade. The signifying forms around which political action mobilizes and with which social rearticualtions are accomplished are attractive and compelling precisely because of the powers legally bestowed upon them." (Coombe 1998: 37)

the forest technicians certainly are excluded from the networks of exchanges, it is also true that they were to a certain extent indifferent to these networks - they had no interest in carrying the data forward into knowledge, or claiming it. In fact, they presented a particular sort of double situation, in which they were both excluded, and also just absent. Even though the exclusion they felt most keenly was a financial one, they cannot be exhaustively described through the idiom of exclusion and inclusion. Their exclusion was manifest in their constant movement in and out of the forest. This movement is at the bottom of the hierarchy of movements that those in the LBA perform. However, the technicians' indifference to this hierarchy allows this movement to become its own, separate perspective. I have tried to explore the various entanglements of the different types of movement, as they separate themselves from and then come to depend upon each other for significance.

I have argued that the movement between office and forest has a radical transformative potential for skill and knowledge.[123] It becomes a scale along which the capacities held by the forest technicians can move, assuming different proportions in and out of the forest. Thus the office and forest are connected through this scale, such that all who travel along it must suffer certain transformations. However, the forest and the office are also different perspectives, giving rise to two separate socialities. They create two distinct spheres from which people can be included or excluded. The forest technicians' indifference, I suggest, is an escape from this oscillation between perspectives; but it remains elusive to describe or to use, for in every attempt, it becomes connected up, or related in some way, to some discourse of exclusion.

[123] The movement from office "into the field" is of course a transformative one in anthropology as well, providing a comparison worth pursuing at a later date.

Conclusion

The Nature of Data - summary of argument

Over the course of this study, I have followed the LBA data from collection to dissemination. I have delved into the specific ways in which those researchers and technicians I worked with effectively made information about the Amazon forest. I have described how data comes to be data - why certain information is considered to be data, and other information, error. I have dwelled on the processes by which data becomes certified and trustworthy, in order for it to be taken up subsequently and employed in the production of scientific subjectivities, scientific communities and claims about the world. I have demonstrated how and why the production of the LBA data is a complex process that overflows itself, and is not sufficiently described using the conventional and particular terms of a representational, binary idiom. But in refusing the language of representation, I do not intend to deny it. I have sought, instead, to *extend* this idiom substantially in each chapter. As I proposed at the beginning of the thesis, this is the essence of the oscillation between identity and difference that characterizes what it is to take science seriously anthropologically. The scientific practice of the LBA seems to be both representational, and non-representational. These two identifications hold within them a myriad of different binary opposites. The LBA simultaneously affirms such descriptions, and breaks away from them. In the same way, studying the data that the LBA produces engenders a simultaneous sense of identity and estrangement. Those researchers I worked with do understand reality as singular, and nature as given and they do not imagine that they are creating that which they study. Yet at the same time, my research demonstrates that if one draws out the subtleties of their understanding and their practice then it becomes clear that they do not simply make singular representations of a universal reality. It was the aim of this anthropological description to take account of this oscillation, rather than to deny it by demanding that matters come to rest on one side or the other. My aim was, as Bill Maurer puts it, neither to get "back to business" with empirical fervour, nor to merely "add more voices" (Maurer 2005: 11).

In the first chapter, I described in detail the attempts of a student from the LBA to extract data from the forest. In order to know what she was measuring, she had constantly to battle the threat of the incessant relationality that constitutes the Amazon forest. I argued that her efforts, augmented in great part by the use of different scientific instruments, worked to introduce a position of zero relationality into the forest. Only once such a position had been attained was she able to start measuring the particular relation she was interested in. Describing or understanding what she was trying to achieve demanded a shift in how the binary subject/object (constitutive of representational logic) is understood. It is often assumed that when there is more subjectivity, there must be less objectivity, and vice-versa. However, the notion of objectivity that this implies is clearly not appropriate to capture how E is producing data. She is keen to remove *all* superfluous relations, including not only those that might be considered "subjective", but also those between other objects in the forest, herself, others involved in the project, and so on. Far from denying that her work was reductionist, I argued that it was in fact much *more* reductionist than is assumed by studies that focus on critiquing only the exclusion of the scientist-as-subject from scientific knowledge. The specific relation, so dear to the social sciences, between human subject and inanimate object was, as far as E was concerned, just one amongst many relations that needed to be excised, silenced or ignored. Why, then, I concluded by asking, privilege it amongst all others as the basis for critique? Why not also question the removal of the relation between, say, isoprene and honeybees? Or any of the other myriad of relations that E must occlude in order to achieve a moment of singularity and stability in amongst the quagmire of relations, all in turn relating to each other, that is the Amazon forest?

In the second chapter, I looked again at the relation of nature to data; but I did so through an examination of the mechanics of measurement as the vector of quantification. The necessity of discretization to the process of measurement was apparent in my ethnography. Measurement, one researcher told me, "breaks up" the world. This was an action that occurred often during my fieldwork. The footprints of the tower, parcels of air, molecules of gas – all represent instances of this "breaking up". In as much as discretization is done *to* something, this implies a conception of nature or the world as not discretized. It implies that the "world is in analogue", or that the world is "continous", as my informants would say. I discussed at length the ways in which this trope of discretization is ubiquitous. It does not only appear in the specific terms of digitizing the world. It also underlies quantification as the basic scaling practice of observational science. I explored measurement's most crucial properties using another infamous example of the same dynamic – origin myths, in this case Amerindian. Having shifted the descriptive dynamic from a binary between "real" (world) and "representation" (measurement) to one between the continuous (world) and the discrete (measurement), I argued that these two dynamics do not map onto each other. The discrete is not any less real than the continuous. It is real, however, *in a different way*. I pushed this difference even further by suggesting that the reason that reality does not share its conceptual coordinates with nature is because measurement is actually what introduces the notion of reality into the world. Reality is not “in nature”. Nature, for my informants, is given and unquestioned because it is that which is being measured. Reality, on the other hand, is the form that nature takes when it is measured, or known quantitatively. It is only real because it could be false. For the data to be convincing, it is what it must be shown to be. Knowledge, for my informants, is therefore dependent upon uncertainty, the inevitable fact that what is known is also not-known. It is for this reason that the metrologists, those researchers I accompanied who specialize in mensuration, attend to uncertainty as constitutive of measurement. I concluded by proposing an asymmetrical tripartite understanding of scientific cosmology: nature, uncertainty, and *then* reality/representation. Uncertainty lies as the *means* by which the first is produced from the last. Again, this schema did not aim to deny the representational logic, but to extend it.

In the third and fourth chapters, I examined how the raw data becomes certified data. Data is separated from nature in the same movement through which it is made more and more real and trustworthy. In Chapter 3, I demonstrated how raw data can be used to develop an understanding of the relation between ambiguity and singularity of meaning in the science of the LBA. Anti-representational ANT analyses of network building and the construction of black boxes depend upon opposing ambiguity to singularity. It is their claim that a fact becomes true when its meaning is no longer disputed. The raw data complicates this dynamic somewhat. It appears not as a singular entity, but rather as a "unique ambiguity". Dwelling on this tension in the raw data, I drew attention to how sought after and powerful ambiguity can be when it is unique. Researchers from all over the world come to the Amazon and undergo all sorts of ordeals in order to obtain raw data. Summoning the understanding of nature as chaotic relational excess that I documented in in the first two chapters, I also demonstrated that the "rawness" of raw data is a result of its hyper-relationality or connectivity to the world. It is raw data because it still has too many of the wrong relations in it, relations such as that between the instrument and bees, or lightning and the tower. It therefore has the possibility of being either data, or error – it contains within it the potential to be real or false. In fact, it *is* neither of those things. Instead it encapsulates the tension between the two, or the potential for the relation between the two to emerge. This, as I also argued in Chapter 2 is the effect of measurement: to introduce the relation between real and unreal, or data and error. Attending to the raw data in this way, as it is and not for what it will become, establishes the necessity of taking account of those entities that in a sense pre-exist the networks of ANT, if those networks are taken to be a function of singularization of meaning, or "black-boxing".

I also drew attention to the implicit temporal bounding that occurs in anti-representational, constructivist accounts of the fabrication of facts. I suggested that there is always something that pre-exists the account, and, perhaps provocatively, called these "somethings", "pre-network entities". They are examples of what Bruno Latour calls in one case a "vague, cloudy grey substance" (1999a: 45). If I am right, what is important is not creation or construction, but rather the substitution of the "right relations to the world" for the "wrong relations to the world". The real-represented binary might then be to revealed as a data-error binary instead, where error is all the bits of the world that are not under investigation - ‘the rest’, as it were.

In the fourth chapter, I turned to the process by which this relation between data and error is solidified and fixed. This is called "cleaning the data", and consists in weeding out those wrong relations to the world that constitute the data’s "rawness". Effecting this transformation of the raw data into certified data is no mean task. There are several key mechanisms by which it occurs. The first is the transferral of the raw data into a pre-existing structure in order to turn the undifferentiated mass of numbers into values of phenomena, such as wind speed, or carbon dioxide concentration.

These values are then verified by ascertaining whether they have transgressed the "range" or not. If they have, then the data cleaners begin painstakingly adjudicating whether they are still what one of them referred to as "inside the possible". This involves relating the data values to each other, so that the phenomena in question are re-composed inside the data set. Is it possible that such a value could exist, given the other values that it must relate to? If the answer is yes, then the data is data – and has always been so. If the answer is no, then it is error – and has always been so.

This seems, if time-consuming, at least straightforward. However, what was remarkable was the self-referentiality of these procedures. The range is itself composed of data, and making the data set coherent is achieved, not by establishing its external relations to nature or the world from which the data has come by but instead by establishing its internal relations to itself. There is no reference backwards, or if there is it is only to the *relatório* detailing the conditions of data collection that accompanies the raw data. Moreover, this *relatório* does not accompany the certified data after the cleaning process. It is left behind, and in fact needs to be left behind, in order for the data to become certified and travel onwards and outwards from the LBA. Far from needing epistemological justification, this circularity is necessary for the data to arrive at the state of being referent, or reality, to *itself*, a state in which it is both representation and referent, simultaneously. Certified data therefore reconfigures conventional ideas of correspondence as the representational relation by turning itself inwards: certified, trustworthy data corresponds to itself.

I concluded the fourth chapter and first part of the thesis by introducing the notion of "compaction", drawn in part from the work of Donna Haraway and Marilyn Strathern, in order to clarify this relation of internal correspondence that certified data displays. Through the cleaning process, one side of this internal correspondence compacts into the other. The data becomes more and more real the further away it gets from its raw state, and the further away it gets from the forest; as Bruno Latour suggests, the more work is done on it, the more real it becomes. However, what compaction stresses (that ideas of "black boxing", for example, do not) is the collapse of one side of the reality-representation relation into the other. Compaction is not the progressive eradication of the other side, and neither is it hybridisation of the two. Compaction is when the analogy caves in altogether – but only in order to be separated out again. Separation is the subsequent drawing apart of the terms of the relation, and is thus the realization of the potential for meaning that is inherent in the collapse. Therefore, the fourth chapter ended with a brief synopsis of the process of data production, that aimed to re-think the reality-representation relation in terms of this dynamic. Data begins when "world" as an undifferentiated substrate is separated ("discretized") through measurement to give uncertainty and raw data. Raw data, as an ambiguous material, is then separated through data cleaning to give error and certified data. Certified data is itself a compacted form of the relation between reality and representation, which can go on to be separated in different ways.[124] Undifferentiation, ambiguity and compaction emerge in between different moments of separation. None of these stages – world, raw data, certified data – are either realities or representations. They are a related series of different separations and compactions. They can be seen as real, or as representational or they can be given any number of other denominations.

The Social Life of Data - summary of argument

I ended the first part of the thesis by suggesting that it is the capacities given the data by becoming certified that allow it to become a social, and in a sense socializing, scientific object. This is the theme that animates the second part of the thesis – what might be called the "sociality" of the researchers and technicians of the LBA. This part is made up of two chapters. In the first, chapter five, I explored the way that the LBA data is capable of engendering boundaries and differentiations. I identified two different models of data exchange for the LBA data, models that give rise to two different kinds of scientific community. One is premised on a logic of "flow", and the other on a logic of "exchange". In the first, data is open-access and free, and the point of the exchange is in fact to overcome boundaries, generating the notion of an inclusive, limitless scientific community. In the second, the data flow is controlled through a series of bound exchanges, and is not free. As one of the tenets of this second model is that one has to have a relation to it in order to become a part of it, this system generates and perpetuates its own mutable limits. These limits are keenly felt by those involved. The two models, it bears insisting, are separate and distinct: one is not an idealized version of the other. Their co-habitation accounts for many of the

[124] In Chapter 5, I suggested that one of those ways is for it to be both means and end to knowledge, which might also be seen as the distinction between form and content.

misunderstandings (such as contested accusations of data theft) that occur during scientific collaborations.

In the second part of the chapter, I turned to address the individuals involved in these communities, and their relations to the product of the communities, the data. I argued that there is a relation of mutual co-production between people at the LBA and the LBA data. The data cleaners work on the data, but the data also works on the data cleaners. Many studies of scientific property that have emphasized the impossibility of ownership in scientific practice. In the LBA, however, I found that the crucial relation between data and researcher or technician was indeed understood to hinge on the question of ownership. There was an intricate interplay between "ownership" and the transformation of data into claims about the world. The data cleaners are not in a position to convert the data that they produce into knowledge, but they have invested significant work into it. I suggested that this work gives them a sense of entitlement towards the data that they cannot convert into anything else. Unlike other researchers, who do have the means to convert data into knowledge, or technicians, who do not work on the data at all, the data cleaners are suspended between owning and not owning the data, and therefore between scientific subjectivities. This gives them a feeling of liminality, or of being stuck. The aim of the chapter was to re-animate discussions concerning creativity and subjectivity in scientific practice, by elaborating a generative system of scientific sociality that turns on relations between people and things, most notably data. It is a closed and self-referential system of mutual co-creation.

The sixth and final chapter then challenged this closed and self-referential system by asking how those who have no relation to the data at all could figure in such a system. I denominated these people the "forest technicians". For the most part they are local people from Manaus, who go in and out of the forest collecting data, constructing research sites or ensuring the research sites are maintained in working order. The system sketched out in the previous chapter seems to necessarily place these actors in a position of being excluded from it. This is certainly how they appear, endlessly circulating around the research sites, going in and out of the forest. Paying attention to their exclusion, however, indicated they do not in fact feel excluded in this way. Any sense of exclusion they have relates to job security and finance. In terms of the system of data exchange, they feel indifferent. If their indifference can be masked by the imputation of exclusion, their incessant circulation as index of their exclusion also contains another movement, the movement in and out of the forest. From the perspective of the forest, the forest technicians emerge as agents of their own sort of sociality at the research sites. Although these two domains – the office and the forest – are opposed, it is the movement between them that is the index of their indifference. This movement is transformative, and is neither opposed nor aligned, but mediative. As such, the forest technicians’ indifference appears ethnographically as an important but elusive relation. The implicit concern of this second part of the thesis was to explore how conflicting discourses may not simply reflect each other, but may instead provide entirely new avenues for enquiry. The final chapter therefore suggested that informants’ "indifference" is one relational means to mediate such enquiries. Indifference treads between exclusion and inclusion, visibility and invisibility. But it also inevitably becomes entangled up with those notions, negating itself in the descriptive process.

What Next?

As well as being an ethnography, this thesis was intended as a re-description of scientific practice. The study of data in the LBA presents an opportunity to re-consider the real-representational relational idiom that has long dominated social studies of science, by dwelling in the oscillation between recognition and estrangement that studying science elicits. In this sense, this thesis has constituted a sustained effort to take science seriously anthropologically.[125] It was also intended as an example of making explicit the recursivity that is built into such attempts. The extension, or re-working, of representational logic effected in the first part, and the reconsideration of scientific sociality pursued in the second, are both recursive moves.

[125] This allows for comparisons to be made with other knowledge-practices that have previously been imagined in only one way. At the end of the fourth chapter I suggested one example might be different processes of what in the West might be thought of as objectification and subjectification, or what I call "ontological alchemy", in Amerindian societies.

Both are intended to refigure the forms that come most easily to hand when anthropologists seek to describe scientific practice and knowledge.

Another, perhaps less obvious, set of targets for such refiguring are the notions of complexity prevalent within anthropological theorizing. In the LBA researchers' understanding of the world, the whole seems to end up being always beyond or bigger than the sum of its parts, precisely because it always has the capacity to be smaller. This is, I suggested, a result of the fact that the forest, and by extension the world, is always potentially further divisible. There are always more relations inside the relations, more terms inside the equations, more scales inside the scales. However, data, because it carries its own reality with it, is able to become *more* real than the world. This excess in the data is compounded by the relational capacity that the data is imbued with. It can travel, be copied, indexed, shared, compared. It outstrips the world as it leaves it behind. In this situation the world somehow becomes *smaller* than the sum of its parts: the parts become bigger than that which they are about.[126] As Lev Manovich describes it in his account of the database as cultural form, "Jorge Luis Borges’s story about a map which was equal in size to the territory it represented became re-written as the story about indexes and the data they index. But now the map has become larger than the territory" (Manovich 1999: 85).[127]

Thus one avenue for enlightening and recursive comparison between the LBA practice and anthropological knowledge is mereological. My description of an ever-expanding world, as seen through the eyes of the researchers of the LBA, will not be new to anthropologists. The impossibility of dealing with an ever-increasing complexity has been thoroughly addressed in Marilyn Strathern’s work, most notably in *Partial Connections* (1991). She describes the dizzying complexity that the anthropologist faces in trying to organize his or her data across different scales. Trying to hold something steady by which to compare generates the sensation of endless insufficiency. The closer one looks, the more internal variation appears, eclipsing the importance of the initial boundaries drawn around phenemona in order

[126] I thank Martin Holbraad for this insight.
[127] See also Geoff Bowker (2008), for an investigation of the context and repercussions of this shift for memory practices in the sciences.

to organise one's knowledge of them. This sensation, Strathern maintains, is derived from an understanding of the world as composed of discrete entities that can only ever provide a partial perspective on the entire phenomenon – what Strathern denotes a "pluralist" perspective.

Strathern's alternative trope is that of the fractal – a form in which complexity does not in fact increase but is conserved at every level. Yet although this proposal has proven immensely fruitful for social scientific theorizing (see for example Kelly 2005; Jensen 2007; Holbraad and Pedersen 2009), it is ethnographically bounded and to a certain extent specific. Although the fractal form employed is a Western one, the motive for employing it is Melanesian. Because pluralist Western understanding of the relation of parts to wholes that encompass them (as evinced by the relation of the individual to society, for example) does not underpin Melanesian notions of sociality or knowledge, it is therefore insufficient to describe them. The fractal in Strathern's description thus becomes the form that spans the particular ethnographic engagement and the recursive move that resulted from it. As a descriptive device, it neatly encapsulates one particular ethnographic relation. In the case of my thesis, however, with its particular ethnography, a different form might be required. Such a form would have to confront the specific difficulties involved in studying science. It would have to take account of the specific mereology of data, that generates the sensation that the world is endlessly smaller than the sum of its parts.[128] It could be used to rethink the discussion of partiality, as itself a contemporary anthropological merography,[129] offered in the introduction. If I do not offer such a form here, this thesis is nevertheless a step towards identifying the geometry it might take.

A reflexive comparison of this sort represents an attempt to productively negotiate the line between anthropology and that which it studies. To say this, however, is not to exhaust the recursive potential in the arguments I have presented. This move could be propelled even further if it was turned back on itself. This would not involve asking how anthropological ideas that present themselves as comparable to scientific ones

[128] Preferably, such a form would also capture the extent to which the ethnographic relation in question is governed by the extent to which scientists' data is anthropological knowledge, when ideas of knowledge and data do not map onto each other.
[129] No reference is intended here to Marilyn Strathern's separate but related notion of "merographic connections" (Strathern 1992: 72).

might fare in such a comparison. Rather, the question would be whether the very premise of this study could be refigured by what it contains. In such a move, the very means of the comparison would become the ground that needs to be refigured. "Taking seriously" would figure, not as an end point, but as a constant re-beginning (cf Stengers 2000 [1993]). My question, at the end of this thesis, is: what can taking science seriously teach us about the taking seriously of anthropologists? I can only indicate the salient direction that such a recursive move might take. The discussion around "taking science seriously" with which I started this thesis is based on anthropological conceptions of identity and difference. Anthropological discussions of endo- and exo-anthropology, of the relation between the two, and of what "taking seriously" could signify, are the means by which anthropologists explore the insides and outsides of their own discipline, and their own knowledge.

What is at stake in taking science seriously is not correctly ascertaining how near or far the scientist as Other is to the anthropologist, but rather describing the different ways of conceiving and 'doing' proximity and distance, identity and difference, strangeness and sameness encountered in the field. The Other of the Other may always be other, as Viveiros de Castro suggests; but this means little if ideas of Otherness differ across that gap.[130] This thesis has documented the ways in which data itself encapsulates, and is the result of, knowledge-practices of inclusion and exclusion, internality and externality, difference and indifference. It has described the different ways in which limits and boundaries are construed, imagined, talked about, constructed, dissolved and employed by the technicians and researchers I worked with when they work on their data. I have already cited Roy Wagner: "[A]nthropology will not come to terms with its mediative basis and its professed aims until our invention of other cultures can reproduce, at least in principle, the ways in which those cultures invent themselves" (Wagner 1981 [1975]: 30). This also applies to the ways that those we study invent their Others, and even the ways that Otherness is constituted for them. The question is one of ethnographic specificity all the way down, as it were.[131]

[130] As Strathern writes : "It is all very well for Giddens to state that "all social actors…are social theorists" (1984:335), but the phrase is an empty one if techniques of theorizing have little common ground” (1987:30).

[131] The reason that Viveiros de Castro’s work is so engaging is, I suggest, exactly because of its ethnographic loyalty to the Amerindians he studies, and their concern with negotiating relations with "enemies" - Others (e.g. see Viveiros de Castro 1998). This is what makes it appealing to wider audiences: the ease with which the recursive move can be effected.

But how can anthropology allow that specificity to move beyond itself, in order to deal with and negotiate its own sense of the universal?

References:

Alder, K. (1995) A Revolution to Measure. In *The Values of Precision* (ed.) M. Norton Wise, 39-71. Princeton: Princeton University Press

Barad, K. (2003) Posthumanist Performativity: Toward an Understanding of How Matter Comes to Matter. *Signs: Journal of Women in Culture and Society* 28 (3): 801-831

_________ (2007) *Meeting the Universe Half-Way: Quantum Physics and the Entanglement of Matter and Meaning*. Durham: Duke University Press

Barbosa, L. (2000) *The Brazilian Amazon Rainforest: Global Ecopolitics, Development and Democracy.* Maryland: University Press of America, Inc.

Barley, S.R. and Bechky, B.A. (1994) In the Backrooms of Science: The Work of Technicians in Science Labs. *Work and Occupations* 21 (1): 85-126

Bennett, J.A. (1989) A Viol of Water or a Wedge of Glass. In *The Uses of Experiment: Studies in the Natural Sciences* (eds.) David Gooding, Trevor Pinch and Simon Schaffer, 105-114. Cambridge: Cambridge University Press

Biagioli, M. Jaszi, P. and Woodmansee, M. (2011) Introduction - High and Low: IP Practices and Materialities. In *Making and Unmaking Intellectual Property: Creative Production in Legal and Cultural Perspective*, (eds.) Mario Biagioli, Peter Jaszi, Martha Woodmansee, 1-22. Chicago: University of Chicago Press

Biagioli, M and Galison, P. (2003) Introduction. In *Scientific Authorship: Credit and Intellectual Property in Science* (eds.) Mario Biagioli and Peter Galison, 1-12. London/New York: Routledge.

Biagioli, M. (1998) The Instability of Authorship: Credit and Responsibility in Contemporary Biomedicine. *The FASEB Journal* 12: 3-16

____________ (2003) Rights or Rewards? Changing Frameworks of Scientific Authorship. In *Scientific Authorship: Credit and Intellectual Property in Science*, (eds.) Mario Biagioli and Peter Galison, 253-281. New York: Routledge

____________ (2008) Documents of Documents: Scientists' Names and Scientific Claims. In *Documents: Artifacts of Modern Knowledge,* (ed.) Annelise Riles, 127-157. Ann Arbor: University of Michigan Press

Bijker, W.E. and Law, J. (1992) (eds.) *Shaping Technology/Building Society: Studies in Sociotechnical Change.* Cambridge, Massachusetts: MIT Press

Bourdieu, P. (1988) *Homo Academicus*. Trans. Peter Collier. London: Polity

Bowker, G.C. and Star, S.L. (2000) *Sorting Things Out: Classification and its Consequences.* Cambridge, Massachusetts: MIT Press

Bowker, G.C. (2008) *Memory Practices in the Sciences.* Cambridge, Massachusetts: MIT Press

Braga, S.M. and Fernandes, C.V.S. (2007) Perfórmance de Sensores de Precipitação do Tipo 'Tipping Bucket' (Báscula) — Um Alerta para a Ocorrência de Erros. *Revista Brasileira de Recursos Hídricos* 12: 197-204

Callon, M. (1999 [1986]) Some Elements of a Sociology of Translation: Domestication of the Scallops and the Fisherman of St Brieuc Bay. In *The Science Studies Reader* (ed.) Mario Biagioli. London: Routledge

Candea, M. (2007) Arbitrary Locations: In Defense of the Bounded Field-site. *Journal of the Royal Anthropological Institute* 13: 167-184.

___________ (2010a) *Corsican Fragments: difference, knowledge and fieldwork.* Indiana: Indiana University Press

__________ (2010b) "I fell in Love with Carlos the Meerkat": Engagement and Detachment in Human: Animal Relations. *American Ethnologist* 37 (2): 241-258

___________ (2011) Endo/Exo. *Common Knowledge* 17 (1): 146-150

___________ A "That's a Piece of Data!": Habituation, Interest and the Partible Meerkat. Forthcoming in *Theory Culture and Society.*

___________ B Suspending Belief: Animal Behaviour Scientists as Abstentionists. Forthcoming in *American Anthropology.*

Castells, M. (2007) Communication, Power and Counter-power in the Network Society. *International Journal of Communication* 1: 238-266

Claeys et al. (2004) Formation of Secondary Organic Aerosols Through Photo-Oxidation of Isoprene. *Science* 303: 1173-1176

Clifford, J. (1986). Introduction: Partial Truths. In *Writing Culture: The Poetics and Politics of Ethnography* (eds.) James Clifford and George E. Marcus, 1-26. Berkeley: University of California Press

Collins, H.M. (1985) *Changing Order: Replication and Induction in Scientific Practice.* Chicago: University of Chicago Press

__________ (2002) The Experimenter's Regress as Philosophical Sociology. *Stud. Hist. Phil. Sci.* 33

Crump, T. (1990) *The Anthropology of Numbers.* Cambridge: Cambridge University Press

Daston, L. and Galison, P. (1992) The Image of Objectivity. *Representations* 0 (40): 81-128

Daston, L. (1992) Objectivity and Escape from Perspective. *Social Studies of Science* 22 (4): 597-618

Dear, P. (1992) From Truth to Disinterestedness in the Seventeenth Century. *Social Studies of Science* 22 (4): 619-631

Derksen, L. (2000) Towards a Sociology of Measurement: The Meaning of Measurement Error in the Case of DNA Profiling. *Social Studies of Science* 30 (6): 803-45

Despret, V. (2004) The Body We Care For: Figures of Anthropo-zoo-genesis. *Body and Society* 10 (2-3): 111-134

Edwards, P.N., Mayernik, M.S., Batcheller, A.L., Bowker, G.C. & Borgman, C.L. (2011) Science friction: Data, Metadata and Collaboration. *Social Studies of Science*. 41 (5): 667-690

Edwards, P.N. (1999) Global Climate Science Uncertainty and Politics: Data-laden Models, Model-filtered Data. *Science as Culture* 8 (4): 437-472

___________ (2006) Meteorology as Infrastructural Globalism. *Osiris* 21: 229-250

____________ (2010) *A Vast Machine*. Cambridge, Massachusetts: MIT Press

Fortun, K. and Fortun, M. (2005) Scientific Imaginaries and Ethical Plateaus in Contemporary U.S. Toxicology. *American Anthropologist* 107 (1): 43-54

Fortun, M. (2011) Care, Creation and the Impossible Sciences of GeneXEnvironment Interactions in Asthma: Promising Genomics V. 2. Draft prepared for workshop on Knowledge/Value: Experimental Biologies and Translational Research, University of Chicago November 6-7, 2011.

Franklin, S. and Ragoné, H. (eds.) (1998) *Reproducing Reproduction: Kinship, Power and Technological Innovation.* Philadelphia: University of Pennsylvania Press

Franklin, S. 1995. Science as Culture. *Annual Review of Anthropology* 24: 163-184

Galison, P. (1988) History, Philosophy and Central Metaphor. *Science in Context* 2 (1): 197-212

_______ (1997) *Image and Logic: A Material Culture of Microphysics.* Chicago: University of Chicago Press.

_________ (1999) Trading Zone: Coordinating Action and Belief, in Biagoli, M. *The Science Studies Reader.* London: Routledge

Gerson, E.H and Star, S.L. (1986) Analyzing Due Process in the Workplace. *ACM Transactions on Office Information Systems,* 4 (3): 257-270

Goldman, M. (2009) An Afro-Brazilian Theory of the Creative Process: An Essay in Anthropological Symmetrization. *Social Analysis* 53 (2): 108-129

Gomes Alves, E. (2011) Emissão de Isopreno em Funçåo da Fenologia Foliar de *Eschweilera coriacea* (DC.) S.A. Mori sob Diferentes Condiçœes de Luz e de Temperatura em Floresta Primária da Amazônia Central. Master’s Dissertation, LBA/INPA.

Gooding, D., Pinch, T. and Schaffer, S. (eds.) (1989) *The Uses of Experiment: Studies in the Natural Sciences.* Cambridge: Cambridge University Press

Gooding, D. (1990) *Experiment and the Making of Meaning: Human Agency in Scientific Observation and Experiment*. Dordrecht: Kluwer Academic Publishing

Gregory, C. (1982) *Gifts and Commodities.* London: Academic Press

Guenther, A., Nicholas Hewitt, C., Erickson, D., Fall, R., Geron, C., Graedel, T., Harley, P., Klinger, L., Lerdau, M., McKay, W.A., Pierce, T., Scholes, B., Steinbrecher, R., Tallamraju, R., Taylor, J., Zimmerman, P. (1995) A Global Model of the Natural Volatile Organic Compound Emissions. *Journal of Geophysical Research* 100 (D5): 8873-8892

Hacking, I. (1983) *Representing and Intervening: Introductory Topics in the Philosophy of Natural Science.* Cambridge: Cambridge University Press

_________ (1991) How Should We Do a History of Statistics? In *The Foucault Effect: Studies in Governmentality* (eds.) Graham Burchell, Colin Gordon and Peter Miller. Chicago: University of Chicago Press

__________ (1992) The Self-Vindication of the Laboratory Sciences. In *Science as Practice and Culture* (ed.) Andrew Pickering, 29-64. Chicago: Chicago University Press

Hackmann, W.D. (1989) Scientific Instruments: Models of Brass and Aids to Discovery. In *The Uses of Experiment: Studies in the Natural Sciences* (eds.) David Gooding, Trevor Pinch and Simon Schaffer, 31-66. Cambridge: Cambridge University Press

Hagstrom, W. O. (1965) *The Scientific Community*. New York: Basic Books

_____________ (1982) Gift-giving as an Organizing Principle in Science. In *Science in Context: readings in the Sociology of Science,* (eds.) Barry Barnes and David Edge, 21-34. Milton Keynes: Open University Press

Hankins, T. and Silverman, R. (1995) *Instruments and the Imagination.* Princeton: Princeton University Press

Haraway, D.J. (1991) *Simians, Cyborgs and Women: The Reinvention of Nature*. London: Free Association Books

_____________ (1997) *Modest_Witness@Second_Millenium.FemaleMan©_Meets _Oncomouse™*. New York, London: Routledge

Harman, G. (2009) *Prince of Networks: Latour and Metaphysics*. Melbourne: re.press

Harré, R. (2003) The Materiality of Instruments in a Metaphysics for Experiments. In *The Philosophy of Scientific Experimentation* (ed.) Hans Radder, 19-38. Pittsburgh: University of Pittsburgh Press

Hayden, C. (2003) *When Nature Goes Public: The Making and Unmaking of Bioprospecting in Mexico.*

Henare, A., Holbraad, M. and Wastell, S. (2007) Introduction: thinking through things. In *Thinking Through Things: Theorising Artefacts Ethnographically* (eds.) A. Henare, M. Holbraad, and S. Wastell, 1-31. Oxon: Routledge

Hilgartner, S. and Brandt-Rauf, S.I. (1994) Data Access, Ownership, and Control: Toward Empirical Studies of Access Practices. *Science Communication* 15 (4): 355-372

Hilgartner, S. (2012) Selective Flows of Knowledge in Technoscientific Interaction: Information Control in Genome Research. *BJHS* 45 (2): 267-280

Hine, C. (2006) Databases as Scientific Instruments and Their Role in the Ordering of Scientific Work. *Social Studies of Science* 36 (2): 269-298

Holbraad, M. and Pedersen, M. (2009) Planet M: The Intense Abstraction of Marilyn Strathern. *Anthropological Theory* 9 (4): 371-394

Holbraad, M. (2004) Response to Bruno Latour's "Thou Shalt Not Freeze-frame". Available at http://nansi.abaetenet.net/abaetextos/response-to-bruno-latours-thou-shall-not-freeze-frame-martin-holbraad. Accessed November 2012.

___________ (2005) Expending Mulitplicity: Money in Cuban Ifá Cults. *Journal of the Royal Anthropological Institute* 11: 231-25

___________ (2012) *Truth in Motion: The Recursive Anthropology of Cuban Divination.* Chicago: Chicago University Press

Ingold, T. (2000) *The Perception of the Environment: Essay on Livelihood, Dwelling and Skill.* Routledge: London

Jensen, C.B. and Winthereik, B.R. (2013) *Monitoring Movements: Infrastructures and Partnerships in Development Aid.* Cambridge, MS/London: MIT Press.

Jensen, C.B. (2007) Infrastructural Fractals: Revisiting the Micro-Macro Distinction in Social Theory. *Environment and Planning D: Society and Space* 25 (5): 832-850

__________ (2011) Making Lists, Enlisting Scientists: The Bibliometric Indicator, Uncertainty and Emergent Agency. *Science Studies* 24 (2): 64-84

__________ (2012) Motion: The Task of Anthropology is to Invent Relations. *Critique of Anthropology* 32(1): 47-53

Jiménez, A.C. (2003) Teaching the field: the order, ordering, and scale of knowledge. *Anthropology Matters* 5 (1). www.anthropologymatters.com

____________ (2005) After Trust. *Cambridge Anthropology* 25 (2): 64-78

Kelly, J.A. (2005) Fractality and the Exchange of Perspectives. In *On the Order of Chaos: Social Anthropology and the Science of Chaos* (eds.) Mark S. Mosko and Fred H. Damon, 108-135. Oxford/New York:Berghahn Books

__________ (2012) Figure Ground Dialectics in Yanomami, Yekuana and Piaroa Myth and Shamanism. Unpublished.

Knorr Cetina, K. and Amman, K. (1990) Fixation of (Visual) Evidence. In *Representation in Scientific Practice* (eds.) Michael Lynch and Steve Woolgar, 85-118. Cambridge, Massachusetts: MIT Press

Knorr Cetina, K. (1981) *The Manufacture of Knowledge: An Essay on the Constructivist and Contextual Nature of Science.* Oxford: Pergamon Press

________________ (1982) Scientific Communities or Transepistemic Arenas of Research? A Critique of Quasi-Economic Models of Science. *Social Studies of Science* 12 (1): 101-130

Knowles Middleton, W.E. (1969) *Invention of the Meteorological Instruments.* Baltimore: John Hopkins Press

Konrad, M. (1998) Ova Donations and Symbols of Substance: Some Variations on the Theme of Sex, Gender and the Partible Body. *Journal of the Royal Anthropological Institute* 4 (4): 643-667

__________ (2005) *Nameless Relations: Anonymity, Melanesia and Reproductive Gift Exchange Between British Ova Donors and Recipients*. Oxford: Berghahn

Kumar, M. (2008) *Quantum.* London: Icon Books

Kwa, C. (1987) Representations of Nature Mediating Between Ecology and Science Policy: The Case of the International Biological Programme. *Social Studies of Science* 17 (3): 413 - 442

_________ (2005) Local Ecologies and Global Science: Discourses and Strategies of the International Geosphere-Biosphere Programme. *Social Studies of Science* 35 (6): 923-950

Lahsen, M. and Nobre, C.A. (2007) Challenges of Connecting International Science and Local Level Sustainability Efforts: the case of the Large-Scale Biosphere-Atmosphere Experiment in Amazonia. *Environmental Science and Policy* 10: 62-74

Lahsen, M. (2009) A Science-Policy Interface in the Global South: The Politics of Carbon Sinks and Science in Brazil. *Climatic Change* 97: 339-372

Laidlaw, J. (2000) A Free Gift Makes No Friends. *Journal of the Royal Anthropological Institute* 6: 617-634

Latour, B. and Woolgar, S. (1979) *Laboratory Life: The Construction of Scientific Facts.* Princeton: Princeton University Press

Latour, B. (1983) Give Me a Laboratory and I Will Raise the World. In *Science Observed: Perspectives on the Social Study of Science* (eds.) Karin D Knorr-Cetina and Michael Mulkay, 141-170. London: Sage Publications

_________(1987) *Science in Action: How to Follow Scientists and Engineers Through Society.* Cambridge, Massachusetts: Harvard University Press

________ (1988) *The Pasteurization of France.* Trans. A. Sheridan and J. Law. Cambridge, Massachusetts: Harvard University Press

_________ (1993) *We Have Never Been Modern.* Trans. Catherine Porter. Cambridge, Massachusetts: Harvard University Press.

_________ (1999a) *Pandora's Hope: Essays on the Reality of Science Studies.* Cambridge, Massachusetts: Harvard University Press

_________ (1999b) On Recalling ANT. In *Actor Network Theory and After* (eds.) John Law and John Hassard, 15-25. Oxford: Blackwell Publishing

_________ (2004a) *Politics of Nature: How to Bring the Sciences into Democracy.* Translated by Catherine Porter. Cambridge, Massachusetts: Harvard University Press

________ (2004b) Scientific Objects and Legal Objectivity. In *Law, Anthropology and the Constitution of the Social: Making Persons and Things* (eds.) Alain Pottage and Martha Mundy, 73-114. Cambridge: Cambridge University Press

__________ (2005) *Reassembling The Social.* Oxford: Oxford University Press

__________ (2009) Will Non-Humans Be Saved? An Argument in Ecotheology. The Henry Myers Lecture 2008. *Journal of the Royal Anthropological Institute* 15: 459-475

Law, J. (2003) Networks, Relations, Cyborgs: On the Social Study of Technology. Published by the Centre for Science Studies, Lancaster University, Lancaster, LA1 4YN, UK at http://www.comp.lancs.ac.uk/sociology/papers/Law-Networks-Relations-Cyborgs.pdf

______ (2004) *After Method: Mess in Social Science Research.* Abingdon, Oxon: Routledge

Leach, J. (2011) The Self of the Scientist, Material for the Artist: Emergent Distinctions in an Interdisciplinary Collaboration. *Social Analysis* 55 (3): 143-163

Lee, B. and Brown, S. (1994) Otherness and the Actor Network: The Undiscovered Continent. *American Behavioral Scientist* 37 (6): 772-790

Lenoir, T. (ed.) (1998) *Inscribing Science: Scientific Texts and the Materiality of Communication.* Stanford: Stanford University Press.

Lionelli, S. (2012) Introduction: Making Sense of Data-driven Research in the Biological Sciences. *Studies in History and Philosophy of Biological and Biomedical Sciences* 43: 1-3

Lovelock, J.E. (1972) Gaia as Seen Through the Atmosphere. *Atmospheric Environment* 6: 579-580

Lynch, M.E. (1988) Sacrifice and the Transformation of the Animal Body into a Scientific Object: Laboratory Culture and Ritual Practice in the Neurosciences. *Social Studies of Science* 18 (2): 265-289

Mahli, Y. and Grace, J. (2000) Tropical Forests and Atmospheric Carbon Dioxide. *Trends in Ecology and Evolution* 15 (8): 332-337

Makarieva, A.M. and Gorshkov, V.G. (2007) Biotic pump of Atmospheric Moisture as Driver of the Hydrological Cycle on Land. *Hydrol. Earth. Syst. Sci.* 11: 1013-1033

Mallard, A. (1998) Compare, Standardize, Settle Agreement: On Some Usual Metrological Problems. *Social Studies of Science* 28: 571-601

Manovich, L. (1999) Database as Symbolic Form. *Convergence* 5 (2): 80-99

Marcus, G.E. (1995) Ethnography in/of the World System: The Emergence of Multi-Sited Ethnography. *Annual Review of Anthropology* 24: 95-117

Maurer, B. (2005) *Mutual Life, Limited: Islamic Banking, Alternative Currencies, Lateral Reason*. Princeton: Princeton University Press

McSherry, C. (2003a) Uncommon Controversies: Legal Mediations of Gift and Market Models of Authorship. In *Scientific Authorship: Credit and Intellectual Property in Science*, (eds.) Mario Biagioli and Peter Galison, 226-250. New York: Routledge

___________ (2003b) *Who Owns Academic Work? Battling for Control of Intellectual Property*. Harvard: Harvard University Press

Meillassoux, Q. (2009 [2008]) *After Finitude: An Essay on the Necessity of Contingency*. London: Continuum

Miller, C.A. and Edwards, P.N. (eds.) (2001) *Changing the Atmosphere: Expert Knowledge and Environmental Governance.* Cambridge, Massachusetts: MIT Press

Mimica, J. (1988) *Intimations of Infinity: The Cultural Meanings of the Iqwaye Counting and Number Systems.* London: Berg

Mody, C.C.M. (2001) A Little Dirt Never Hurt Anyone: Knowledge-Making and Contamination in Materials Science. *Social Studies of Science* 31 (1): 7-36

Mol, A. and Law, J. (1994) Regions, Networks and Fluids: Anaemia and Social Topology. *Social Studies of Science* 24 (4) 641-671

Murray, F. (2011) Patenting Life: The Oncomouse Patent. In *Making and Unmaking Intellectual Property: Creative Production in Legal and Cultural Perspective,* (eds.) Mario Biagioli, Peter Jaszi, Martha Woodmansee, 339-412. Chicago: University of Chicago Press

Nelson, B. and Barley, S.R. (1992). *Practice Makes Perfect: Emergency Medical Technicians and the Social Negotiation of a Skilled Occupational Identity.* Working Paper, University of Pennsylvania, Centre for the Education of the Workforce, Philadelphia.

Nobre, A. D. (2004) Unraveling the Mysteries of Carbon in Amazonia: LBA Moves Forward, but Stumbles on Complexities of Ecosystems. *Folha Amazônica* 6 (12). LBA Project Publication.

_________ (2010) Floresta e Clima: Saber Indígena e Ciência. In *Manejo Do Mundo: Conhecimentos e Práticas dos Povos Indígenas do Rio Negro, Noroeste Amazônico,* 38-45. Organized by Aloísio Cabalzar. FOIRN/ISA Publication.

North, J.D. (1989) *The Universal Frame: Historical Essays in Astronomy, Natural Philosophy and Scientific Method.* London: The Hambledon Press

Norton, S.D. and Suppe, F. (2001) Why Atmospheric Modeling is Good Science. In *Changing the Atmosphere: Expert Knowledge and Environmental Governance.* Cambridge, Massachusetts: MIT Press

Norton Wise, M. (1995) Introduction. In *The Values of Precision* (ed.) Matthew Norton Wise, 3-13. Princeton: Princeton University Press

O'Connell, J. (1993) Metrology: The Creation of Universality by the Circulation of Particulars. *Social Studies of Science* 29: 129-173

Pentland, B. (1991) *Making the Right Moves: Towards a Social Grammar of Software Support Hotlines.* PhD Dissertation, MIT, Philadelphia.

Pedersen, M. (2012) Motion: The Task of Anthropology is to Invent Relations. *Critique of Anthropology* 32(1): 59-65

Pickering, A. (1981) The Hunting of the Quark. *Isis* 72 (2): 216-236

Pickering, A. (ed.) (1992) *Science as Practice and Culture.* Chicago: University of Chicago Press

Porter, T.M. (1995) *Trust in Numbers: The Pursuit of Objectivity in Science and Public Life.* Princeton: Princeton University Press

Rheinberger, H-J. (1995) From Microsomes to Ribosomes: "Strategies" of "Representation". *Journal of the History of Biology* 28 (1): 49-89

______________ (1997) *Toward a History of Epistemic Things: Synthesizing Proteins in the Test Tube.* Stanford: Stanford University Press

______________ (2010) *An Epistemology of the Concrete: Twentieth-Century Histories of Life.* Durham: Duke University Press

Riles, A. (2001) *The Network Inside Out*. Ann Arbor: The University of Michigan Press

_______ (2010) Collateral Expertise: Legal Knowledge in the Global Financial Markets. *Current Anthropology* 51 (6): 795-818

______ (2011) *Collateral Knowledge: Legal Reasoning in the Global Financial Markets*. Chicago: Chicago University Press

Scarselletta, M. (1992) *Button Pushers and Ribbon Cutters: Observations on Skill and Practice in a Hospital Laboratory and Their Implications for the Shortage of Skilled Technicians.* Working Paper, University of Pennsylvania, Centre for the Education of the Workforce, Philadelphia.

Schaffer, S. (1989) Glass Works: Newton's Prisms and the Uses of Experiment. In *The Uses of Experiment: Studies in the Natural Sciences* (eds.) David Gooding, Trevor Pinch and Simon Schaffer, 67-104. Cambridge: Cambridge University Press

__________ (1992) Late Victorian Metrology and Its Instrumentation: A Manufactory of Ohms. In *Invisible Connections: Instruments, Institutions, and Science* (eds.) R. Bud and S.E. Cozzens, 23 - 56. Bellingham: SPIE.

__________ (2000) Modernity and Metrology. In *Science and Power: The Historical Foundation of Research Policies in Europe*, (ed.) Luca Guzzetti. 71-92. Luxembourg: EU, 2000

Schickore, J. (2001) Ever-Present Impediments: Exploring Instruments and Methods of Microscopy. *Perspectives on Science* 9 (2): 126-146

Schor, T. and Moraes, A. (2011) Programas de Pesquisa em Meio Ambiente e o Urbano: Um Ensaio Sobre a Ausência. *Revista GEONORTE* 2 (3): 1-24

Schrempp. G. (1992) *Mythical Arrows: The Maori, the Greeks and the Folklore of the Universe.* London: The University of Wisconsin Press

Shapin, S. (1989) The Invisible Technician. *American Scientist* 77(6): 554-563

Smith, B.H. (2006) *Scandalous Knowledge: Science, Truth and the Human.* Durham: Duke University Press

Star, S.L. (1983) Simplification in Scientific Work: An Example from Neuroscience Research. *Social Studies of Science* 13 (2): 205-228

________ (1991) Power, Technology and the Phenomenology of Conventions: On Being Allergic to Onions. In *A Sociology of Monsters: Essays in Power, technology and Domination,* (ed.) John Law, 26-56. London/New York: Routledge.

Stengers, I. (2000 [1993]) *The Invention of Modern Science.* Trans. Daniel W. Smith. Minneapolis: University of Minnesota Press

________ (2005) The Cosmopolitical Proposal. In *Making Things Public: Atmospheres of Democracy,* (eds.) Bruno Latour and Peter Weibel, 994-1003. Cambridge, Massachusetts: MIT Press

________ (2010 [2003]) *Cosmopolitics I.* Trans. Robert Bononno. Minneapolis: University of Minnesota Press

_________ (2011) Comparison as a Matter of Concern. *Common Knowledge* 17 (1): 48-76

Strathern, M. (1987) Out of Context: The Persuasive Fictions of Anthropology [and Comments and Reply]. *Current Anthropology* 26 (3): 251-281

__________ (1988) *The Gender of the Gift: Problems with Women and Problems with Society in Melanesia.* Berkeley: University of California Press

___________ (1991) *Partial Connections*. Maryland: Rowman and Littlefield Publishers, Inc.

__________ (1992a) *Reproducing the Future: Anthropology, Kinship, and the New Reproductive Technologies.* Manchester: Manchester University Press

__________ (1992b) Parts and Wholes: Refiguring Relationships in a Post-plural World. In *Conceptualizing Society*, (ed.) Adam Kuper, 75-106. London: Routledge

__________ (1996) Potential Property: Intellectual Rights and Property in Persons. *Social Anthropology* 4:17-32

_________ (2002) Not Giving the Game Away. In *Anthropology, By Comparison* (eds.) Andre Gingrich and Richard G. Fox, xiii-xvii. London: Routledge

__________ (2003) Emergent Relations. In *Scientific Authorship: Credit and Intellectual Property in Science* (eds.) Mario Biagioli and Peter Galison, 165-194. London/New York: Routledge.

Suchman, L. (2009) Agencies in Technology Design: Feminist Reconfigurations. Keynote, 5th European Conference on Gender and ICT, Bremen, Germany.

Tansey, E.M. (2008) Keeping the Culture Alive: The Laboratory Technician in Mid-twentieth-century British Medical Research. *Notes. Rec. R. Soc.* 62: 77-95

Taylor, J. S. (1998) Image of Contradiction: Obstetrical Ultrasound in American Culture. In *Reproducing Reproduction: Kinship, Power and Technological Innovation* (eds.) Sarah Franklin and Helena Ragoné. Philadelphia: University of Pennsylvania Press

Thompson, C. (2005) *Making Parents: The Ontological Choreography of Reproductive Technologies.* Cambridge, Massachusetts: MIT Press

Tsing, A. L. (2005) *Friction: An Ethnography of Global Connection.* Princeton: Princeton University Press

Venkatesan, S. (2011) The Social Life of a "Free" Gift. *American Ethnologist* 38 (1): 47 -57

Vilaça, A. (2000) Relations Between Funerary Cannibalism and Warfare Cannibalism: The Question of Predation. *Ethnos* 65 (1): 83-106

________ (2002) Making Kin Out of Others in Amazonia. *Journal of the Royal Anthropological Institute* 8: 347-365

Viveiros de Castro, E. (1998) Cosmological Deixis and Amerindian Perspectivism. *Journal of the Royal Anthropological Institute* 4 (3): 469-488

________________ (2001) GUT Feelings About Amazonia: Potential Affinity and the Construction of Sociality. In *Beyond the Visible and the Material: The Amerindianization of Society in the Work of Peter Rivière* (eds.) Laura M. Rival and Neil L. Whitehead. New York: Oxford University Press.

________________ (2002) O Nativo Relativo. *Mana* 8 (2):113–148

_________________ (2004) Perspectival Anthropology and the Method of Controlled Equivocation. *Tipití: Journal of the Society for the Anthropology of Lowland South America* 2 (1): 3-22

________________ (2007) The Crystal Forest: Notes on the Ontology of Amazonian Spirits. *Inner Asia* 9 (2): 153-172

_________________ (2011a) Zeno and the Art of Anthropology: Of Lies, Beliefs, Paradoxes and Other Truths. *Common Knowledge* 17 (1): 128-145

_________________ (2011b) Zenos Wake. Trans. Ashley Lebner. *Common Knowledge* 17 (1): 163-165

Wagner, R. (1977). Scientific and Indigenous Papuan Conceptualisations of the Innate: A Semiotic Critique of the Ecological Perspective. In *Subsistence and Survival: Rural Ecology in the Pacific*, (eds.) T. Bayless-Smith and R. Feachem. London: Academic Press

_________ (1981 [1975]) *The Invention of Culture.* Chicago: University of Chicago Press

_________ (1986) *Symbols That Stand For Themselves*. Chicago: University of Chicago Press

_________ (1991) The Fractal Person. In *Big Men and Great Men: Personifications of Power in Melanesia* (eds.) Marilyn Strathern and Maurice Godelier, 159-173. Cambridge: Cambridge University Press.

_________ (2010) *Coyote Anthropology.* Lincoln: University of Nebraska Press.

Zimmerman, A.S. (2003) *Data Sharing and Secondary Use of Scientific Data: Experiences of Ecologists*. PhD dissertation, University of Michigan.

_____________ (2008) New Knowledge from Old Data: The Role of Standards in the Sharing and Reuse of Ecological Data. *Science, Technology and Human Values* 33 (5): 631-652