

Material Origins of a Data Journey in Ocean Science: How Sampling and Scaffolding Shape Data Practices



Gregor Halfmann

Abstract This chapter discusses the epistemological relevance of material interactions during the early stages of a data journey. It shows that processes taking place when research technology makes physical contact with the objects targeted in research endeavours shape the subsequent data journeys and the practices of creating scientific knowledge. The chapter is based on a case study of ecological monitoring in ocean sciences and zooms in on the practice of sampling the oceans' ecosystems with mechanical sampling devices that are towed regularly by commercial ships. I propose an understanding of materiality as the integration of physical matter from various sources so as to constitute a new entity, in this case a research sample containing plankton organisms. The material integration is followed by material continuity, the preservation of the sample throughout several if not all stages of the research process without a change of medium. This two-fold understanding is an attempt to ground the notion of "materiality" epistemologically rather than ontologically. As shown with empirical examples, material interactions are the origin of resistances or challenges which unfold throughout the research process as scientists intend to create knowledge by manipulating and analysing physical objects. The scientific practices are shaped by investigating, resolving, and working around these challenges.

1 Introduction

This chapter tracks physical interactions during the creation of research samples and discusses their epistemological significance. On the basis of a case study in ocean science, I argue that interactions between materials of the research technology and of the natural systems studied by scientists shape practices of creating and using scientific data; scientists deliberately study material interactions in order to account for uncertainties and to maintain commensurability of data that have been created decades apart. Understanding the epistemological significance of "materiality" in scientific practices is thus important for studies of data journeys.

G. Halfmann (✉)

Department of Sociology, Philosophy and Anthropology, University of Exeter, Exeter, UK

© The Author(s) 2020

S. Leonelli, N. Tempini (eds.), *Data Journeys in the Sciences*,
https://doi.org/10.1007/978-3-030-37177-7_2

A variety of studies in the philosophy and sociology of science are concerned with the material nature of scientific objects and practices. Inducing a ‘clash of materials’ (Rheinberger 2011: 344) between biological entities and research technologies is central to many experimental practices in the life sciences. Such a clash may lead to the formation of objects, which are described as “material”. However, a wide range of objects with fundamentally different formation processes and physical characteristics are used in the life sciences, for example anatomical preparations (Rheinberger 2015), model organisms (Ankeny and Leonelli 2011), or “material models” in the form of species collections in museums (Griesemer 1990). What “materiality” implies for knowledge production has been elaborated by scholars in some cases, showing that material interactions and knowledge production processes are often intertwined, but in a variety of ways.¹ This chapter complements these accounts by tracking the epistemological impacts of material interactions at selected stages of the formation and processing of research samples.

While many scholars have focussed on specific kinds of material objects or material aspects of their case studies, the terms “material” and “materiality” tend to remain rather loosely defined. Quite often, it seems that “material” is used to signal difference or opposition to other classes of objects or processes, which may be labelled “non-material”, “virtual”, “theoretical”, “mathematical”, “ideational”, or the like. The opposition seems to bear on differences in an entity’s physical constitution, stability, or tangibility, but also relates to its ontological status: mathematical theories or ideas certainly differ ontologically from a sampled biological species.

Debates over the meaning of “materiality” have ensued in some cases; for example, Morgan (2003) and Parker (2009) debate how to understand “materiality” with respect to scientific experiments. Parker (2009: 492–93) criticises that computer simulations are not seen as material experiments by many; she further suggests that the emphasis on “stuff” may be misplaced and that epistemologically, the behaviour of a system is more relevant than its ontological characteristics. In science and technology studies, the meaning of materiality has been discussed in relation to a growing interest in ontology; Woolgar and Lezaun (2013: 326) argue that characteristics that may qualify an object as “material” should be treated as practical achievements and “materiality” should therefore be understood as an ‘upshot of practices’ of a certain kind. These examples show that materiality in scientific practices deserves deeper scholarly consideration; a closer study of materiality may provide classifications involving “material”, “non-material”, or other types of entities with crucial context and a more solid grounding.

In this chapter, I propose an understanding of materiality as the integration of physical matter from various sources so as to constitute a new entity; the material integration is followed by the preservation of the entity throughout several if not all

¹For example, the materiality of anatomical preparations results in an ‘indexicality’ of the object that points to itself rather than representing something else (Rheinberger 2015: 323); standardised material characteristics of model organisms make them usable as ‘genetic tools’ (Ankeny and Leonelli 2011: 316); the materiality of species collections provide an epistemological robustness to potential changes of theoretical perspective (Griesemer 1990: 83).

stages of the epistemic process without a change of medium. Material integration and material continuity are a two-fold characteristic applicable to objects that scientists create and use as well as to scientific practices. This understanding is an attempt to ground the notion of “materiality” epistemologically rather than ontologically. Empirical examples presented in this chapter show that material interactions are the origin of resistances or challenges, which unfold throughout the research process as scientists intend to create knowledge by manipulating and analysing physical objects; scientific practices are shaped by investigating, resolving, and working around these challenges.

Material integration appears as a characteristic that is applicable to virtually any entity considering formational processes on a biological or chemical level. However, in combination with material continuity, my understanding of materiality leaves aside data practices that involve “jumps” to an entirely different medium. Understanding materiality in research processes requires a focus on preservation and overlaps between different stages of epistemic practices, which can be likened to scholarly accounts of biological reproduction as I discuss later in this chapter.

By focusing on sampling and subsequent research practices, my chapter zooms in—as the title indicates—on the “origins” or the very early stages of a data journey. The beginning of a journey is not necessarily the moment, in which things move physically (or virtually) for the first time. Many would argue that a personal journey begins with thorough planning and smart packing; many choices made at this stage—which route to travel, which shoes to wear—depend on material aspects and conditions such as terrain or expected weather conditions. These conditions create challenges, which shape the actual movement and influence the journey’s outcome. The journey, as an unfolding process or development, is enabled, facilitated, or “scaffolded” by these choices and by the artefacts, infrastructures, and agents a traveller has decided to utilise, for example boots, maps, or travel agents. This chapter is not about the data journey per se, but about early stages of an epistemic process; I use the plural form “origins” to account for the difficulty of pointing at one distinct moment, at which the journey begins.² A great deal of thinking, planning, and preparing is necessary for data (and for persons) to travel successfully (Leonelli 2016: 40, [Learning from Data Journeys](#)); the origins of these preparations, that is the processes and conditions that cause or provoke certain preparatory measures, are scattered across various domains,³ including, as I intend to show in this chapter, material interactions at the sampling stage.

²I use the term “origin” with caution, in particular in relation to material objects; Rheinberger (2011: 338–9) writes that with respect to “traces”, which are ‘material manifestations’ in experimentation before they are turned into representations, an origin does not exist and has never existed. With “origins of a data journey”, I intend to highlight a number of processes leading up to the creation of data and the data journey, without implying that a concrete origin in space and time is tangible.

³The institutional context of research or the history of a research field, from which research activity is inspired and research methods are passed on, are examples of other domains that introduce restrictions on data practices.

Various stages of processing and manipulation of physical objects are strongly pronounced in my case study, the Continuous Plankton Recorder (CPR) Survey. Research samples containing marine organisms are created by deployment of mechanical sampling devices on commercial ships crossing the oceans. Samples are then analysed in four distinct steps in a laboratory in Plymouth, UK; these include microscopic identification and counting of hundreds of different taxa ranging from single-celled phytoplankton to zooplankton organisms measuring several millimetres. All samples are archived for potential re-analysis in the future. I illustrate with several examples in this chapter that material interactions between the mechanical sampling device and marine organisms require specific practices, which “scaffold” the creation and the interpretation of scientific data.

I understand scaffolding as dynamic structures of conceptualisations, practices, theories, technologies, or personal relationships, which are applied to entities in order to facilitate the development of specific capacities or skills. Wimsatt and Griesemer (2007) have coined the concept of scaffolding in relation to the development of culture and it has since been applied to various domains, including scientific practice. A rich collection of essays (Caporael et al. 2014b) demonstrates the applicability in three very broad domains—evolution, culture, and cognition—and encourages scholars to analyse their own work in terms of scaffolding. An example of its applicability in science is Wylie (2016: 1), who explains how ‘interpretive scaffolding’ is used in archaeology to determine how material traces of the past can be used as evidence; Wylie points to epistemological consequences of scaffolding by arguing that scaffolding is always provisional and new ways of data interpretation are capable of calling assumptions based on an established scaffold into question.

As the following empirical sections show, the CPR Survey is grounded in the analysis of physical objects probably as much as archaeology; yet, my case is quite different, because the same type of evidence—physical samples containing marine organisms—is created repeatedly over multiple decades. Scientists must then be able to compare old data with new data, which is a common challenge in environmental sciences that study long-term changes of natural systems (Edwards 2010). Besides discussing the material origins of scientific data, this chapter illustrates how the usability of data from different decades is scaffolded by implementing data practices that preserve methodological continuity.

After introducing the CPR Survey and tracking some material interactions and their epistemological implications, I discuss my understanding of “materiality” and elaborate how the material origins of scientific data require different forms of scaffolding and thereby shape data practices.

2 The Continuous Plankton Recorder Survey

The CPR Survey is an ongoing, long-term research programme run by the Sir Alister Hardy Foundation for Ocean Science (SAHFOS) from Plymouth, UK, since 1990 until SAHFOS merged with the Marine Biological Association of the UK (MBA) in

2018.⁴ However, the survey itself is much older. The CPR was invented by fisheries ecologist Alister Hardy in the 1920s for the purpose of monitoring zooplankton, the key food source of larval fish (McQuatters-Gollop et al. 2015: 2). The design of the CPR and the steps of sample analysis were developed experimentally in a ‘pioneer period’ in the 1930s and the early 1940s (Reid et al. 2003: 130). Since the 1950s, the basic methods of sampling and analysis have remained unchanged (Reid et al. 2003: 131–32). With datasets covering more than 70 years, the CPR Survey is one of the longest running time series in environmental and marine science (McQuatters-Gollop et al. 2015: 2). The methodological stability is one of the most important aspects of the CPR Survey; it is vital for its reputation and prestige in the scientific community, but it introduces constraints to scientific practice, as the survey’s lab manager David Johns explains:

“The whole idea is that you keep the methodology the same. You don’t want to make any mistakes with methodology, it has got to be the same. We pride ourselves on our 70-year time series, that’s what we want.” (DR0934: 5)

The CPR Survey has a long and eventful history; it was close to shut down in the 1980s, when long-term marine monitoring programmes in Europe were terminated at an alarming rate (Duarte et al. 1992).⁵ Unlike many other programmes, the projected closing of the CPR Survey led to an international initiative strongly supported by the International Council for the Exploration of the Sea (ICES) and the Intergovernmental Oceanographic Commission (IOC) of UNESCO; a rescue fund was put together and established SAHFOS as a charity organisation in 1990 (Reid et al. 2003). SAHFOS’ core work was ‘the running and safeguarding’ of the CPR Survey, according to its former Director (Owens 2015: 2). Running the survey consists of producing data related to plankton distributions from the analysis of samples, which are created through the deployment of CPRs. In addition to this core activity, SAHFOS increasingly engaged in ‘ancillary activities and associated science’ (Owens 2015: 2).⁶

A CPR is a mechanical filtering device that is towed by commercial ships on their regular shipping routes. Bands of silk inside the CPR filter the seawater and are processed into individual samples measuring around ten by ten centimetres. As of summer 2017 more than 5 million nautical miles have been sampled with CPRs in total and more than 250,000 samples have been analysed.⁷ The CPR Survey oper-

⁴At the time of my research, the CPR Survey was still conducted by SAHFOS, and the name therefore appears throughout the chapter and my references. Since April 2018 the CPR Survey is officially run by the MBA and the name “SAHFOS” has now largely disappeared from websites and official statements related to the CPR Survey.

⁵At that time, long-term ecological monitoring ‘was considered weak science, akin to stamp collecting’ (Reid et al. 2003: 141); around 40% of European monitoring programmes were shut down in the late 1980s (Duarte et al. 1992).

⁶The survey’s staff members are involved in the development and testing of new technology, in policy-driven work, or in education and outreach. Several research fellows conduct research in environmental change, molecular ecology, marine biodiversity, sustainable resources, and health and well-being of marine food sources (SAHFOS 2015).

⁷<<https://www.sahfos.ac.uk/>> [accessed 26 June 2017].

ates mainly in the North Atlantic and the North Sea, where most of the circa 25 regular towing routes are located. All samples are archived and stored in Plymouth for potential re-analysis in the future. Research based on CPR data sets has contributed significantly to the understanding of spatio-temporal dynamics of oceanic plankton and their responses to anthropogenic pressures and climate variability. The data are also used to inform UK and European marine policy-making and management of the seas (McQuatters-Gollop et al. 2015: 2).

In today's ocean science landscape, the CPR Survey is one of the oldest, yet only one of many projects that engage people without scientific credentials or institutional affiliations in sampling or data creation. To meet the economic challenges of sampling the world's oceans on increasingly finer spatial scales and with temporal regularity, a growing number of projects take advantage of recreational and professional seafarers, who regularly interact with the oceans. Picking up the current wave of citizen science and fuelled by technological innovation, marine science is often seen as a prime example of scientific fields with high potential for contributions by citizen scientists (Lauro et al. 2014). The CPR Survey does not refer to itself officially as "citizen science", although a wide range of non-scientists volunteer to make the survey possible. Among them are the captains, chief officers, bosuns, and crews aboard ships, but also ship owners and managers, stevedores, terminal managers, heavy cargo operators, and engineering companies (DR1960: 6).⁸ The collaboration is crucial for setting up a ship for towing CPRs and for proper handling and transportation of boxed CPRs in high security areas in the ports' container terminals. For each ship and each tow, the survey relies on a number of volunteers, who make sure that a CPR arrives at the right ship at the right time. The collaborative practice of the CPR Survey has epistemic implications in its own right; most importantly, the geographical scope of the sampling and the CPR data depends on the locations of frequented shipping routes. The social dimension of the CPR Survey, in which research culture meets seafaring culture, offers opportunities for sociological and epistemological research. This chapter, however, focuses on the epistemology of the CPR Survey's material dimension.

3 Material Interactions and their Epistemological Implications

The following sub-sections describe two examples of material processes within the CPR Survey and their epistemological implications. These implications become manifest in data practices such as methods of creating data by sample analysis, but also in the outcomes of those practices, for example in databases and publications.

⁸SAHFOS often used the term "volunteers" to refer to the non-scientists involved in the survey. There is no formal contract with the non-scientists, except for the engineering companies who are commissioned to install davits or blocks on the ships that enable towing of a CPR. The shipping crews, but not the companies or ship owners, are compensated with £60 per tow (DR1960: 6).

The two processes are the deformation of plankton organisms during sampling and interactions between the silk and the organisms.

3.1 *Deformation of Plankton Organisms and Identification*

A CPR is a steel device that is shaped similar to a bobsleigh and weighs around 90 kg. When a CPR tow is scheduled to begin, crew members use the ship's winch to lower the CPR into the sea. SAHFOS emphasises that the sampling is never to interfere with the ship's normal business; a ship thus never stops or slows down for the deployment of a CPR. The steel body hits the water at up to 20 knots, putting significant tension on the steel wire, the body, and the internal mechanics of the CPR.⁹ The wire is paid out until a coloured mark settles on the sea surface, indicating that the CPR has reached the desired depth between seven and 10 m. The pointed nose of the CPR has a small opening of around 1.5 cm², through which seawater enters a tunnel inside the CPR that leads to the filtering silk (SAHFOS 2016: 18; Reid et al. 2003: 126). The tunnel widens, so that the water pressure and the speed of flow inside the tunnel reduce significantly. A layer of silk (the filtering silk) spans across the tunnel, acts as a filter and retains a share of organisms and materials that have entered the tunnel. While the CPR is being towed, a propeller attached to the external body drives a mechanism that pulls the silk continuously across the flow of water. The silk that has crossed the tunnel is met by a second layer of silk (the covering silk), which is drawn by the same mechanism. The covering silk goes on top of the filtering silk, so that the organisms are held between the two layers.¹⁰ The silk rolls are drawn together into a closed chamber filled with a formalin solution. The organisms cannot survive this process, but the formalin prevents the decay of their bodies.

Plankton organisms often get damaged and deformed during the sampling process. They may knock against the steel walls of the CPR or against other organisms that are already on the silk.¹¹ If towed through a plankton bloom, areas of the silk can actually get clogged with organisms, which affects the volume of filtered sea water (Hunt and Hosie 2006). The biggest cause of deformation is, however, the sandwiching of organisms between the two silk layers. With regard to some of the larger zooplankton species,¹² the survey's lab manager David Johns explains that "the stuff

⁹In a video of a CPR deployment, the device jumps on the sea surface for several seconds before submerging. When the CPR is hauled in, it sometimes smashes against the ship's hull strong enough for the steel body to be damaged and require refitting in the survey's workshop (DR1960).

¹⁰Two bands of silk are marked, cut, folded, rolled up, and placed inside the internal cassette by hand before a CPR is deployed. A metre of silk covers around one hundred nautical miles, so up two five metres of silk are rolled up for each of the two silk rolls.

¹¹This cause of deformation is alleviated to some degree by the widening of the tunnel and the reduction of flow speed inside the CPR by a factor of around 1/30 (Batten et al. 2003: 196; DR2901: 2).

¹²Only zooplankton species larger than two millimetres are identified and counted the way described here. Smaller plankton, including single-celled phytoplankton, are identified with up to 625x magnification; Richardson et al. (2006: 35).

is squashed” and “it is very, very flat”, when it arrives in Plymouth (DR0934: 19). The organisms thus look very different under a microscope in the survey’s lab than out in the ocean or in taxonomic reference literature; Johns explains how the altered appearance by deformation affects the identification process:

“Textbooks are obviously really useful, but it is not the same as looking down and actually seeing a physical specimen there. ... They do look quite different, so you need to manipulate the organism.” (DR0934: 18–19).

In order to be manipulated, turned around, and viewed from different angles, the zooplankton organisms are manually scraped off the silk and placed into a Bogorov tray under a different microscope for identification and counting.¹³ Johns explains:

“It is just so much easier to identify them. You can’t do it on the silk very easily. It is so much easier, you take them off, put them into that tray, add some fluid and then you can manipulate them easily, flip them around. Because a lot of them, depending on how they are lying, they can hide their identification features, so you need to kind of manipulate them 360.” (DR0533: 6)

The deformation during sampling and the way some organisms—especially those with spiny body features—are caught up in the silk requires manipulation and removal of organisms in order to create data. In this stage of the analysis, which is called the “zooplankton eyecount”, all organisms larger than two millimetres are taken off the silk for identification and counting and are put back onto the silk afterwards. Data are created by counting different species or taxonomic groups and recording the result with tally marks in a hand-written notebook right next to the microscope.¹⁴

The organisms’ altered appearance also requires the sample analysts to have specific identification skills and experience. New analysts go through a training phase, which lasts several months until they are allowed to work on samples even from the survey’s most frequent sampling routes all by themselves. Samples from areas that are not sampled as frequently as the North Atlantic and the North Sea can be particularly challenging, because the encountered species and the ecology are very different. Some analysts have therefore specialised in samples from certain areas after years of practice and interacting with other analysts in the lab (DR0533: 10). Johns explains that “probably most of [the training] is informal and on-the-job stuff” (DR0934: 18), due to the specific characteristics of the CPR samples; the skills and experience are best acquired in practice and in cooperation with experienced analysts. One of the experienced sample analysts describes the interaction in the lab, by which they gain expertise:

¹³ Manipulation and turning around of organisms is also necessary, because some species are difficult to distinguish; for example *calanus finmarchicus* and *calanus helgolandicus*, two of the most important zooplankton species in the North Atlantic and the North Sea, look very similar and are identified primarily by the shape of their fifth pair of swimming legs; Richardson et al. (2006: 47).

¹⁴ The data in the notebooks are later entered into the digital database manually by two sample analysts together in order to avoid transcription mistakes and to notice unusual looking results that might indicate an error in identification or counting (DR0533: 2).

“We are always looking at each other’s samples all the time. It’s not that a day goes past where you are not going to go a look at someone else’s stuff ...” (DR8112: 10)

The removal of materials from the sample and the ways of acquiring expertise and experience are examples of how data practices are shaped by material interactions at the sampling stage.

In 2011, SAHFOS published the *Fish larvae atlas of the NE Atlantic* (Edwards et al. 2011), which illustrates how deformation during sampling constrains exactly what kinds of data can possibly be created during sample analysis. The atlas covers geographical distributions of fish larvae of nine different taxa for the years 1948–2005. More than 10,000 archived silk samples have been re-analysed with new molecular methods, because fish larvae are not routinely identified in the microscopic analysis:

Due to the size of the fish larvae and the sampling method, they can often be damaged and identification to species level is not always possible using traditional microscopic methods. (Edwards et al. 2011: 2)

As the fish larvae are often too damaged for visual identification, they are only counted and recorded in the survey’s database as one taxonomic group. The database’s content and the knowledge of the ocean ecosystem are thus shaped by material interactions that occur during sampling.

3.2 *Silk Specifications and Quantification*

Albeit having changed silk suppliers several times throughout the history of the survey, silk with identical specifications has been used for sampling since the beginning of the CPR Survey. The silk bands have a mesh size of around 270 μm and are quality controlled and prepared in a standardised way, which includes marking, stamping, folding, cutting, and putting the silk onto a roll that is going to be placed inside the CPR.¹⁵ Smooth fabrics such as nylon and much finer mesh sizes are typically used in plankton science today. The 270 μm is indeed large compared to the size of some species that are routinely recorded, as lab manager Johns explains:

“We had people saying that there is no way that we can see coccolithophores, they said ‘no, it is going to go straight through your mesh, because they are only ten microns.’ But they do stay there, so we took photos and we published some of it and say ‘actually, we can see these.’” (DR0934: 6)

Coccolithophores are a group of unicellular, eukaryotic phytoplankton species, which are around a magnitude smaller than the gap between the silk threads; yet, a constant portion of those species are retained. That is because the silk has a certain

¹⁵Marking and stamping is required for calculating the cutting points after each tow under consideration of the ship’s average speed; each sample is intended to correspond to ten nautical miles of a tow, but the length of silk pulled by the mechanisms over that distance depends on how fast the ship has sailed.

roughness and the individual threads are spinous, so that small organisms stick to them; the silk also has a leno weave, which has two twisted threads going in one direction and one thread in the other direction, whereas most nylon fabrics used for filtering are heat-fused so that the junctions are smooth. Phytoplankton can thus get caught in the tiny gaps between the twisted silk threads (Richardson et al. 2006: 61; DR0934: 6).

Some interactions between certain types of organisms and the sampling technology are in fact multi-layered, because the presence of larger organisms also affects the efficiency, at which small phytoplankton are retained.¹⁶ Large zooplankton may have spiny body features, on which smaller organisms may get caught. As a growing amount of plankton covers the silk, the filter efficiency tends to increase:

As more and more organisms are filtered onto the mesh the open apertures are progressively clogged and reduce the effective mesh size. So as more large organisms are retained, smaller organisms, which at the start of the sampling would have been extruded, will be retained progressively more effectively (Batten et al. 2003: 206).

In general, a significant amount of small phytoplankton still flow through the silk and return into the open ocean, while most of the large zooplankton is retained. The material processes are complex and have led to experimental investigations regarding the effects of clogging with different mesh sizes (Hays 1994; Hunt and Hosie 2006). Some gelatinous plankton species can particularly enhance clogging (Richardson et al. 2006: 61). Batten et al. (2003: 206) explain the challenge posed by such interactions between organisms of different sizes and texture and the silk:

The effect is hard to quantify since the ambient concentrations of organisms (needed to determine the true proportion retained) will never be known for a specific patch of sea water at a specific time.

The materiality of the silk and the plankton organisms thus have implications that relate to the quantities of specific organisms on the silk, which are represented in the data created by the analysts. More specific, the data created by sample analysis hardly reflect the total numbers of plankton organisms at a specific space and time in the ocean. Richardson et al. (2006: 61) state that ‘there is increasing evidence that the CPR substantially underestimates absolute numbers’. The CPR data are thus often referred to as “semi-quantitative”. This characteristic of the CPR Survey, which is a result of material processes, does not mean that data are false or useless; however, the materiality shapes the way data are used by scientists:

Notwithstanding the semi-quantitative nature of CPR sampling, there is considerable evidence that it captures a roughly consistent fraction of the in situ abundance of each taxon and thus reflects the major patterns observed in the plankton. (Richardson et al. 2006: 61)

The semi-quantitative character of the data could be viewed as a shortcoming; however, as Johns explains, the consistency of the sampling is valued higher than potential increases of precision:

¹⁶The distinction I make between small phytoplankton and large zooplankton is a simplification and does not reflect the spectrum of shapes and sizes of the organisms on a silk sample.

“We want to keep that consistent time series. And there are a lot of potential sort of foibles in the dataset. But the fact that it has always been done in the same way ... You get lots of people who, it’s not an accuse, but who would say ‘well, you under-count certain things’. Well yeah, we do, but they have been consistently under-counted for sixty years. So you can just ignore the abundance values and just look at the trend to see what is happening. So yeah, if you were starting [the survey] from scratch, you would do it completely differently.” (DR0533: 4)

Other “foibles”¹⁷ result, for example, from the analysis of phytoplankton and zooplankton smaller than two millimetres, for which each sample is sub-sampled. In case of phytoplankton, only around 1/10,000th of a silk area is looked at under the microscope. The analysts further use a number of fixed abundance categories, which are subsequently converted into estimates for the quantity of organisms of a specific taxon on a sample. Richardson et al. (2006: 63) explain that ‘abundance estimates from individual plankton samples are inherently imprecise because of variable zooplankton behaviour such as diel vertical migration and local weather conditions that can concentrate or disperse fine-scale patches (Robertson 1968), as well as the “broad-brush” counting procedures.’

As CPR data do not reflect total quantities of organisms in the ocean, the data are usually not expressed in units such as organisms per cubic metre of sea water; instead, they remain expressed in the unit ‘numbers per sample’, which is an estimate derived from the hand-written records (Richardson et al. 2006: 62).

Batten et al. (2016) is a localised study in fisheries ecology and an example of how semi-quantitative data are used. The study uses indices calculated from CPR data to explain variability of the Prince William Sound herring’s first year growth. Annual abundance anomalies for groups such as large zooplankton, small zooplankton, or diatoms were calculated and then correlated with estimates of herring growth rates calculated from scale size measurements. Figures in the study use ‘organisms (zooplankton) or cells (diatoms) per Continuous Plankton Recorder (CPR) sample’ as a unit (Batten et al. 2016: 428); the authors also explain the relation between the silk’s mesh size and filter efficiency, and clarify what their data may represent:

Only an undefined proportion of the phytoplankton and microzooplankton community ... is enumerated by CPR sample analysis. The data shown here then do not necessarily indicate whether more or less chlorophyll or ciliates were available, but as the CPR is an internally-consistent sampler, they do indicate when relatively more, or less, of the large diatoms and hard-shelled microzooplankton were present and available as a food source. (Batten et al. 2016: 429)

The specifications of the used silk and material interactions at the sampling stage between the silk and plankton organisms thus affect how many organisms end up on the silk, the quantities subsequently recorded by analysts in their notebooks, and how the data can be used to create knowledge of the ocean ecosystem.

¹⁷ Johns seemed to be searching for the right term before saying “foibles”. However, the term seems very fitting, as it refers to a ‘minor flaw or shortcoming’, but not as a complete fault or failure. Persons or things with foibles are still valued and useful, despite minor shortcomings; <<https://www.merriam-webster.com/dictionary/foible>> [accessed 24 August 2017].

4 Material Integration and Continuity

The previous sections illustrate how many of the sample's material characteristics that restrict how the object can be manipulated and used originate when the CPR is in the water. By contrast, the materials themselves, the silk, the steel, and the organisms, have their respective origins in factories, in plankton life cycles, or even further back. In the course of a CPR tow, physical parts of both the sampling technology and the ocean ecosystem not only “clash” against each other; they become integrated. A variety of effects during integration—some of which are described above—lead to the formation of a novel object, the silk roll, which is later processed into individual samples.

Material integration is a constitutive phase and can be regarded as the realisation of an ‘apparatus-world complex’, a term used by philosopher Rom Harré (2003: 28–31), who explains that a technical device is capable of being ‘integrated into a unitary entity by fusion with nature’ (Harré 2003: 28); furthermore, ‘the apparatus and the neighbouring part of the world in which it is embedded constitute one thing’ (Harré 2003: 29).¹⁸ The point is that the material integration realised in the CPR Survey is a constellation that results in the constitution of a new research object with properties that have been shaped during integration by material interactions.¹⁹ Both the plankton organisms and the silk are physically transformed during the integration: the organisms are immediately deformed and the silk assumes a different colour. The silk as well as the organisms are constitutive parts of the newly formed object and a research sample in the CPR Survey could not exist without either one.

My understanding of “integration” as the constitution of a new research object resonates with Tempini’s (this volume a, b) account of assembling and integrating data from various sources to create new digital datasets. There is obviously a strong contrast between a sample integrated physically from silk, ocean water, and marine organisms and digital data that have been integrated from various datasets by computational commands; however, epistemologically, both integration procedures are geared towards forming objects that are analysable and meaningful in specific epistemic contexts.

In my case, it is important that the very materials that have been integrated are preserved throughout various stages of transportation, unloading, cutting, analysis,

¹⁸Rheinberger’s (2010: 217–218) description of an ‘intersection’ as a ‘surface’, ‘plane’, or ‘point of contact’ between a technical device and the object studied by scientists is similar to Harré’s apparatus-world complex; according to Rheinberger, an interface is a ‘fertile analytical constellation’, which certainly resonates with the idea that new entities are “born” during sampling.

¹⁹While this is not describing a case of reproduction, my view of silk rolls as novel objects, from which individual samples are created, is inspired by Griesemer’s (2014: 39–40) view of hybrids as individuals in biological reproduction; individuality is not an intrinsic property of certain objects, but can be understood as designating a relation between attention, abilities, and interest of the person tracking a phenomenon and properties, relations, behaviours, and activities attributed to what is being tracked. My account tracks materiality and contrasts with a view of the sample as a mere assembly of materials which could easily be disassembled to its original components.

and long-term storage. In the CPR Survey, material continuity is achieved between the silk roll's formation process out in the oceans and the object that is placed under a microscope and eventually archived in Plymouth. In his account of biological reproduction, Griesemer (2014: 26–27) emphasises the notion of 'material continuity' and material 'overlap' between parent and offspring when 'organized material parts' are transferred between the two; form or information are transferred materially and not by any kind of impression or translation to a different medium. Although being pressed severely into the silk, the plankton material usually remains sufficiently organised for the sample analyst to identify and count the organisms using specific tools and methods of manipulation.

Rheinberger (2015: 323–325) asserts preparations a materiality and durability similar to the research samples in the CPR Survey: Preparations 'participate in, are part of, the very materiality of the object under scrutiny'; their 'configuration' is expressed in physical, biological, and chemical properties (Rheinberger 2015: 323). A CPR silk sample has assumed a specific configuration that makes it analysable and the configuration is preserved by material continuity.²⁰ It is important, however, that "preservation" and "continuity" are not intended to imply that samples are immutable or "frozen": Due to the formalin, the organisms' green colour fades over time²¹; their spatial arrangement on the silk changes when plankton are removed and put back onto the silk during the zooplankton eyecount; and samples in the archive might get contaminated and slowly decay, impeding the ability to perform a re-analysis. Material continuity is an absence of "jumps" from one medium to another, as in the hand-written recording of plankton counts or the digitisation of hand-written notes.²²

Material integration and material continuity frame an understanding of "materiality" that—despite being based on the physicality of objects and practices—emphasises the epistemological significance of material objects over characteristics that categorise objects ontologically. The next section discusses exactly how materiality shapes scientific practices.

²⁰Rheinberger (2015: 323) further claims that 'preparations are renderings, not representations' with a 'particular indexicality' that points to themselves and not to something that is represented by the preparation. The material characteristics of the silk samples seem to point primarily to the processes involved in their formation; additionally, the bias between the number of organisms on the sample and plankton distributions in the ocean poses questions regarding the samples' potential use as representations. These issues relating to scientific representation require deeper discussion elsewhere.

²¹The survey derives a set of data from the colour of returning silk samples, as sample colour is used as an indicator of relative phytoplankton biomass in the geographical area of the tow. Due to fading of the colour, the assessment is performed when the silk roll is cut into individual samples and can only be performed once.

²²The lack of translation to another medium is another reason why considering samples as straight-up representations is problematic (see note 20); a sample is a product of continuity starting with the fusion of materials in the oceans, and not by intentionally writing or imprinting information onto a medium.

5 Scaffolding Sample Analysis and the Creation of Knowledge

A CPR sample's physical properties require specific epistemic practices that are applied to the sample or to the data that have resulted from the analysis. The examples described in this chapter are the removal of plankton organisms from the silk, on-the-job transfer of identification skills, and the consideration of relative quantities and trends instead of total quantities. Regarding the removal of large zooplankton from the silk, the scraping together of organisms, the Bogorov tray, the additional microscope, and the manipulation of organisms are artefacts and practices, which scaffold the identification and counting of the organisms. Without this step, the identification would hardly be possible, be much more difficult, or at the very least take much longer to perform. The plankton analyst faces what Caporael et al. (2014a: 15) call a 'productive resistance or challenge', which can be overcome through scaffolding. The aided identification results in a growing volume of scientific data created from an individual sample, and eventually in growth of the database and of the data's interpretive scope. Besides development and maintenance, growth, as a change of size or status without change of organisation, is a plausible function or goal of scaffolding procedures, as Caporael et al. (2014a: 15–16) remark. Similar to a scaffold that is removed from a building after construction work has finished, the additional tray is removed, the organisms are placed back onto the silk and evenly spread out. Except for an altered distribution of the larger organisms, which has never been recorded in any way before the removal of organisms, no visual characteristic of the sample indicates that the scaffolding procedure and the identification of large zooplankton have been performed.

The second example, the on-the-job training of analysts, is a scaffold that develops the skills and capacities of the laboratory staff. Frequent interactions between experienced analysts and new staff members scaffold the acquisition of identification skills, which could hardly be learned without the informal exchanges. Challenges and resistance are caused by the deformed appearances of the organisms, the specific composition of various species on samples depending on the region they are from, or any kind of unusual or surprising encounter on a sample. This type of on-the-job development of capacities and resolving of challenges is an example of what is called 'developmental agent scaffolding' by Caporael et al. (2014a: 15), which is characterised by cooperation and response between agents and their targets rather than just by application of an artefact or structure. The scaffolding in this example is anything but permanent, as people in the lab are not constantly assisting each other; it is utilised as needed, either if new analysts receive basic training, if a special expertise is going to be acquired, or if an analyst is simply in doubt about an organism's taxonomic identity.

The third example of scaffolding relates to the interpretation of the semi-quantitative data created by sample analysis. Although the distribution of organisms on a sample is not representative of the species' total quantities in the ocean, researchers are capable of creating knowledge about the oceans with the data. The use of the data is scaffolded by multiple studies carried out throughout the history

of the CPR Survey into the technical details and uncertainties introduced by material interactions such as clogging of the silk. This is how the survey has accumulated ‘considerable evidence’ (Richardson et al. 2006: 61) that CPRs filter each taxon consistently and that the data reflect the patterns and trends of the plankton in the ocean. SAHFOS has likewise conducted studies regarding the effects of different ship speeds: The average speed of ships has almost doubled since the 1950s and in general, not all ships tow CPRs at the same speed due to season, weather, or other restrictions (SAHFOS 2016: 19; Batten et al. 2003: 200–02).²³

Knowledge and evidence accumulated from these studies scaffold long-term consistency of the sampling and data analysis methods; the consistency, in turn, scaffolds commensurability and comparability of data created decades apart. A wide range of knowledge claims about the ocean ecosystem, especially those based on averaged data, depend on this commensurability. Only because the methods of sampling and data creation have been maintained for multiple decades, the CPR data are as valuable and relevant for plankton science as they are today.

As Caporael et al. (2014a: 16) explain, ‘maintenance seems more different from development than it really is’; in a dynamic system, ‘maintenance sustains a steady state, that is, it preserves organization in the face of stress, deterioration, and change, so maintenance is a change operation’ (Caporael et al. 2014a: 16). In the face of uncertainties, the inner consistency of the CPR Survey is maintained, although potential “foibles” (as the lab manager called them) may be maintained in the data as well. After decades of performing sampling and analysis the same way, the practice has become historically “entrenched” (Wimsatt 2014). However, the use CPR data still hinges on the abilities to evaluate the data’s accuracy and potential bias; each study of the survey’s materiality develops this ability. Along with the material interactions themselves, such scaffolds shape the data practices in my case.

Similar to other scaffolds, efforts aimed at understanding the materiality are expended on different time scales than the CPR Survey as a whole, because they are normally time-limited projects explicitly concerned with one detail or interaction. These studies are not completely invisible, as they are frequently published in scientific journals or referenced in publications using the data. In terms of scaffolding, however, this referencing seems more like a certificate that a development has happened or that a particular aspect of the survey is being maintained. The scaffolding itself, that is the actual practice aimed at development, has been removed, whereas the developed skill or capacity has been internalised.²⁴

²³The effects of the towing speed on the average depth and filter volume of the CPR are still not fully understood; experiments from 2015 showed greater depth with higher towing speeds, but earlier studies suggested a constant towing depth independent of speed (SAHFOS 2016: 19; Batten et al. 2003: 201–02). The average increase of speed from around 10 knots in the 1950s to around 20 knots today had a negative effect on the towing stability. By 1970 more and more CPRs were actually torn off and lost. As a consequence, a stronger and more flexible steel wire was introduced since 1976 (Batten et al. 2003: 199).

²⁴“Internalisation” is also a characteristic of scaffolding; a capacity, a skill, or sometimes the entire scaffold may be internalised by the developed structure, so that it is not visible from the outside; the internalised scaffold (for example a new method, or new knowledge) may then become a stable platform for new scaffolding procedures (Wimsatt and Griesemer 2007: 245). In the CPR Survey,

6 Conclusion

My study of an example of long-term ecological monitoring in ocean science emphasises the importance of samples and material interactions during their formation for epistemic processes and data practices. Materials of the sampling device interact with materials of the research target in ways that require transient and dynamic scaffolding activities²⁵; scientists apply specific practices and techniques to material objects in order to achieve results and progress that would not be realisable otherwise or only realisable with much more difficulty and under much higher economical costs. The continuity of methods, how scientific practice can remain unchanged in the context of historical developments, deserves particular emphasis and certainly offers opportunities for intriguing philosophical study. Without scaffolding the continuity of sampling and data practices, much of the data in my case study would hardly be usable at all to study long-term changes of the ocean ecosystem. Temporary scaffolds are necessary in order to keep an historically “entrenched” scientific method stable for decades and in order to learn about sources of uncertainties in the resulting data.

This chapter approaches the materiality of scientific objects by regarding it as the integration of physical parts from different sources into one novel entity and as the realisation of material continuity—a preservation of physical matter without any “jumps” to a different medium—throughout the epistemic process; this approach is not intended as a readily generalisable definition of the term “materiality”. The aim of this chapter was to flesh out the epistemological relevance of material interactions by showing how such interactions between research technologies and research targets can shape data journeys.

References

- Ankeny, Rachel A., and Sabina Leonelli. 2011. What’s So Special About Model Organisms? *Studies in the History and Philosophy of Science* 42: 313–323.
- Batten, Sonia D., et al. 2003. CPR sampling: The technical background, materials and methods, consistency and comparability. *Progress in Oceanography* 58: 193–215.
- Batten, S.D., S. Moffitt, W.S. Pegau, and R. Campbell. 2016. Plankton Indices Explain Interannual Variability in Prince William Sound Herring First Year Growth. *Fisheries Oceanography* 25: 420–432.
- Caporael, Linnda R., James R. Griesemer, and William C. Wimsatt. 2014a. Developing Scaffolds: An Introduction. In *Developing Scaffolds in Evolution, Culture, and Cognition*, ed. Linnda

any new capacity or knowledge about the survey itself may function as a platform for further development of capacities in the future.

²⁵Using Pickering’s (1993: 567) words, the structure of such transient and dynamic activities can be interpreted as ‘activities in an evolving field of human and material agencies reciprocally engaged in the play of resistance and accommodation’ or in other words, as a “mangle of practice”.

- R. Caporael, James R. Griesemer, and William C. Wimsatt, 1–20. Cambridge MA/London: The MIT Press.
- , eds. 2014b. *Developing Scaffolds in Evolution, Culture, and Cognition*, Vienna Series in Theoretical Biology. Cambridge, MA/London: The MIT Press.
- Duarte, Carlos M., Just Cebrián, and Núria Marbá. 1992. Uncertainty of Detecting Sea Change. *Nature* 356: 190.
- Edwards, Paul. 2010. *A Vast Machine: Computer Models, Climate Data, and the Politics of Global Warming*. Cambridge, MA/London: The MIT Press.
- Edwards, M., et al. 2011. *Fish Larvae Atlas of the NE Atlantic. Results from the Continuous Plankton Recorder Survey 1948–2005*. Plymouth, UK: Sir Alister Hardy Foundation for Ocean Science.
- Griesemer, James R. 1990. Material Models in Biology. In *Proceedings of the Biennial Meeting of the Philosophy of Science Association*, Volume 2: Symposia and Invited Papers, 79–93.
- . 2014. Reproduction and the Scaffolded Development of Hybrids. In *Developing Scaffolds in Evolution, Culture, and Cognition*, ed. Linnda R. Caporael, James R. Griesemer, and William C. Wimsatt, 23–55. Cambridge, MA/London: The MIT Press.
- Harré, Rom. 2003. The Materiality of Instruments in a Metaphysics for Experiments. In *The Philosophy of Experimentation*, ed. Hans Radder, 19–38. Pittsburgh, PA: University of Pittsburgh Press.
- Hays, Graeme C. 1994. Mesh Selection and Filtration Efficiency of the Continuous Plankton Recorder. *Journal of Plankton Research* (4): 403–412.
- Hunt, Brian P.V., and Graham W. Hosie. 2006. Continuous Plankton Recorder Flow Rates Revisited: Clogging, Ship Speed and Flow Meter Design. *Journal of Plankton Research* 28: 847–855.
- Lauro, Federico M., Svend Jacob Senstius, Jay Cullen, Russell Neches, Rachele M. Jensen, et al. 2014. The Common Oceanographer: Crowdsourcing the Collection of Oceanographic Data. *PLoS Biol* 12: e1001947.
- Leonelli, Sabina. 2016. *Data-centric biology: a philosophical study*. Chicago, IL: Chicago University Press.
- . this volume. Learning from Data Journeys. In *Data Journeys in the Sciences*, ed. Sabina Leonelli and Nicolò Tempini. Cham: Springer.
- McQuatters-Gollop, Abigail, et al. 2015. The Continuous Plankton Recorder Survey: How Can Long-term Phytoplankton Datasets Contribute to the Assessment of Good Environmental Status? *Estuarine, Coastal and Shelf Science*.
- Morgan, Mary S. 2003. Experiments Without Material Intervention: Model Experiments, Virtual Experiments, and Virtually Experiments. In *The Philosophy of Experimentation*, ed. Hans Radder, 216–235. Pittsburgh, PA: University of Pittsburgh Press.
- Owens, Nicholas J.P. 2015. 'Director's Review', in *2014 Annual Report*. Plymouth, UK: Sir Alister Hardy Foundation for Ocean Science.
- Parker, Wendy S. 2009. Does Matter Really Matter? Computer Simulations, Experiments, and Materiality. *Synthese* 169: 483–496.
- Pickering, Andrew. 1993. The Mangle of Practice: Agency and Emergence in the Sociology of Science. *American Journal of Sociology* 99: 559–589.
- Reid, P.C., J.M. Colebrook, J.B.L. Matthews, J. Aiken, and Continuous Plankton Recorder Team. 2003. The Continuous Plankton Recorder: Concepts and History, From Plankton Indicator to Undulating Recorders. *Progress in Oceanography* 58: 117–173.
- Rheinberger, Hans-Jörg. 2010. *An Epistemology of the Concrete: Twentieth-Century Histories of Life*. Durham/London: Duke University Press.
- . 2011. Infra-Experimentality: From Traces to Data, from Data to Patterning Facts. *History of Science* 49: 337–348.
- . 2015. Preparations, Models, and Simulations. *History and Philosophy of the Life Sciences* 36: 321–334.
- Richardson, A.J., A.W. Walne, A.W.G. John, T.D. Jonas, J.A. Lindley, D.W. Sims, D. Stevens, and M. Witt. 2006. Using Continuous Plankton Recorder Data. *Progress in Oceanography* 68: 27–74.

- Sir Alister Hardy Foundation for Ocean Science. 2015. *2014 Annual Report*. Plymouth, UK: Sir Alister Hardy Foundation for Ocean Science.
- . 2016. *2015 Annual Report*. Plymouth, UK: Sir Alister Hardy Foundation for Ocean Science.
- Tempini, Niccolò. this volume-a. The Reuse of Digital Computer Data: Transformation, Recombination and Generation of *Data Mixes* in Big Data Science. In *Data Journeys in the Sciences*, ed. Sabina Leonelli and Niccolò Tempini. Cham: Springer.
- Tempini, Niccolò. this volume-b. Visual Metaphors: Howardena Pindell, Video Drawings, 1975. In *Data Journeys in the Sciences*, ed. Sabina Leonelli and Niccolò Tempini. Cham: Springer.
- Wimsatt, William C. 2014. Entrenchment and Scaffolding: An Architecture for a Theory of Cultural Change. In *Developing Scaffolds in Evolution, Culture, and Cognition*, ed. Linnda R. Caporael, James R. Griesemer, and William C. Wimsatt, 77–105. Cambridge, MA/London: The MIT Press.
- Wimsatt, William C., and James R. Griesemer. 2007. Reproducing Entrenchments to Scaffold Culture: The Central Role of Development in Cultural Evolution. In *Integrating Evolution and Development: From Theory to Practice*, 227–323. Cambridge, MA: MIT Press.
- Woolgar, Steve, and Javier Lezaun. 2013. The Wrong Bin Bag: A Turn to Ontology in Science and Technology Studies? *Social Studies of Science* 43: 321–340.
- Wylie, Alison. 2016. How Archaeological Evidence Bites Back: Strategies for Putting Old Data to Work in New Ways. *Science, Technology, & Human Values* 42: 203–225.
- . this volume. Radiocarbon Dating in Archaeology: Triangulation and Traceability. In *Data Journeys in the Sciences*, ed. Sabina Leonelli and Niccolò Tempini. Cham: Springer.

Gregor Halfmann earned his Master of Science degree in Physical Oceanography at the University of Hamburg and a Master of Arts in History and Culture of Science and Technology at Technical University of Berlin. In 2014, he joined the “Epistemology of Data-Intensive Science” research project at Egenis, Centre for the Study of Life Sciences, University of Exeter, where he completed his PhD thesis in Philosophy of Science, titled “Seafarers, Silk, and Science: Oceanographic Data in the Making”, which focuses on epistemological, material and social conditions of data practices in an empirical example of long-term monitoring of oceanic plankton ecology. He is now back in Germany working as a teacher.

Open Access This chapter is licensed under the terms of the Creative Commons Attribution 4.0 International License (<http://creativecommons.org/licenses/by/4.0/>), which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons license and indicate if changes were made.

The images or other third party material in this chapter are included in the chapter’s Creative Commons license, unless indicated otherwise in a credit line to the material. If material is not included in the chapter’s Creative Commons license and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder.

