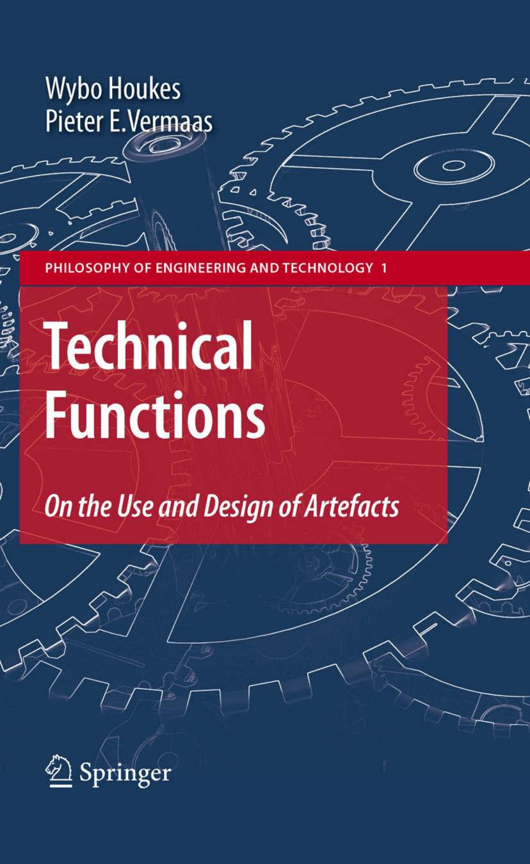

Wybo Houkes
Pieter E. Vermaas

PHILOSOPHY OF ENGINEERING AND TECHNOLOGY 1

Technical Functions

On the Use and Design of Artefacts

Springer

Philosophy of Engineering and Technology

Philosophy of Engineering and Technology

VOLUME 1

Further volumes of this series:
http://www.springer.com/series/8657

Wybo Houkes · Pieter E. Vermaas

Technical Functions

On the Use and Design of Artefacts

Springer

Wybo Houkes
Philosophy & Ethics of Technology
Eindhoven University of Technology
The Netherlands
w.n.houkes@tue.nl

Pieter E. Vermaas
Department of Philosophy
Delft University of Technology
The Netherlands
p.e.vermaas@tudelft.nl

ISSN 1879-7202
ISBN 978-90-481-3899-9 e-ISBN 978-90-481-3900-2
DOI 10.1007/978-90-481-3900-2
Springer Dordrecht Heidelberg London New York

Library of Congress Control Number: 2010922452

Printed on acid-free paper

Springer is part of Springer Science+Business Media (www.springer.com)

Contents

Preface

This book is about the functions of technical artefacts, material objects made to serve practical purposes; objects ranging from tablets of Aspirin to Concorde, from wooden clogs to nuclear submarines. More precisely, the book is about using and designing artefacts, about what it means to ascribe functions to them, and about the relations between using, designing and ascribing functions. In the following pages, we present a detailed account that shows how strong these relations are. Technical functions cannot be properly analysed without taking into regard the beliefs and actions of human beings, we contend.

This account stays deceptively close to common sense. After all, who would deny that artefacts are for whatever purpose they are designed or used? As we shall show, however, such intentionalist accounts face staunch opposition from other accounts, such as those that focus on long-term reproduction of artefacts. These accounts are partly right and mostly wrong — and although we do take a common-sense position in the end, it is only after sophisticated analysis. Furthermore, the results of this analysis reveal that technical functions depend on a larger and more structured set of beliefs and actions than is typically supposed. Much work in the succeeding pages goes into developing an appropriate action-theoretical account, and forging a connection with function ascriptions.

This goes to show that artefacts and their functions are a complicated and rewarding topic for philosophical analysis. To be sure, function talk about artefacts does not present philosophers with the same problems as function talk in biology. Throughout this book, we treat artefacts and their functions as an autonomous topic of inquiry. This is an implicit (and sometimes explicit) rejection of accounts that assimilate artefacts and organisms, or that treat all functional discourse on a par. Accounts of technical functions have long been treated as a straightforward corollary of accounts of biological functions. We show that, once technical functions are a topic in their own right, the straightforward connection to biology is lost.

Our account is primarily a construction. We have constructed an analysis of using and designing artefacts in terms of plans; and we have constructed an account for function ascriptions by a set of three conditions. There may be alternatives to one or all of these results. To make it possible for others to construct such alternatives, we set out our 'design specifications' in the introductory chapter. It is possible to go beyond the (indeed rather elementary) phenomena on which these desiderata are based. Then, our proposals may no longer be useful and more sophisticated constructions might be called for.

Although all material included in this book is original to it, we addressed the main topics in a series of earlier papers. The use-plan analysis of using and designing given in chapter 2 was first presented in 'Design and Use as Plans' (*Design Studies 23*, 2002; with Kees Dorst and Marc J. de Vries). In 'Actions versus Functions' (*Monist 87*, 2004), we gave a modified and shortened version of it, and argued that it undermines function essentialist views in metaphysics — an argument presented in greater detail in chapter 7 of this book. The ICE-function theory went through its own *Werdegang*. In embryonic form, it was added to the critical analysis of etiological theories in 'Ascribing Functions to Technical Artefacts' (*British Journal for the Philosophy of Science 54*, 2003). A more developed form was presented in 'Technical Functions' (*Studies in History and Philosophy of Science 37*, 2006). The present work contains the fully matured ICE-theory, which is properly integrated with the use-plan analysis. Integration steps in designing are seldom trivial, and this one is no exception: the ICE-theory presented in chapter 4 of this book is significantly different — and significantly more successful in terms of the standards set out in chapter 1 — than earlier versions.

We worked on the precursor papers and a first draft of this monograph while we were both post-doctoral researchers in the 'Dual Nature of Technical Artefacts' program at Delft University of Technology. We are grateful to the other researchers in this program, Maarten Franssen, Peter Kroes, Anthonie Meijers, Jeroen de Ridder and Marcel Scheele, for numerous comments on equally numerous drafts, and more general discussions.

Many people outside Delft provided comments on our ideas, at their various stages of development. We are especially indebted to Stefano Borgo, Larry Bucciarelli, Massimiliano Carrara, Randall Dipert, Kees Dorst, Sven Ove Hansson, Philippe Huneman, Ulrich Krohs, David de Léon, Tim Lewens, Françoise Longy, James McAllister, Joe Pitt, Beth Preston, Hans Radder, Norbert Roozenburg and Marzia Soavi for their response to talks and written material. Two anonymous readers from Springer provided helpful comments on the penultimate draft.

Research on this work was made possible by the Netherlands Organisation for Scientific Research (NWO).

Chapter 1

Introduction

This book is about many of the most mundane objects surrounding us. It is about the objects that we use at home, outdoors or at work; objects otherwise as diverse as tea bags, television sets, bridges and microchips. Throughout this book, we shall refer to such objects as 'technical artefacts'. Typically, these are tangible, material objects that serve a practical purpose, either incidentally or regularly. By calling these objects 'artefacts', we take them as objects that have been created, sometimes by ourselves, but more commonly by others. We may occasionally build a makeshift bridge for personal use but most of the bridges that we encounter were built by other people. By calling these objects 'technical artefacts', we bring into focus the skills involved in taking objects as serving our practical purposes. It takes experience about materials to create that makeshift bridge, and those other bridges are typically the work of a specific group of trained professionals, namely engineers. Engineers design most of the objects that we use, and some of those objects are available only because engineers designed them: the creation of microchips exceeds the skills of most people, but evidently not those of all people. By focussing on technical artefacts our analysis excludes, in first instance, objects such as laws and organisations ('social artefacts'), statues and symphonies ('aesthetic artefacts' or 'works of art'), and theories and models ('scientific artefacts'). But we do not limit our analysis to engineering; this book is about technical artefacts broadly conceived, analysing objects ranging from everyday items such as tea bags and television sets, to technologically complex objects such as bridges and microchips. Our analysis gives an integrating account of this spectrum of objects, which may even be expanded to include natural objects such as stones and batches of water that serve practical purposes. In short, our analysis is about useful material.

In this book, we focus on what appears to be — and in fact is — the central feature of technical artefacts: their intimate connection to teleology.

The need for an analysis of the teleology of technical artefacts is not self-evident. After all, artefacts have been described in teleological terms for ages; and whereas these terms have become problematic in other domains in which they were once applied — most notably, of course, in biology — they continue

W. Houkes, P.E. Vermaas, *Technical Functions*, Philosophy of Engineering and Technology 1, DOI 10.1007/978-90-481-3900-2_1,

to be used in the domain of technical artefacts, apparently to everyone's satisfaction. Philosophical scrutiny seems uncalled for. Yet, on closer inspection, artefact teleology is more problematic than one might think. This is illustrated by the notion of function. Nothing seems more common-sensical than describing artefacts in terms of their technical function or functions: even artefact kinds that are not explicitly named after their functions are easily categorised in terms of functions. There may be counterexamples, i.e., artefacts that *cannot* be characterised functionally, but since it is sufficiently difficult to find such examples, functional artefacts are the rule. Consequently, some philosophers have gone as far as claiming that functions are *essential* to artefacts. Yet, despite the general emphasis on functional characterisations of artefacts, there is no consensus about who and what determines technical functions. Moreover, most of the existing attempts to resolve this issue have been so sketchy that they raise more problems than they solve.

One traditional answer is that the intentions of agents fix the functions of technical artefacts: technical functions are characterised as intended effects. But developing this answer to full-fledged theories, theories that we will call *intentional function theories*, only leads to further questions: which agents? Can every user determine his or her individual functions? Or is function-setting the prerogative of engineers who design the technical artefacts; is function-setting somehow part of their professional duties? If so, which of the many intentions, wishes and beliefs of these agents are relevant for determining functions? Another, traditional answer is Robert Cummins' (1975) function theory, in which the functions of an item correspond roughly to the causal contributions the item makes to systems containing it. This theory, which we call the *causal-role function theory*, raises issues about how to single out the right causal contributions as functions of artefacts. Artefacts make all sorts of those contributions and not all correspond to their functions. If the intentions of agents single out functional contributions, we are back at the questions raised by the intentional function theories. Moreover, artefacts may — unfortunately — sometimes fail to work by not making the contribution to realising the practical purpose for which they are used. In that case Cummins' theory cannot take that non-existing contribution as the function of the 'malfunctioning' artefact. A third and less traditional answer is that intentions are largely irrelevant in determining artefact functions. Instead, these functions are shaped by evolutionary forces of variation and selection, much like those that shape the biological world. And indeed, artefacts have to survive in an even more competitive environment than many organisms, and the history of technology — especially in the last two centuries — is one of continuous mass extinction. There is no doubt that there are many similarities between the natural and the artificial realm. The way functions are determined may be one, but also the *evolutionist function theories*[1] this third perspective leads to, raise several questions, most obviously concerning the relevant processes of selection and the remaining role of purposive design

[1] In philosophy of biology and philosophy in general, these evolutionist function theories are better known as etiological theories.

and use.

The core chapters of this book are devoted to a new theory about artefact functions that falls squarely in the intentionalist tradition. We shall explore the similarities between the natural and the artificial realm insofar as functions are concerned, and we shall find that they are insufficient to overthrow the tradition, but more than sufficient to refine existing intended-effects accounts. Defending intentionalism on artefact functions is, despite the entrenchedness of this perspective, surprisingly difficult, and we shall find that answering the questions to which the intended-effects account gives rise requires incorporating elements of the causal-role and of the evolutionary perspective. The function theory that results from this operation is called the *ICE-theory*, to honour its three ancestors while putting the intentionalist *I* first. But although we give designer intentions priority in determining the functions of artefacts, we can only avoid the problems hinted at above by endorsing a rather liberal and decidedly non-standard account of designing. On this account, Alexander Graham Bell, who developed the first telephone to aid the hard-of-hearing, counts as a designer but so do later engineers, who adapted telephones for use as a general communication device, and even innovative consumers who use their telephone to listen in on their sleeping children.

We can bring this last balancing act between the priority of the intentions of designing pioneers, of redesigning engineers and of innovative users to a successful conclusion only by firstly analysing artefact teleology in general. As the earlier questions about agents and intentions show, an accurate intentional theory of technical functions requires an analysis of the using and designing of these artefacts. The analysis that we present is an action-theoretical one that incorporates some epistemological notions. This task involves some trailblasing. Of course, the theory of action is a well-established part of contemporary philosophy, but to the best of our knowledge only Randall Dipert (1993, 1995) has attempted to apply it to artefacts. We acknowledge Dipert's work as an important source of inspiration. In our exploration of the unfamiliar terrain of artefact use and design, we draw upon more general action-theoretical analyses. In particular, we adopt the notion of plan and reshape it for our own purposes, leading to a 'use-plan' analysis of artefact using and designing.

By addressing the phenomenon of artefact teleology, this book breaks new ground. Our function-theoretical project is, to a large extent, located within a well-articulated — some might say, over-articulated — philosophical debate. Yet, since most work in this field has been concerned with the understanding of biological functions, both critical distance and considerable sophistication were needed to arrive at a function theory for technical artefacts that is worthy of its name. By contrast, constructing an action-theoretical analysis of artefacts, for which there are relatively few reference points in the literature, led us to complement typical analytic step-by-step arguments with a more explorative mode of analysis.

Most broadly, we aim to provide the groundworks for a philosophical analysis of artefacts. This goal is realised in several ways. First and foremost, we study and analyse the basic concepts in terms of which artefacts are described, concepts such as 'using', 'designing' and 'function'. We offer arguments for taking some of these concepts as basic and others as peripheral. In fact, one of the results of our efforts is that functionality is not as important for describing artefacts as it is often taken to be. This calls for a change of focus: for properly understanding technical artefacts philosophers, but also engineers, should consider the intentional actions that involve these artefacts instead of merely regarding them as functional objects. Second, by analysing and clarifying artefact functionality and teleology, we examine the intuitive distinctions between technical artefacts and other objects — in particular, between artefacts and natural objects including biological organisms. And finally, we show that several features in the domain of technical artefacts can be accounted for by terminology and themes familiar from analytic philosophy: we analyse actions in terms of *rationality* and *plans*, which provide a background for a theory of *functions*; and we draw on resources mined in disciplines such as *action theory* and *epistemology*. This choice means that we approach both artefacts and the actions in which they play a role largely from a *normative* rather than a descriptive perspective. We do not offer a theory about how people actually use or design artefacts, or how they in fact describe them in functional terms; instead, we seek to provide a framework for evaluating some aspects of these activities, and we theorise about rational and proper artefact use, and about justifiable function ascriptions.

Conceptual engineering

Before presenting an overview of this book, a few remarks about our method are in order. Our aim to develop a function theory for technical artefacts based on an action-theoretical analysis of artefact using and designing, calls for a careful choice of means. The first part, developing a function theory, is a theme that has become increasingly familiar in philosophy over the last decades, although it is regarded by many as an exhausted field, characterised by what two authors have memorably called 'the dull thud of conflicting intuitions.'[2] Indeed, function theory occasionally gives the impression of philosophical angler's tales. Unlikely events like nuts and screws falling into machines and making them work and Bibles preventing people from being shot in the heart appear to be the yardstick by which the performance of function theory is measured. The problem is not so much that an appeal to intuitions is needed to account for these cases, but that intuitions about them are weak and almost bound to diverge, and that it is unclear what is at stake in accounting for these cases in the first place. For our project, this problem seems to arise with a vengeance, because we plan to cover some unfamiliar ground: the domain of technical artefacts. Although we do attempt to phrase our analysis in philosophically familiar terms, such as 'rationality' and 'justification', the fact remains that, because few philosophers have thought and theorised about artefacts, intuitions are likely to be unschooled, weak and divergent.

[2] Bigelow and Pargetter (1987, p. 194).

The proper-accidental desideratum:
A theory of artefacts should allow that artefacts have a limited number of enduring proper functions as well as more transient accidental functions.

The malfunctioning desideratum:
A theory of artefacts should introduce a concept of a proper function that allows malfunctioning.

The support desideratum:
A theory of artefacts should require that there exists a measure of support for ascribing a function to an artefact, even if the artefact is dysfunctional or if it has a function only transiently.

The innovation desideratum:
A theory of artefacts should be able to ascribe intuitively correct functions to innovative artefacts.

Table 1.1: Four desiderata for a theory of artefacts

Our response is not to avoid an appeal to intuitions, but to make this appeal as explicit and circumscribed as possible. In line with the subject matter of this book, we take an engineer's attitude towards our intuitions: we list our intuitive, phenomenological 'data' and then translate them into clear specifications — or desiderata, as we shall call them — for a theory of technical artefacts.[3] We take these desiderata, and these alone, as touchstones for our own theory. Furthermore, our phenomenological data are relatively unassuming, leading to minimal specifications for an effective function theory. Still, it is possible that someone disagrees with our choice of phenomena and intuitions, or that someone doubts whether these should be accounted for in terms of functions. Some of these disagreements and doubts will be addressed later, but we are not apologetic about our choice of desiderata: it is a choice and therefore, to some extent, arbitrary. Yet the only way of disagreeing productively with our choice is to construct alternative desiderata, and an alternative theory, leaving the ultimate choice to the users of both theories. Similarly, our tightly circumscribed goal means that we regard successful uses of our conceptual apparatus for other purposes than satisfying our desiderata as, at most, beneficial side-effects of our efforts.

Table 1.1 lists our four desiderata for theories of technical functions. As said, each of these desiderata captures an aspect of everyday involvements with artefacts and reflects an assumption that this aspect ought to be accounted for in terms of the functions of these artefacts. Thus, two choices are made for every desideratum: we choose to select one aspect of artefact using or designing as especially salient, and we choose to hold a theory of functions accountable for this aspect. In the remainder of this section, we shall briefly justify both

[3]This way of appealing explicitly and exclusively to certain intuitions is not original to our book. Our method is similar to, for instance, Jackson's recent attempt at defending conceptual analysis (Jackson 1998).

choices.

The four desiderata reflect, consecutively, the following four phenomena: use versatility, possible lack of success, physical restriction and innovation. Each of these is a broad and variegated phenomenon, best described by means of multiple examples and real-life narratives; in this respect, they are on a par with phenomena such as object persistence and personal identity, which are also encountered in many different forms and guises. For the sake of brevity, we only give a few short and simple illustrations, just to show the intuitive appeal and wide scope of the phenomena and the resulting desiderata.

First, artefact use is versatile. Virtually every artefact can be used for different purposes and in different ways. Chairs may be used for sitting on, for resting one's legs on while sitting, for standing on; and one can sit up straight or slouch in a chair. Cars may be used for transporting people from one place to another, for relaxation, even for ramming the front of a store to make possible a robbery. Not all of these uses are on a par, however. Chairs are most standardly or most appropriately used for sitting on. The force of this standard is not just one of numbers. Using a car for personal transportation is common practice; intentionally ramming it into a storefront is not just unconventional, but also a crime. By contrast, standing on a chair to change a light bulb is at most mildly frowned upon and not uncommon. Standing on a swivel chair to change the light bulb that hangs over a staircase may raise eyebrows, but presumably for different reasons than it does in other situations. Moreover, these differences in evaluating common and uncommon ways of using artefacts are not just a matter of intuitions: the warranties of many products contain void clauses mentioning *improper* use. The existence of such void clauses, like that of written restrictions and prohibitions in general, shows that alternative uses are possible and that some are disapproved of. These cases show that different standards and sanctions are at work in evaluating the many ways in which artefacts are used. For the moment, we refer indiscriminately to all these standards by calling artefact use 'limited'. Many existing philosophical analyses of functions contain a distinction, or call for a distinction, that seems the perfect counterpart of the phenomenon of limited artefact-use versatility. For it is common practice to make a distinction between an item's proper function(s), which are more or less persistent, and its more transient accidental features, and to maintain that a function theory needs to honour and explicate this distinction.[4] Thus, it makes

[4]Many theories in the literature aim at distinguishing proper and accidental functions, albeit often not in exactly those terms. One of the earliest places to find this distinction is Larry Wright's seminal paper on functions: 'Very likely the central distinction of this analysis is that between the function of something and other things it does which are not its function [...] This is sometimes put as the distinction between a function and something done merely 'by accident'.' (1973, p. 141). An even more influential theory of function is explicitly presented as one of proper functions, in contradistinction to accidental functions or 'functioning as': E.g., 'I have said that the definition of 'proper function' is intended to explain what it is for an item to have a function or purpose, but not what it is for an item to function as something.' (Millikan 1989, p. 290). More recently, Beth Preston (1998b) formulated a theory that makes a distinction between proper functions and system functions, and has explicitly extended it to the realm of artefacts. More precisely, she argues that a function-accident distinction as desired by Wright can be drawn in an account of proper functions and is impossible to draw

sense to propose the *proper-accidental* desideratum as the way in which a theory of technical functions should accommodate this aspect of artefact use.

Another aspect of artefact use is that success is never guaranteed. This lack of success is often a matter of degree. A television, for instance, may show unsharp images, may fail to broadcast a number of channels, may not respond to the remote control anymore, or it may have stopped working altogether, smelling of burned circuitry. Artefact failure may show qualitative differences as well. If a television set does not show any images, this provides a strong reason for calling it 'broken'; but if a twice-used tea bag can no longer be used to make a decent cup of tea, there is no such reason. Malfunctioning does not coincide with ineffective use. Again, we ignore all these fineries for the moment, despite their obvious practical relevance. Instead, we simply take the possibility of unsuccessful use of artefacts as a second salient aspect or phenomenon. Like use versatility, this phenomenon can be translated into a desideratum that is already familiar from the existing literature on functions:[5] a function theory should introduce a concept of a proper function that allows for malfunctioning, i.e., for the fact that an item has a proper function that it may be incapable of performing. This desideratum is more specific than the phenomenon, since it only concerns proper functions and use, whereas artefact breakdown is also a real possibility in cases of non-standard use. But again, we trade phenomenological scope for analytic precision, at least for the time being. In chapter 5 on malfunctioning, we take ample opportunity to remedy this shortcoming.

The third aspect is intimately related to both the first and the second, but it brings to light a different phenomenon. Artefact-use versatility is not just limited by peer pressure and void clauses. That artefact use is effective or not is not just a matter of skill and/or convention. If this needs demonstration, it is provided by cross-breeding some of the examples used so far: using a tea bag for store-ramming is probably not a crime in those legal systems where you cannot be prosecuted on the basis of bad intentions alone; and making a decent cup of tea by means of a chair is beyond even the most creative and skillful user. In these cases, the artefact is not fit for its use. Taking a tiny step in the analysis of the phenomenon, it may be said that the physical structure of, for instance, ordinary tea bags (i.e., those not filled with nitroglycerine) simply make it impossible to use them effectively for ramming a storefront. There is also a positive side to these physical limitations of artefact use, namely that we can, in many cases, reasonably expect some measure of success when using an artefact. 'Folk physics' seems sufficient to predict that one can successfully sit on a chair, or on the hood of a car. To capture both the restrictive and the positive side of this phenomenon, we call it 'support', and we introduce a third desideratum of the same name: a theory of artefacts should require that there

---

in an account of system functions (1998b, §I).

[5] The malfunctioning desideratum is implicitly introduced as a criterion for assessing function theories by Millikan (1989), Neander (1991b) and others, when they criticise the theory presented by Cummins (1975) for not being able to describe malfunctioning items as functional.

exists a measure of support for ascribing a function to an artefact, even if the artefact is malfunctioning or if it has a function only transiently.

A fourth and final aspect of artefact use is that artefacts with innovative uses, standard and otherwise, appear regularly. Such innovations come in gradations. The majority, such as flat-screen televisions, allow for minor changes in existing ways of using artefacts.[6] Others, such as cellular phones-with-camera, combine two existing artefacts in one. The most extreme and rare examples may be the first airplane and the first nuclear plant, which allow for truly innovative, unprecedented ways of using artefacts. Even for these examples, one might want to point out predecessors, toning down the strength of the innovation. Nevertheless, innovative artefacts — or artefacts that can be plausibly advertised as such — are highly valued in our modern society. Reflecting this phenomenon, we take it as our final, *innovation* desideratum that a theory should be able to ascribe intuitively correct functions to innovative artefacts.

We end this methodological section with a word of warning. Although we have now mainly used the phenomena to introduce a list of desiderata for theories of technical functions, we keep in mind the underlying phenomena while constructing a theory that satisfies the desiderata. It might be the case that no single function theory can account for all of the phenomena: they might be mutually incompatible.[7] This incompatibility may lead one to adopt a kind of function pluralism, in which several notions satisfy the desiderata together. Alternatively, one might come to the conclusion that some of the phenomena should not be accounted for in terms of technical functions in the first place. Of course, accounting for all aspects of artefact use in terms of one concept would be elegant and thus certainly worth striving for. In this book, we show that a theory of technical functions can bear most of the phenomenological burden that we have now saddled it with.

Outline

After this introduction, we present the action-theoretical background for our theory of technical functions. In chapter 2, we analyse artefact using and designing in terms of one central, versatile concept, that of 'use plan', i.e., a more or less standardised way of manipulating objects in order to realise a practical goal. After introducing use plans by means of examples and some analysis, we show how artefact using can be reconstructed as the execution of a use plan, and how this reconstruction can be used for evaluative purposes. We introduce a number of standards on use plans, and show how they can also be applied to artefact use. Then, we turn to a closely related activity, namely designing. Artefacts are commonly taken to be man-made objects; and conversely, designing is usually understood, in philosophy and elsewhere, as the creation of artefacts — or perhaps as creating blueprints, i.e., descriptions of a procedure for creating

[6]Many innovations sought after in industry might even have only an indirect relation to artefact use, namely those innovations that concern ways of producing artefacts more reliably or cost-effectively.

[7]This possibility was raised by Preston (2003) when she considered earlier formulations of these four desiderata given in (Vermaas and Houkes 2003).

the artefacts. For our purposes, designing is reconstructed more broadly. We characterise it as, first and foremost, constructing use plans intended to contribute to other agents' realising their goals. Sometimes and, if so, secondarily this is supplemented with creating artefacts and/or blueprints. We denote this secondary activity as 'product designing', and explicitly reconstruct it as a variety of our broad notion of designing. This notion brings out the central role that constructing use plans and communicating them to other agents play in designing. Closing the chapter, we construct an evaluative framework for using and designing, based on our use-plan analysis.

Work on the action-theoretical background already goes some way towards fulfilling our phenomenological requirements. Yet it mainly prepares the ground for the next chapters, in which we analyse the teleology of artefacts more directly in terms of their functions. As said before, artefacts are primarily characterised in philosophy as intentionally created functional objects. This functional approach has led to various function theories for artefacts and may inspire various others. To prepare the ground for our own theory of functions, we present a critical review of existing and possible theories of technical functions in chapter 3. This review establishes that the functional descriptions of artefacts in currently existing function theories cannot meet our desiderata. This review, especially our presentation of the various theories, is instrumental to the development of our alternative, which does meet the desiderata. We organise the review by considering three archetypical function theories — the intentional, causal-role and evolutionist function theories — which span up the spectrum of currently existing function theories.

This review of function theories provides the means for taking the next step in our analysis. In chapter 4, we add a function theory of our own making to the description of artefact using and designing as developed in chapter 2. We construct our function theory on the basis of the use-plan approach, or more specifically, on the basis of our reconstruction of designing in terms of use plans. It consists of two definitions of what it means to justifiably ascribe functions to artefacts, relative to use plans of those artefacts. Together with their action-theoretical background, these definitions amount to a theory of technical functions that meets our desiderata. Since the function theory incorporates elements from the intentional, causal-role and evolutionist function theories, we call it the ICE-theory.

One complication arises. The ICE-theory formally meets the four desiderata. Yet the extent to which it meets the *malfunctioning* desideratum leaves something to be desired. Looking back at the phenomenology, the theory turns out to describe only some cases in which one might want to call an artefact 'malfunctioning'. On the scale of our desiderata, this is only a minor disadvantage, but there are at least two reasons to set it straight. Firstly, artefact malfunctioning is a very common and variegated phenomenon, so that a theory of technical functions that only accounts for some fragments of it is impoverished, whatever its other virtues might be. Secondly and more importantly, the mediocre performance of our theory with respect to the malfunctioning desideratum might reveal a tension between our four requirements. As said above, we require a measure of support for every function ascription, even to severely malfunctioning artefacts. Meeting this support desideratum might inevitably decrease a theory's performance for the malfunctioning desideratum. If this is so, one might conclude that the desiderata are incompatible, not that our theory is inadequate.

In chapter 5, we address this complication. We start by identifying various types of artefact malfunctioning, in order to describe more precisely to what extent our theory meets the malfunctioning desideratum. We continue with presenting two routes to broaden the extent in which the ICE-theory describes malfunctioning artefacts. The first consists of a new argument about what it means to believe that an artefact has a capacity to perform a function. The argument aims at establishing that such a belief is not equivalent to the stronger belief that the artefact is in a state in which it can exercise that capacity. The reason is that it allows the belief that the artefact can be brought in or returned to such a state in a technologically acceptable way, say, by maintenance or repair. Acceptance of this argument broadens the scope of the ICE-theory significantly; more precisely, it covers malfunctioning claims of all the types identified previously in the chapter. This shows that our four desiderata are mutually compatible, but also opens the possibility that other function theories meet this desideratum more easily. The second route broadens our analysis in another way, by reflecting on the normative content of malfunctioning claims, and that of claims about artefacts in general. We reveal two types of normative content, one related to practical reasons, and found in all function ascriptions, and one related to the privileging of use plans and the role of professional designers, which is found in proper-use claims and some malfunctioning claims. The second type in particular leads us beyond a theory of technical functions to its social and action-theoretical background, and reveals a complicated network of social, social-epistemic and practical recommendations and requirements.

Once our ICE-theory of technical functions is in satisfactory shape, we explore its merits in other domains. We start within engineering, and then move step-by-step, via physics and chemistry, towards biology. In chapter 4, when formulating our two definitions of what it means to justifiably ascribe functions to artefacts, we presented them as central to our theory but not as exhaustive. Our reason for this caveat is that agents can arrive at functional descriptions of technical artefacts without considering use plans. And because these descriptions cannot be plausibly reconstructed as function ascriptions relative to such plans, we enriched our function theory with a subsidiary notion of plan-less ascriptions of *functional roles* to artefacts. In our discussion of functional descriptions in engineering in chapter 6, we first consider such plan-less functional roles in more detail, and examine to what extent they challenge the central, plan-relative notion of functions. We prove that functional descriptions of components by engineers can be taken as plan-relative: references to plans in the relevant definitions can be suppressed when applied to components. The plan-relative notion of artefact functions thus remains to stand central in engineering. Yet, somewhat paradoxically, the added subsidiary notion of ascribed functional roles enhances the possibilities to apply the ICE-theory to domains other than engineering or technology in general. Physics and chemistry also seem to contain functional descriptions. Systems can be described as measurement or preparation devices, and physical and chemical substances can be characterised as conductors or solvents. These functional descriptions can easily be accommodated in the ICE-theory as function ascriptions relative to use plans and as plan-less functional roles, respectively.

In biology, clearly the main and most interesting field for applying function theories, functional descriptions do not fit our use-plan analysis. In the ICE-theory, biological functional descriptions may be taken as the ascription of functional roles, but that is insufficient for reproducing the standard practice in biology to consider items as having (proper) functions. This negative result leaves two options for contributing to the analysis of biological functions, which we discuss in the second half of chapter 6. The first, critical option consists of arguing that contemporary attempts in the philosophy of biology to analyse biological functions as 'as-if' technical functions fail. The ICE-theory shows that taking biological items as if they are technical artefacts implies the implicit but implausible acceptance of a series of teleological concepts, in particular that of use plans, within the domain of biology. The second, more daring option is to accept explicitly the teleology presupposed by the ICE-theory, in all domains that feature functional descriptions. We show that the ICE-theory can be generalised to a uniform function theory in which agents ascribe, relative to goal-directed patterns, functions to items that are part of those patterns. This generalised ICE-theory applies to biology but has the consequence that also biological functions are ascribed by agents and become teleological.

The upshot of our discussion is that many functional descriptions in technology, physics and chemistry can be taken as function ascriptions relative to use plans, showing the flexibility and versatility of the ICE-theory. For biology, the ICE-theory leads to a dilemma: the teleological background of the ICE-theory forces one to either give up on a uniform analysis of functions and seek for a separate theory for biological functions, or to hold on to such a uniform analysis on pain of accepting this teleological background.

This ends our analysis of using, designing and functional descriptions. In chapter 7, the coda of this book, we study some consequences for understanding the nature of artefacts. The material presented in this book offer grounds, perhaps surprisingly, that this nature is not essentially functional. We review our major results to show how they undermine function essentialism, a popular view on the nature of artefacts. Then, we study an alternative view of artefacts, as objects manipulated in the course of executing use plans. Although this characterisation of artefacts as useful materials has some drawbacks, it is superior to function essentialism and can be profitably combined with another common characterisation of artefacts, namely as man-made objects. In summary, we arrive at an analysis of artefacts as objects with a twofold dual nature: they are objects that have intentional characteristics and that have physical characteristics, as well as objects that are used and that are man-made. Functional descriptions are relevant to the first, intentional-physical duality since these descriptions allow users and engineers to connect and disconnect teleological and structural descriptions of artefacts. Hence, technical function is a useful concept, that serves as a conceptual hinge between the two natures of artefacts.

Outlook

This book presents useful material. It provides an action-theoretical analysis of artefact using and designing, and a theory of function ascriptions to artefacts that is phrased in terms of the beliefs and intentions of various agents, and in terms of the evidence for the relevant beliefs. Together, these results constitute the groundworks for a philosophy of artefacts, in which 'using', 'designing' and 'function' are inevitably central concepts. We do not, however, paint a complete picture of technical artefacts. Our function theory is spelled out in considerable detail, and, together with its background, it accounts for some essential aspects of the phenomenology of artefact using and designing. Large parts of this phenomenology, and other concepts needed to account for it, remain to be explored.

One important self-set limitation concerns the topic of functionality. Our analysis focuses on technical artefacts and their technical functions: we study material objects that were designed to serve some more or less immediate practical purpose, and we study only those practical purposes. We ignore useful objects of questionable functionality, such as art works, useful objects of questionable materiality, such as computer software and tax-evasion schemes; and purposes for material objects that are not immediately practical, such as the social status that your clothes or car bring you. Future research may find a place for these other artefacts and these other aspects of technical artefacts in the picture we painted. In addition, any work on functions faces the question of whether and how it applies to the various types of items to which functions are ascribed. Among these, biological items are historically and currently most prominent. In chapter 6, we sketch some ways in which our ICE-theory, or a generalised version of it, can be applied to biology, the other natural sciences and the humanities. Detailed research is needed to examine whether the theory can stand its ground as an immigrant in these wider and more crowded domains.

Our function-theoretical project also sets limitations for the action-theoretical background. Firstly, as shall become clear in the next chapter, we intentionally neglect several features of artefact using and designing that we deem irrelevant for our theory of functions. Secondly, we largely ignore other artefact-related activities, such as production, repair and recycling. Functionality and usefulness may be distinctive features of artefacts, but it is a wide-spread and apparently crucial intuition that artefacts are objects that are produced by intentional agents. Therefore, in the final chapter, we briefly address this intuition and examine to which extent we can accommodate it in our own, useful-material account of artefacts; in all other chapters, we focus almost exclusively on the activities of using and designing, which suffice as a background for our function theory.

Finally, the philosophy of artefacts has a natural place as part of the philosophy of technology. This discipline is, to a large extent, devoted to ethical and societal issues concerning technology. Our action-theoretical analysis of artefact using and designing may shed new light on these issues. It offers one way of analysing and evaluating the impact that designers and their artefacts have on society: an analysis in terms of testimony and the communication of use plans, which may be more or less rational. Both the philosophy of technology and science and technology studies have offered different images. Most of these are not aimed at analysis, and few researchers in these other disciplines share our focus on evaluation rather than description. Yet combining these perspectives requires more than a simple division of labour between analysis and synthesis or description and evaluation, and we shall not even attempt such a combination here. Yet, since both perspectives are based on the phenomenology of human dealings with artefacts, combining them is bound to lead to a more in-depth image of a domain that is so close to us that philosophers have tended to overlook it.

Chapter 2

Using, designing and plans

We start this book by presenting the action-theoretical background for our theory of technical functions. To this effect, we study two commonplace actions that involve artefacts: using and designing. We analyse both of these activities in terms of one central concept, that of use plans, *i.e., ways of manipulating objects in order to realise practical goals.*

Most of the chapter is devoted to presenting our analysis of using and designing. In section 2.1, we introduce our action-theoretical perspective, and in section 2.2, our central concept of use plans. Since this is just the start of a background story, we tell it by means of simple examples and provide no more than a basic analysis. In sections 2.3 and 2.4, we analyse using and designing in terms of plans: the former is cast as the execution of a use plan, and the latter as the construction and communication of a use plan. For our purposes, this broad notion of designing suffices. To make room for ordinary usage, we describe in section 2.5 how product designing may be distinguished as a type of designing that also involves the description or production of a previously unavailable object.

Our use-plan analysis is not only useful as a background for a theory of technical functions. In the final part of the chapter, we show how focussing on use plans can provide a framework for evaluating artefact using and designing. In section 2.6, we introduce standards on use plans, and in section 2.7, we demonstrate how these standards may be applied to the execution and construction of use plans. This leads to a characterisation of distinct types of using, which we call 'rational using' and 'proper using', and to several standards on designing.

W. Houkes, P.E. Vermaas, *Technical Functions*, Philosophy of Engineering and Technology 1, DOI 10.1007/978-90-481-3900-2_2,

2.1 Artefacts and actions

Using things to achieve a goal is such a commonplace activity that most of the time it comes naturally. Most people do not have to think twice about wearing their glasses, cycling along a road, or speaking into a telephone. These actions all involve objects other than human bodies, but performing them is as natural as waving at a friend or wiggling your toes. Artefacts are not just a dominant part of our environment: using artefacts is such a dominant part of our lives that it is easier to think of examples of artefact use than of things that we literally do with our bare hands.

Philosophers, like everybody else, tend to overlook artefact use. But as important as observing the 'natural' world and theorising about it are, these activities are highly atypical. Most people spend most of their time doing things with things, using stuff to achieve some goal or another. Even philosophers need more than their brains to get their papers published. Insofar as philosophers have looked at artefact use, they were mainly working in or inspired by the phenomenological tradition. Moreover, authors such as Merleau-Ponty and Heidegger were mainly concerned with the very fact that artefact use is so natural, and they did not address artefact design at all.[1]

As a background for our function theory, we need to focus on both artefact use and artefact design, and we need a framework for analysing the relation between the two. Unlike work done in the phenomenological tradition, we focus on the role of reasoning, deliberation and evaluation in human involvements with artefacts. This places us in the tradition of philosophical action theory. This discipline centres on issues such as the distinction between intentional action and mere behaviour, the differentiation of actions, and the reasons instead of — but in relation to — the causes for actions.

Reasoning and deliberation apply to artefact use and design because our typical involvement with artefacts is goal-directed. Indeed, the notion of *using* an item seem to indicate that this item is a means to an end: playing with a tennis ball while having a conversation is not naturally described as using the ball. Describing some activity as using of an object x invites the question what x is used *for*. Using is using with an aim or goal. Similarly, 'designing' carries strong connotations of purposiveness and intentionality. Describing an activity as 'goal-directed design' does not seem right, because goal-directedness appears to be an integral part of designing. In this chapter, we bring out the goal-directedness of artefact using and designing, because we are convinced that artefacts only have functions by virtue of the fact that human beings are capable of (goal-directed) use and design of artefacts. Action theory provides the means

[1] This 'unobtrusiveness' of artefacts as used objects is the cornerstone of Heidegger's description of the everyday world in the *First Division* of *Being and Time*, and of his distinction between the ready-to-hand items (*Zuhandenes* or *Zeug*) of everyday life and the present-at-hand objects (*Vorhandenes*) of science. To quote one typical description of this phenomenon of unobtrusiveness: 'Equipment can genuinely show itself only in dealings cut to its own measure (hammering with a hammer, for instance); but in such dealings an entity of this kind is not grasped thematically as an occuring thing, nor is the equipment-structure known as such even in the using.' (Heidegger 1962, p. 98)

to express this conviction, and to argue for it.

In focussing on evaluation and on goal-directedness, as we do, the activities of using and designing artefacts lose some of their naturalness. Making explicit the reasons why using an artefact is an effective way of realising a goal is far less straightforward than simply using it; and designing an artefact that successfully contributes to a particular goal is typically even more difficult. One type of complication has drawn the bulk of philosophical attention, namely that virtually all artefact use and design involves skills that people have to acquire through personal experience.[2] Anyone who has watched a small child trying to use a spoon knows that even apparently elementary artefact use involves such skills.

But although skills are essential to using artefacts, there is another, frequently overlooked reason why artefact use is complicated: in the vast majority of cases, it involves more than one action. Just as we have to learn to use artefacts skillfully, we have to learn to take the appropriate actions in the appropriate order. Take, for instance, using a toaster for making toast. This requires taking the toaster from where it is stored, plugging it in (needed only if one does not always keeps one's toaster ready for action), putting pieces of bread into the toast rack, setting the toasting time, pulling down the toast rack by means of a switch, etc. Taking the bread from the rack before the rack is pulled down or forgetting to plug in the device is not going to result in toast. This description is by no means intended to make people doubt their capacity to use a toaster, but only to show that the use of this artefact is quite naturally described as a sequence of actions, some of which might be conditional on other actions. Several of the actions require manipulations of the artefact and its components (the rack, the switch). Only taken together, as a series, they are expected to lead to realisation of the goal. Once someone decides to use a toaster in order to achieve a goal, she does not decide to manipulate this artefact in some arbitrary way: instead, this decision concerns an ordered series of actions, which constitutes a determinate way of realising a goal by means of the artefact.

We seek to reconstruct the process of deliberation that underlies artefact use. Therefore, we need to look at the series of actions that constitutes use, and we need a concept that is tailor-made to describe such series. This concept is that of *use plan*: a goal-directed series of actions, including manipulations of the artefact and its components. Some philosophers of action have emphasised the role of plans in *all* intentional action,[3] both inspired by and inspiring work on artificial intelligence. We do not consider these general merits of a plan-centred approach. Doing so would lead us into several philosophical debates that would complicate our background story without significant advantages for our function theory. Therefore, we remain neutral on, for instance, the question

[2] Classical analyses of the role of skills in using artefacts and in human life in general are Ryle's (1949, ch. 2) description of knowing-how, Merleau-Ponty's (1962) example of the blind man's cane, and Polanyi's (1962) plea for tacit and personal knowledge.

[3] Our presentation in this chapter is influenced in particular by the work of Michael Bratman (1987, 1999) and John Pollock (1995).

whether a planning approach is superior to the desire-belief model of intentional action. For our purposes, it does not matter whether planning can be reduced to a series of intention formations, as long as the (possibly irreducible) concept of use plans is available for use in our action-theoretical background.

In section 2.3, we show explicitly how artefact use can be reconstructed as the execution of a use plan. First, we introduce our key concept in some more detail.

2.2 Use plans

'Plan' is a versatile notion, and quite common in everyday language — unlike notions such as 'artefact' and 'function'. Because it is so versatile, the notion is useful for giving a rich, intuitively appealing account of artefact use and design. Yet versatility may also lead to vagueness. The *Oxford Concise English Dictionary*, for instance, describes a plan as 'a detailed proposal for doing or achieving something', which immediately raises questions concerning the level of detail, the nature of the proposal, and the measure of expected success in planning. Is a plan a written, spoken or merely mental proposal? How successful does it need to be to count as a proposal for 'doing' something, let alone for 'achieving' something? Can it leave room for improvisation or does it need to take into account all eventualities to count as a full proposal? And to whom is the plan proposed?

Most of these issues may be resolved by stipulation. We choose to regard plans as complex, mental items that consist of *considered* actions, not of *actual* actions.[4] If the considered actions that constitute the plan are carried out, the plan is *executed*. This execution is a physical process that involves the human body and possibly other material objects. If artefacts are involved, we call the plan a *use plan*. Execution of a use plan thus involves using artefacts. We refer to the mental process of planning as *constructing* or *designing* a use plan. This process results in a more or less durable mental state, similar to a belief or intention, and different from a wish or a fancy.

To illustrate some features of use plans, consider a relatively straightforward activity such as making tea. The physicochemical process underlying the brewing of tea is well-known: flavours are extracted from tea leaves or other herbs, either fresh or dried, by putting them in hot water. Still, there are many different ways of making tea, all of which involve taking a series of rather elementary actions. One, not very refined, 'camping-ground' way is to put a substantial amount of dried tea leaves in a cup, pour boiling water over it, wait a short while, and to start drinking. Alternatively, one may use a sieve or a special tea strainer and a second cup to remove the tea leaves from the beverage before drinking it. Many companies provide leaves in tea bags, along with instructions to put them in containers with a specific amount of freshly boiled water, and to

[4]Proponents of a planning approach in the theory of action, such as Bratman and Pollock, do not explicitly resolve this ambiguity. For a position in AI planning largely compatible with our stipulative resolution, see Pollack (1990).

brew the tea for a specific amount of time before removing the bag and pouring the tea into a cup. Dutch tea bags have strings attached, so that they can be suspended in a tea pot and removed at will. Many American tea bags are to be used in percolators instead. Sophisticated tea drinkers prefer the use of fresh, cold water, boiled very briefly to prevent loss of oxygen and a consequent loss of taste, and they insist that the tea leaves be stirred gently, if at all, to prevent extraction of tannins along with the flavours. Finally, the mundane activity of brewing and drinking tea has been taken to extremes of sophistication in Chinese and Japanese tea ceremonies. These consist of large numbers of steps, only some of which involve the actual brewing and drinking of tea.

These various ways of making tea, from the camping-ground way to the tea ceremonies, are recognisably different. Among other things, they include different steps, or similar steps in a different order, and they involve the use of different objects. This can be captured by describing each way of brewing tea as the execution of a different plan. The camping-ground way, for instance, involves executing the following plan:

1. *Put dried leaves into a sufficiently large cup.*
2. *Pour hot water into the cup.*
3. *Wait.*
4. *Drink.*

whereas a tea drinker who is both sophisticated and Dutch might execute the following plan:

1. *Boil fresh, cold water.*
2. *Pour the water in a tea pot.*
3. *Suspend a tea bag in the pot.*
4. *Wait.*
5. *Remove the bag from the pot.*
6. *Pour tea from the pot into a cup.*
7. *Drink.*

The two ways of brewing tea can straightforwardly be reconstructed as use plans — goal-directed series of actions, including manipulations of an artefact and its components. After all, they contain various steps, including the manipulation of various objects: tea leaves, hot water and a cup in the former, a tea bag, boiled water, a tea pot, and a cup in the latter. Moreover, the two ways of brewing tea can be reconstructed as executing *different* use plans: apart from the inevitable waiting and drinking, the plans do not have a single step in common, and they involve the manipulation of different sets of objects. Still, because the two plans have roughly the same goal state, and since both include the manipulation of tea leaves, both are naturally described as plans for 'brewing tea'. The goal and most salient manipulated object[5] together determine a coarse-grained description of this use of an artefact.

[5] This salience is hard to describe exactly, but expresses the intuition that separating one's garbage, in case it involves the manipulation of tea leaves, is not *using* tea.

Authors who have defended a planning approach in the theory of action have emphasised several useful features of plans, most of which presumably transfer to use plans. For our purposes, the standards on plans and the evaluative possibilities entailed by these standards are the most useful feature. We treat these at some length in section 2.6. Another feature is the communicability of plans, which is discussed by Michael Bratman (2000). Plans make explicit which steps ought to be taken in order to realise a goal. These steps and their ordering can be communicated verbally: if an agent who knows how to realise a certain goal tells another how he went about realising it, he communicates a series of actions. Communicating this 'procedural' aspect of artefact use does not, of course, immediately give the other agent the capacity to realise the goal, since some or all steps in the plan may require skills that the other agent does not possess. Still, this lack of skills does not, in principle, prevent the agent from understanding the procedure: to a large extent, the procedural and operational aspects of artefact use, i.e., the steps to be taken and the actual taking of the steps, may be separated. Recipes are an obvious case in point: the procedure to make lamb curry can be communicated from one agent to another through writing a recipe. Reading this recipe does not produce the capacity to make lamb curry, but it does transfer relevant, procedural information. The plan is a significant part, perhaps even the core, of this information.

Other features that we shall use repeatedly, albeit largely implicitly, involve the structure of plans. Plans are more than sets of actions: the main reason to introduce the notion is, as described above, that the ordering of these actions matters. Therefore, plans can be described in terms of their structure. Pollock (1995, ch. 6) develops this structural approach by presenting plans as graphs, and defending that plans are structurally similar to programs in involving conditionals, loops and variable-binding. We do not need to go this far, but some structural features are useful for telling plans apart. At some points in our work, we need to distinguish use plans from one another. For this purpose, we do not require intricate sets of necessary and sufficient conditions; instead, we employ the following features:

Aim:

The simplest and most telling way to differentiate use plans is to check whether executing them is intended to realise the same goal. A use plan for making tea is different from a use plan for going to Amsterdam. The context of executing use plans may generate some complications; for instance, the sophisticated-Dutch use plan for making tea as part of a tea ceremony might be distinguished from a plan containing the exact same actions, but just aimed at producing a cup of tea to drink while watching a soccer match. Yet this problem looks sufficiently similar to the well-known complications regarding coarse-grained and fine-grained actions to suspend further discussion here and leave it to the general philosophy of action.

Objects used:

Use plans can also be differentiated by the objects manipulated in them. The camping-ground plan for making tea involves the manipulation of a different set of objects than the sophisticated tea-bag plan for making tea. Therefore, even though their goals are the same and they are both use plans for tea leaves, it makes sense to speak of two plans for making tea rather than two versions of the same plan. We primarily apply this criterion by identifying the entire set of objects manipulated in the course of executing the use plan. Indeed, our use-plan approach does not distinguish between the primary object in use and what may intuitively be called 'auxiliary items'.[6]

(Ordering of) considered actions:

Occasionally, the goal state and objects involved are insufficient to differentiate two use plans that are intuitively different. Conceivably, one might use the exact same objects involved in the sophisticated-Dutch tea-making plan to make a cup of tea in a different way. In this case, a detailed comparison of the considered actions involved in both plans is in order. In many cases, merely studying the structure of their ordering suffices to see whether plans are different; but in some cases, inspection of the content of the actions is required. Here, complications are bound to arise,[7]but we rarely, if ever, need this level of sophistication for our function-theoretical purposes.

[6] Nor does the approach distinguish between artificial and natural objects that are manipulated while realising a use plan. In chapter 7, we take up the question whether the use-plan analysis can be regarded as an implicit definition of 'artefact'.

[7] Listing the steps does not completely individuate plans, since the description of actions is notoriously sensitive to their context. To give one overdramatised example, suppose one takes part in a car race on a motorway at night, perhaps in an attempt to emulate James Dean. It may be plausibly claimed that the use plan of a car is similar when racing and when driving to work — for instance, if the goal is described as 'driving from A to B'. Specifying the goal in this coarse-grained way allows a coarse-grained description of the actions included

2.3 Planning in use

Deliberation may play two roles in using artefacts. Firstly, someone may deliberate about the way to realise a pre-set goal. Taken in this sense, one may deliberate about using an artefact or not: you might drive to the supermarket or decide to walk there instead. Secondly, someone may deliberate about how to manipulate the artefact in order to realise the goal. Thus, you might think about how to drive a car, in case you have forgotten the details of your driving lessons. The two roles are clearly interrelated, since the choice for using an artefact will usually depend on whether the agent knows how to use it — in evaluative terms: the former deliberation will be unsuccessful if the latter is. To combine both examples: you have a good reason to walk to the supermarket in case you have forgotten your driving lessons.

In this section, we discuss the role of deliberation in the process of artefact use. This discussion makes explicit the relation between use plans and artefact use. This brings out some advantages of our approach, but it also shows a problem: our reconstruction seems to paint a very limited or even straitjacketed picture of artefact use. The real benefits of our use-plan approach do not lie in its psychological accuracy, but in its ability to provide a basis for evaluating using and designing, and to raise and answer philosophical questions regarding artefacts and their functions.

To make our reconstruction explicit, we elaborate slightly on the household example by which we introduced the notion of use plans. Suppose Anna desires a piece of toast and decides to use her trusty toaster. She takes the machine from the cupboard and plugs it in. Then she puts two slices of bread in the toast rack and pushes down the switch. The toaster has a small dial that determines when it will switch off. According to the manual, users should determine which setting of the dial ('1' to '5') goes with their preferred 'level of browning' of the toast, put the dial at that level and then wait until the toaster switches off and the toast rack lifts. Yet what Anna usually does is turn the dial to some higher-than-desired setting (say, '5'). Then, she uses either her experience with this type of bread or frequent checks of the 'level of browning' to determine when she wants the toasting to stop. At that moment, she turns the knob to the lowest setting, causing the toaster to switch off and the toast rack to lift. In other words, Anna uses the dial as a glorified on/off-switch.

This episode may be described as a series of decisions or intention-formations: Anna's decision to use a toaster, to push down the switch, to turn the dial to a higher-than-desired setting, etc. With each decision, a new intention is formed, which is subsequently realised, after which another decision is taken and a new intention formed on the basis of the current situation. This is the standard reconstruction of intentional actions offered in action theory and rational choice

in the use plan. For instance, part of the use plan of one's car in a rally may be to check regularly for approaching opponents in the rear-view mirror. Taken in a coarse-grained way, this racing plan is similar to the lightweight use plan that is executed when driving a car in all circumstances. Hence, lists of considered actions will not always suffice to determine whether plans are identical.

U.1.	user u wants to bring about some goal state g, and believes it does not obtain.
U.2.	u chooses, from a set of available alternatives, a use plan p for bringing about g that involves the intentional manipulation of objects $\{x_1, x_2, \ldots\}$.
U.3.	u believes that p is effective, i.e., that executing p will bring about g.
U.4.	u believes that his or her physical circumstances and set of skills support realising p.
U.5.	u intends to execute p and acts accordingly.
U.6.	u observes g' as the outcome of p and compares g' with g.
U.7.	u considers whether g has been brought about or not. If not, u may decide to re-execute p, to repeat step U.2, or to abandon his goal. If U.2 is repeated, u may reconsider his desired goal state g, select another use plan, or do both.

Table 2.1: Reconstruction of using

theory. This model may be defensible,[8] but this case and using artefacts in general are more naturally described as the execution of a use plan. Anna's plan to use her toaster consists of a number of steps: plugging it in, pushing down the switch, turning the dial clockwise, waiting and observing, and turning the dial counter-clockwise. If Anna considers her manipulations of the toaster at all (we will raise that point in a moment), she does not consider them individually, but only as steps in realising a goal.

Hence, we propose to reconstruct artefact use as the execution of a use plan, as given in table 2.1.[9] Some remarks are in order. First, the beliefs mentioned in the various steps are used to assess the relative merits of the available alternatives; they form the 'belief base' of the use plan.[10] We examine this belief base further when we look at plans from an evaluative point of view. Second, the beliefs mentioned are oversimplified. In many cases, plans are chosen or constructed on the basis of comparative and/or less apodictic beliefs, such as 'I believe I'm better at carrying out p than p', although I'm not quite sure under the present circumstances'. Third, we have not included the option that the user constructs her own use plan for achieving the desired goal. The reason is that a user who constructs a use plan is involved in designing, on the analysis presented in the next section. The present reconstruction thus only applies to what we later characterise as *passive using*.

The reconstruction of using artefacts presented in table 2.1 can be illustrated by means of the toaster example. Anna deliberates on a way to obtain toast,

[8] A general argument why planning cannot be reduced to a series of intention formations is presented by John Pollock (1995, §5.2).

[9] Houkes, Vermaas, Dorst, and de Vries (2002, §1) and Houkes and Vermaas (2004, §§1–2).

[10] In AI planning, the belief base is usually called the 'knowledge base' of the plan; our terminology is chosen in accordance with the tradition in epistemology to distinguish beliefs from knowledge. Thanks to Marcel Scheele for pointing this out.

given her desire for and present lack of toast. As said above, this deliberation is not understood as a sequence of intention-formations, but as the adoption of a single plan consisting of a sequence of intentional actions for attaining the goal. This plan can be chosen from a (small) set of available alternatives, such as using a toaster, buying readymade toast, putting slices of bread in an oven. Anna chooses a plan from this set on the basis of her beliefs about the effectiveness of the plan, and about her skills and her circumstances. Once she has settled on a plan, she executes it, and judges whether the result is satisfactory. Suppose the toast comes out distastefully burned. Anna may then repeat the execution of the plan with a fresh slice of bread, change her plan (e.g., turn the knob earlier or use an oven instead of a toaster), or leave her desire to eat toast unsatisfied.

We propose our reconstruction mainly as a background for our theory of functions, and secondarily as an evaluative framework. Still, one may wonder to what extent it reflects actual artefact using. At first glance, our planning approach seems to overestimate the role of deliberation. As noted in the beginning of this chapter, using artefact is unobtrusive, often almost second nature; it does not typically involve elaborate planning.[11] In the remainder of this section, we sketch how this feature of artefact use can be accommodated in our analysis.[12] Readers who are prepared to accept our reconstruction on its possible evaluative merits, or as a background alone, may skip to the next section.

To illustrate two possible limitations of the planning approach, consider once again the use of a toaster. Reconstructed as the execution of a use plan, this use involves getting the toaster, plugging it in, putting slices of bread into the toast rack, pushing down the switch, etc. Most steps of this plan can in turn be constructed as a number of steps, e.g., the 'sub-plan' of getting slices of bread. For other steps, such a construction would be contrived. Anna does not, for instance, deliberate on how to push down the switch. In any plan, there is a point at which further deliberation becomes strongly counterproductive — no one will deliberate for a week on what to eat — or further sub-planning is infeasible, e.g., for simple bodily movements. Here, deliberation typically ends. A second limitation is that one does not generally plan every tiny detail of achieving a goal in advance, even if this would be possible in principle. To allow for mistakes in deliberation or the beliefs on which it is based, to accommodate changes in the situation and other eventualities, plans are usually incomplete. In the toaster case, for instance, Anna probably does not deliberate on where to put the toaster, but decides while realising the plan. The first limitation reflects a principled incompleteness of planning: it is impossible to specify all included actions in every detail. The second limitation reflects a more practical incompleteness: it is seldom convenient or wise to specify all actions, conditional

[11]This observation is easily turned into an objection against a planning approach to artefact use. Such an objection might also draw inspiration from Lucy Suchman's (1987) work. Suchman's primary purpose is to show that the 'planning model of agency' is incapable of describing interactions between agents and human-machine interaction and that the model should be replaced with an appeal to 'situated action'.

[12]A more elaborate defense of the descriptive accuracy of our planning approach to artefact use is presented in Vermaas and Houkes (2006b, §3).

on detailed situations, in a plan.[13]

Our reconstruction can accommodate both limitations. To start with the principled one: in realising plans, we necessarily draw on mental and motor skills.[14] Such skills are usually appealed to without further analysis: a user manual of a toaster that specifies how the user should hold her hand when pressing the switch of the toast rack would be odd; nor is a user expected to think long and hard about this step in executing the plan.[15] Skills are embedded in use plans as part of the belief base, represented as step U.4 of our reconstruction. As we shall see when discussing the evaluation of plans, this appeal to skills can be evaluated through the standard of belief coherence. Thus, if the user justifiably believes to possess the relevant skills, and if these skills are the ones presupposed by the designer of the artefact, the implicit decision to stop deliberating and start drawing on skills is a good one; otherwise, it is not. Thus, the evaluation depends both on the executor of the plan and on the nature of the task. In the case of a toaster, little harm is done if it turns out one does not in fact know how to stop the toasting process; but if one uses a parachute, it may be best to double-check whether one knows the ropes before drawing on possibly non-existent skills.

The practical incompleteness is captured in a different way. In many cases, e.g., when one is unsure about the circumstances in which a plan will be executed, it makes little sense to construct an elaborate plan for using an artefact. Elaborate plans are likely to be defeated by some aspect of the later situation, and time and effort may be wasted on attempts to continue executing the original plan. Hence, it is preferable to construct a *lightweight plan*[16] or *plan schema*,[17] which specifies only the starting actions and a broad route that leads to the goal state. Such lightweight plans are adaptable to the situation in which they are realised. Suppose one decides to use a toaster and includes in one's use plan a step of turning the knob to level 3 (out of five), only to find that one's spouse has bought a new, more sophisticated toaster with ten levels. Then one has to revise one's plan before using this toaster. If one uses a toaster by turning the dial to the maximum level and turning it back once the toast is sufficiently browned, there is no need to revise the plan. Hence, the less detailed plan is more flexible and more responsive to the specific, possibly changing situation in which it is realised.

This responsiveness can be accommodated in the planning approach. Plans

[13] These two phenomena are not always distinguished, as we think they should. Pollock, for instance, observes that plans are typically 'partial' (1995, p. 28), which corresponds to our practical limitation, but later takes up this theme again by noting that planning all desires is impossible in principle, because it would lead to an infinite regress (1995, p. 248).

[14] Pollock (1995, p. 248) refers to such skills as 'built-in procedures' or 'procedural knowledge'. We use the more colloquial 'skill'.

[15] Although the role of deliberation in our actions should not be overemphasised, a surprising amount of planning goes on in seemingly elementary actions. Haggard (1998), for instance, following in a research tradition reviewed in Rosenbaum (1991), describes the advance planning of changing one's grip on an eight-sided object.

[16] A similar point about the practical importance of lightweight planning is made by Leudar and Costall (1996), who use a 'dialogical' notion of plan different from ours.

[17] Pollock (1995, pp. 248–251).

are based on beliefs that the agent has about herself, the artefact and the environment. These beliefs enter into the standards for plans by determining the context of evaluation, as shall become clear in section 2.7 A plan to make fire with a cigarette lighter in the middle of a storm may be purposeful, but it is nevertheless a bad plan. Moreover, additional information can lead to quick revisions of a plan, given sufficient sensitivity to the context. Previous planning may even, through focussing mechanisms, be helpful in acquiring relevant beliefs about the environment, which can then be used to revise or complete the plan. In this way, planning and flexibility with respect to changing aspects of the situation in which the plan is executed may support each other.

2.4 Designing plans

We now turn our attention from using artefacts to the closely related activity of *designing*. The relation between using and designing, or users and designers, is probably intuitively evident, but our use-plan approach allows us to analyse it explicitly. The basic idea, developed in the upcoming sections, is that, in the technical realm, designing is primarily — sometimes even exclusively — constructing and communicating use plans. This characterisation does not require designing to result in material objects that were previously non-existent. Some design processes do, and such processes are valued highly in most contemporary societies. On our liberal characterisation, activities that result in new material objects are a subtype of designing, called *product designing*, and even this type of designing primarily involves the construction of use plans and only secondarily the construction of material objects or their blueprints.

Similarly liberal characterisations of designing are found occasionally in the literature on design methodology. Eekels and Poelman (1998, ch. 4) and Roozenburg and Eekels (1995, §4.3) explicitly adopt an action-theoretical basis for their account of designing, and so do Brown and Blessing (2005). Likewise, the theory of technical systems by Hubka and Eder (1988) is broad enough to include actions of human operators with technical systems in the description of designing. In general, however, design methodologists focus on the creation of new material objects, and characterise designing accordingly (e.g., Gero (1990) and Pahl, Beitz, Feldhusen, and Grote (2007)). The latter descriptions of designing are too narrow for our purposes. A more general difference between the work of design methodologists and our work is that methodologists aim at providing rules and methods in order to improve actual practices of designing, whereas we aim at reconstructing designing, in order to develop our philosophy of artefacts and, eventually, our theory of technical functions. Our reconstruction is not aimed at describing actual design practices or at revealing how designers, in the course of their work, arrive or have to arrive at their intentions, beliefs and decisions.[18] Given these differences in focus and aim, we develop our characterisation of designing from our own use-plan approach. We return to design

[18]It is possible to defend some of the more descriptive aspects of our reconstruction of designing; see (Houkes 2008).

methodology only after our reconstruction of product designing, in section 2.5.

Our characterisation of designing is broad and liberal. It allows us to conceptualise the differences between the activities or 'roles' of using and designing, and to distinguish various kinds of designing. We already contrasted product designing from designing that consists exclusively of the construction and communication of use plans. Another distinction is epistemic: some design processes build upon specialised scientific and technological knowledge, others require only common sense and some basic knowledge of physical regularities and properties.[19] We denote the former type as 'expert designing', and the latter as 'common sense designing'.[20] A third difference is that between professional and amateur designing. In our highly technological society, designing has increasingly become associated with a limited group of professionally trained and licenced persons. Yet, on our characterisation of designing, everyone, i.e., also those without such professional qualifications, is bound to be involved in it every now and then. These distinctions[21] matter for our understanding of designing in general and for the development of our theory of technical functions. If designing is product designing, use plans are typically not the sole output. The product designed should also be described, minimally by ostentation when the designer concerned has created the artefact, but more typically in terms of blueprints and instructions for manufacturing the product. If designing is amateur designing, information on how to maintain or repair the objects used may be given or not. For professional designing, especially when it is product designing, higher expectations are warranted. Users expect their newly designed cars to come with information about how to maintain it and about how to solve failures. Moreover, if maintenance demands and trouble-shooting exceed the skills of the users concerned, one may expect professional product designing to have led to the distribution of information to technical experts who can assist users in maintenance and repair. Such additional output of designing can, in coherence with our action-theoretical approach, be described in terms of (separate) plans for the manufacturing, maintenance, repair and possibly also disassembly of artefacts.

The differences in designing and the additional output and expectations play a role in our arguments, in section 4.3 and chapter 5, that our theory of technical functions meets the four desiderata for such theories. Broadly speaking, the legal and more tacit social constraints on engineering designing — which is

[19]Some cognitive psychologists denote such widely available beliefs and knowledge about properties and regularities as 'Generalized Event Representations' or 'General Event Knowledge' and have studied their role as prerequisites of planning, inferential reasoning and other cognitive activities. See, for instance, Nelson (1986), Hudson and Fivush (1991) and Hudson, Sosa, and Shapiro (1997).

[20]This terminology partly shows the lack of more appropriate terms. We do not claim that common-sense designing does not show expertise, nor that designing that is based on scientific principles (alone) always shows expertise. Those claims would run counter to many studies in both social epistemology and science and technology studies, e.g., (Collins and Evans 2002) and Goldman (1999, 2001a, 2001b). Some further reflections on the expertise of designers can be found in (Houkes 2006).

[21]In (Vermaas and Houkes 2006a, §§4–7) we have given a rudimentary framework for distinguishing various types of designing along these lines.

typically professional expert product designing — are much stricter than the expectations in play in everyday creative use of objects — which is typically amateur common-sense plan designing. We expect more when a newly designed car is introduced than when the next-door neighbour tells us about a new use for an existing kitchen appliance.

Reconstructing designing

To bring out the connection between using and designing, our reconstruction of designing builds upon our previous reconstruction of using. Since using artefacts is a goal-directed activity, it seems natural to reconstruct designing as an activity that contributes, directly or indirectly, to realising these goals. Since using is the execution of a use plan, a goal-directed series of considered actions, a natural way of contributing to goal-realisation is to provide users with a use plan that may be executed. This analysis makes for a straightforward connection between using and designing: designers support users by constructing new plans for attaining new or existing goals; conversely, every activity of constructing a use plan is, on our model, called 'designing'. The narrower notion of product designing as creating new artefacts or blueprints thereof, consists of constructing (and communicating) a use plan for realising a particular goal, and of describing some previously non-existent objects manipulated in the execution of this plan. Product designing is sufficiently prominent in actual practice to save its reconstruction for the next section.

For showing the connection between using and designing, a full reconstruction of designing is not strictly needed; from a user's perspective, the only aspect of designing that needs reconstructing is the transfer of a use plan. But if we reconstruct designing itself as a rational activity in line with the planning model of using, it makes sense to regard designing as the execution of a (design) plan with a belief basis of its own. The goal of this design plan, like that of a use plan, is to realise a state of affairs, namely the existence of a use plan that enables prospective users to realise a certain goal. To be rational, the design plan ought to aim at the construction of a use plan for goals that the designer thinks are viable and based on the designer's beliefs about the objects used, the agent executing the plan and the circumstances in which the plan is executed. Finally, if the plan executor is not the designer herself, the plan ought to be communicated to other agents. If we take all these requirements into account, we can come up with the general reconstruction of designing given in table 2.2.[22]

To give an example of plan designing that is not also product designing, suppose that someone, let us call her Juliet, is phoned by Romeo, a friend who has locked himself out of his house and wants to get in. After learning that Romeo's key is on the inside of his frontdoor lock, Juliet decides to help Romeo by telling him how to retrieve his key (adjusted goal state g', which is identical to g in this case). Despite Romeo's suggestion that she lend him her crowbar (familiar use plan p_1), she tells Romeo to take the minutes of last

[22] Houkes, Vermaas, Dorst, and de Vries (2002, §2).

Goal contribution:

D.1.	designer d wants to contribute to realising goal state g.
D.2.	*goal adjustment*: d believes that g' is the closest viable approximation of g.
D.3.	d intends to contribute to realising g'.

Plan construction:

D.4.	d intends to construct a new plan p for realising g'.
D.5.	*effectiveness*: d believes that p, which includes intentional manipulation of objects $\{x_1, x_2, \ldots\}$, is effective, i.e., that executing p will bring about g'.
D.6.	*competitiveness*: d believes that p improves on familiar use plans $\{p_1, p_2, \ldots\}$ with respect to effectiveness in realising g', i.e., that executing p realises g' more effectively than executing any member of $\{p_1, p_2, \ldots\}$.
D.7.	*physical support*: d believes that object x_1 has physicochemical capacities $\{\phi_{1,1}, \phi_{1,2}, \ldots\}$, that object x_2 has physicochemical capacities $\{\phi_{2,1}, \phi_{2,2}, \ldots\}$, etc., and that these capacities make successful execution of p possible.

Communication:

D.8.	*users*: d believes that p will be executed by prospected executors $\{u_1, u_2, \ldots\}$.
D.9.	*goal consistency*: d believes that realising g' by executing p will be compatible with goals $\{g_1, g_2, \ldots\}$ of $\{u_1, u_2, \ldots\}$.
D.10.	*skill compatibility*: d believes that $\{u_1, u_2, \ldots\}$ will have the skills required to execute p, i.e., to manipulate $\{x_1, x_2, \ldots\}$.
D.11.	*means-end coherence*: d believes that $\{u_1, u_2, \ldots\}$ will have recourse to auxiliary items needed to execute p and that $\{u_1, u_2, \ldots\}$ are capable of subsidiary planning in executing p.
D.12.	*circumstance support*: d believes that $\{u_1, u_2, \ldots\}$ will execute p in physical circumstances that support this execution.
D.13.	if d believes that $\{u_1, u_2, \ldots\}$ is not identical to $\{d\}$, he intends to communicate p and relevant parts of his belief base to $\{u_1, u_2, \ldots\}$.

Table 2.2: Reconstruction of designing

week's meeting, which they both found in their office mailboxes today, take off the paperclip holding the stack of papers together, shove one of the papers under his front door at the height of the lock, bend the paperclip into a straight wire, and prod the lock with it, until he hears the key fall on the paper, and finally gently pull the paper back to him, making sure that the key stays on it. By constructing this use plan p, which involves manipulation of paper, paperclip and key (objects x_1, x_2, $\ldots$), Juliet is plan designing (D.4). In telling Romeo to execute the constructed plan (D.13), she expresses that *he* is the one who will execute the plan and she expresses her belief that this plan will be effective (D.5) and that it is better than smashing a window or various other plans that she or Romeo might have considered (D.6). Juliet's plan is based on various physical-support beliefs that she has, e.g., about the flexibility of the paperclip, the size of the lock, the position of the key, and the space underneath Romeo's front door. Moreover, in communicating this plan to Romeo, Juliet expresses some trust in the skills of her friend (D.10) and beliefs about his other goals (D.9). In this case, these steps are rather trivial: Juliet assumes that Romeo knows how to bend a paperclip and that he is not so attached to this object that he does not want to change its original shape. Furthermore, she implicitly believes Romeo to be capable of some subsidiary planning (D.11) — e.g., in opening the briefcase in which he keeps the minutes, or in putting the paper in an appropriate place beneath the front door — and she believes that Romeo's circumstances support execution of this plan. The latter belief is, again, quite trivial, since there are few circumstances that prevent execution of this plan. If there were, or if Juliet would like Romeo to take particular care in executing some steps of the plan, she might communicate some of her belief base, or reasons behind this base, to him, in addition to communicating the plan itself (D.13).

Designing, even in this anecdotal key-retrieval example, is a more complicated process than our reconstruction suggests. Still, the given list of steps serves our immediate purposes: it makes possible an evaluative perspective on designing, and it brings out the relation between using and designing — also mainly insofar as this relation has evaluative aspects. We end this section with brief glosses on some parts of our reconstruction.

General presentation

We divide our reconstruction into three stages, which may in reality be connected by various feedback loops. These three stages concern the major aspects of designing as a goal-oriented activity: the selection of a goal to which designing is to contribute; the construction of a plan that is directed towards this goal; and the communication of this plan to other agents than the designer himself. These 'prospected executors' are, of course, none other than the users who proceed to execute the designed use plan p in accordance with the reconstruction presented in section 2.3. Furthermore, we present designing as an activity that contains several intentions and/or actions (to contribute to a goal, to construct a plan, to communicate), each with (part of) their respective belief bases. The various parts of the belief base of p have been named (in italics) for future reference.

Goal adjustment and users

Steps D.2 and D.8 show the descriptive and reconstructive side, respectively, of our analysis of designing. Goal adjustment (D.2) is a familiar aspect of design practices. In deciding how a particular goal is to be met (i.e., in constructing a plan), designers frequently change this goal. They may, for instance, decide that the original goal state cannot be realised, or that use plans realising it can only be executed by exercising rare or professional skills; alternatively, they may put additional constraints on the way in which the goal state is to be realised, reflecting concerns of durability, safety or cost effectiveness. By contrast, the 'users' belief (D.8) is an idealisation, since it suggests that designing is always intended to contribute to realising some pre-existing need. Actual designers are not so exclusively responsive, but may communicate the desirability of certain goal states along with the use plans for realising them, thus fulfilling a previously non-existing need. It may be cynical to regard such 'pro-active' designing as rule rather than exception, but there is no denying that it exists. Hence, our analysis is a rational reconstruction in that it does not represent the motivation of the designer beyond the recognition of a goal, which is needed to assess the rationality of designing.

Effectiveness and competitiveness

The two beliefs D.5 and D.6 are the core of our evaluation of designing, presented in section 2.7. Designing does not live up to its most important standard if the use plan constructed is not effective. In actual practice, competitiveness may be a more urgent matter than mere effectiveness. Designing an effective use plan may not be difficult, but the activity is a waste of cognitive capacity, time and money if it results in a use plan that is outperformed by familiar, equally effective use plans. Furthermore, newly designed use plans are often marketed as improvements upon existing plans, to facilitate their acceptance. Use plans have, in other words, to fill a practical niche or to compete successfully with existing occupants of such niches in order for designing to be worthwhile. Despite this practical relevance, we largely ignore considerations of competitiveness in this book.

Physical support

The physical support belief D.7, which is important for our theory of function ascriptions, specifies that the designer believes that the objects manipulated in the use plan have various necessary, enabling, supportive or otherwise contributing and non-detrimental capacities. Designers need not have elaborate knowledge about the physical properties and behaviour of the objects, but what beliefs they have must support their belief in the effectiveness of the use plan. Minimally, the physical-support belief is that the manipulated objects together, in the context of executing the plan, have the capacity to realise the goal state due to their physical and chemical structure.

Communication

All parts of the belief basis listed under the communication stage concern the *execution* of the use plan rather than its *construction*. In acquiring these beliefs, the designer should, so to speak, put himself in the users' shoes, in order to find out which skills and auxiliary items are available to the users, with which other goals plan execution ought to cohere, and in which circumstances users will typically execute the plan.

Our reconstruction gives a reason for the role of communication in designing: if the designer is not the only prospected executor of the use plan, designing without communication misses part of its point. Since the designing-using relation is analysed in terms of use plans, the use plan is what minimally ought to be communicated. There are many ways to do this. The lists of considered actions given as examples of use plans in section 2.2 may resemble the content of manuals. Indeed, manuals are common means of communicating the designed use plan to users and they share much of the structure of use plans. They are usually phrased as a series of instructions for actions, linearising the partial ordering of plans; typically, manuals do not say that actions may be taken in several orders. Another feature of manuals, which reflects the incompleteness of use plans, is that many instructions are somewhat indeterminate: they specify that auxiliary items should be used, or that certain actions should be taken, without saying how. One underlying reason is that it is impractical and counterproductive to specify every instruction in detail, since it would make prospective users even less inclined to read the manual. But there is also a conceptual reason: the use plans expressed in manuals are themselves incomplete, leaving room for skills and improvisation.

But manuals are just one possible means of expressing and communicating a plan. Our reconstruction is compatible with the existence of other means. Many advertisements, for instance, show examples of intended artefact use: people cleaning their floors with revolutionary new devices or making coffee with a new type of percolator. Like every communicative act, the message included in these plan-communicative advertisements functions against a background of common knowledge and presuppositions. Many advertisements present a new plan and artefact in contrast with more traditional plans and artefacts, with which the viewer is assumed to be familiar. By making this contrast, the advertisement brings across the message that the new plan and artefact are more convenient, effective, or otherwise desirable than the traditional or rival plans. Of course, not every advertisement centres on the communication of a use plan — although most show some sequences of actions — but the fact that some do is sufficient to show that there are other communicative means than manuals.

Still another means for communicating plans are training and product demonstrations. Using some artefacts may involve elaborate procedures. For these, communicating the entire use plan through an explicit demonstration may be the best means available; although all actions included in the plan could be expressed as instructions in a manual, the chances of anyone learning how to use a complicated artefact from a manual are small. Board games provide a helpful example.[23] Other cases in which demonstrations and training are effective means for communication are those in which many uncommon or novel skills are appealed to. Here, seeing someone else go through all the requisite motions and successfully use the artefact, or being trained while using the artefact may be far more effective than verbal communication of the use plan.

Whatever means is chosen, the communication of relevant parts of the belief base, or of the reasons for parts of the belief base, is context-sensitive; an example was given in the key-retrieval example of plan designing. Still, all of these means of communicating use plans go beyond merely presenting the artefact to users.[24] An exception may be common-or-garden user products that by their physical structure directly indicate how and for what they can be used. A button is immediately recognised as something to push on, and a cup with a handle is immediately recognised as something to fill and hold. However, these cases can be analysed as cases in which the communication of (parts of) the use plan is established via generally established signs incorporated in the physics of the artefact. We have all found at some time that buttons can be pushed, and handles held — such action patterns are learned at a very early age and may even be innate. Hence, designers can add such features as means for passing information about the use plan to users. These features are sometimes called 'visual cues',[25] revealing that they communicate action-theoretically relevant information to users.

[23]Unless the game strongly resembles a familiar game, the entire procedure for playing it must be communicated to rookie players. For playing simple games, a rule book may be effective, but complicated games are learned far more quickly and effectively from other players, by playing a test game or in actual game play.

[24]In the sociology of artefacts inspired by Akrich (1992) and Latour (1992), artefacts are described as texts (*scripts*) that are inscribed in objects by their designer. Users decide how to use the artefact on the basis of their interaction with these scripts: scripts 'define a framework of action together with the [users] and the space in which they are supposed to act' (Akrich 1992, p. 208). Although this definition does not exclude a broader range of communicative means, the emphasis lies on features of the used object alone. If one accepts this narrow reading of communication, then designers pass information to users by the physical properties of artefacts only.

Disregarding some underlying philosophical divergences between our view and that of Akrich and Latour, we regard the latter as too narrow to describe the way in which users learn how to use artefacts. Surely, it holds for some artefacts: such everyday artefacts as coffee percolators and corkscrews have typical physical properties by which their use can be recognised. But it does not hold in general. Consider, for instance, artefacts that are white odourless pastes, such as filler, condensed milk, anti-ageing and medical onguents, etc. Users need the containers of these pastes to tell them apart and to recognise their use. These containers contain explicit information in pictures or printed words about how to use the paste. Or the containers remind users of information (demonstrations) provided in commercials or in shops. This information originates in part from the designers: they have, for instance, communicated to the marketing department how to present the paste. If one considers more complicated artefacts, it becomes even clearer that designers pass information to users through other channels than the physical properties of the artefacts. Take, for instance, a new atomic submarine or the latest MRI scanner. Navy personnel does not simply read off the operating principles from the properties of the submarine. Instead, users are explicitly trained in the use of these artefacts. In (Houkes and Vermaas 2006), we describe the differences in analysing these examples in more detail.

[25]Norman (1990).

2.5 Designing products

Our general characterisation of designing says nothing about the origin of the objects x_1, x_2, ..., manipulated while executing the designed use plan. If these objects already exist, the designer need only point them out to the users as part of his communication of the use plan. If some or all objects do not exist, the designer has to make them available to users in order to believe that the executors of the use plan have recourse to these objects (i.e., in order to satisfy requirement D.11). Object creation can be added as an extra stage to the general characterisation of designing, turning it into a characterisation of product designing. The reconstruction given in table 2.3 shows more precisely how product designing is 'embedded' in plan designing.[26] In the product-identification phase, the task to be fulfilled in product designing proper is determined. This is done partly on the basis of a (conceptually) previous phase of designing, in which the set of objects to be manipulated in executing the use plan is fixed (step D.5 of our general reconstruction of designing). Once it is determined which of these to-be-used objects does or do not yet exist, the activity of product designing can start.[27] To simplify the reconstruction, we assume that only one object is created; let that object be x_n. For determining the success of product designing, the product designer should have physical-support beliefs that determine the desired capacities of the product to be designed, or a use plan, in which the product is manipulated, which can be tested for its effectiveness. The reconstruction of product designing can therefore be seen as an expansion of the physical support step D.6 of our general reconstruction of designing.

Product designing proper is a recursive activity, in which the designer sets himself tasks of describing objects and their components. The goal state g_n defines the initial design task, and an additional subsidiary task g_{nm} is set for each component. The step in which tasks are set, (PD.3), is followed by a series of recursive steps: a decomposition step (PD.4), which involves a configuration of multiple components, all with their own capacities, which together realise the composite object with its desired capacities; a design task setting step (PD.5) for all those components that do not exist either; and a final integration step (PD.6), in which the overall design task defined by g_n is considered to be fulfilled. Since the activity starts with the (justified) belief that the product to be designed does not yet exist, product designing without one or more decomposition steps amounts to discovering that the desired product exists after all. Such product designing would be the limiting case of off-the-shelf engineering. At the other extreme, the recursive decomposition process must end at some point, on pain of infinite regress.

The recursive phase of product designing creates a hierarchy of plans and components. If the designer intends to contribute to realising a goal state, a process of subsidiary planning is started, which ends in considered actions

[26] Houkes, Vermaas, Dorst, and de Vries (2002, §3).

[27] It is possible to produce an object with a set of capacities without presupposing the construction of a use plan, i.e., to make something without having a use for it in mind. We consider such planless designing in sections 6.2 and 6.3.

Product identification:

PD.1. *non-existence*: designer d believes that object x_n with physicochemical capacities $\{\phi_{n,1}, \phi_{n,2}, \ldots\}$, which is one of the objects $\{x_1, x_2, \ldots\}$ manipulated in the course of executing p, does not exist.

PD.2. *design task*: d intends to contribute to realising the goal state g_n, consisting of the existence of a description of an object x_n with physicochemical capacities $\{\phi_{n,1}, \phi_{n,2}, \ldots\}$.

Product designing proper:

PD.3. *task setting*: d intends to describe an object x_n with capacities $\{\phi_{n,1}, \phi_{n,2}, \ldots\}$ for realising g_n.

PD.4. *decomposition*: d believes that a composite of components $\{c_{n1}, c_{n2}, \ldots\}$, where c_{n1} has capacities $\{\phi_{n1,1}, \phi_{n1,2}, \ldots\}$, c_{n2} has capacities $\{\phi_{n2,1}, \phi_{n2,2}, \ldots\}$, etc., has the desired capacities $\{\phi_{n,1}, \phi_{n,2}, \ldots\}$.

PD.5. *component design task*: for each component c_{nm}, d intends to contribute to bringing about the goal states g_{nm}, consisting of the existence of a description of an object c_{nm} with capacities $\{\phi_{nm,1}, \phi_{nm,2}, \ldots\}$. If d believes that this component c_{nm} already exists, then design task g_{nm} is fulfilled by describing this object. If d believes that this component c_{nm} does not exist, then another decomposition step PD.4, component design task step PD.5 and integration step PD.6 is made in order to fulfil the design task g_{nm}.

PD.6. *integration*: d believes that the various design tasks g_{nm} are fulfilled simultaneously, i.e., that the object x_n composed of the described components $\{c_{n1}, c_{n2}, \ldots\}$ has the capacities $\{\phi_{n,1}, \phi_{n,2}, \ldots\}$.

PD.7. d intends to communicate the description of object x_n, possibly along with instructions for production and assembly of it and/or its components, to appropriate agents.

Table 2.3: Reconstruction of product designing

for which no further planning is needed. Such further planning ends when a component with the desired capacities is available; then, no further design task is set that is aimed at producing this component. To integrate all the results of subsidiary design tasks, the product designer needs to believe that the components can be integrated in such a way that the composite object indeed has the desired capacities and that the original use plan can be effectively executed by manipulating, among other things, this newly produced object. These two beliefs amount to bottom-up checks on the fulfilment of the original task of product designing. Finally, this reconstruction assumes that the designer is not the producer of the designed product: product designing proper starts with the intention to describe an object, i.e., to produce a blueprint, not with the intention to produce the object. Hence, some results of product designing have to be communicated to agents who will manufacture the product (PD.7). This final step is added to our reconstruction to make it more realistic: designers nowadays seldom make artefacts.

As an illustration of product designing, consider a nail clipper. By our description, the process of designing a nail clipper does not start with a designer intending to make a nail clipper. Instead the designer intends to help people to realise some goal, with a use plan that may be executed to realise this goal, and with some beliefs about the objects that are manipulated in executing this plan. Suppose that this plan features an appliance that the designer believes is non-existent. Identification of such a product, and its desired capacities, is an example of subsidiary planning. Given the designer's intention to contribute to realising the overall goal state, he has to contribute to realising the existence of this object, which is needed to realise the goal. Hence, product designing of a nail clipper starts. This process results in further design tasks once the product designer decomposes the object to be produced into components that have capacities of their own, such as a unit for applying manual force, a unit for applying force to the nail, units for transferring forces, and a unit for collecting clipped nails. If the designer believes that there are objects available that have the desired capacities, the only task left is to integrate them in such a way that the composite object has the capacities desired of it; otherwise, decomposition leads to further design tasks for each of the non-existent components.

Product designing presupposes plan designing. The design tasks that constitute product designing require a goal that is set in the goal-contribution and plan-construction stages of designing in general. If there is no use plan in which the new object is manipulated, and to which its physical capacities lend support, the object is, so to say, without practical import, and product designing is pointless. As the example of the nail clipper shows, the use plan may determine the choice of components, and thus direct product designing even at this more fine-grained scale. This conceptual priority of plan designing over product designing does not entail a chronological ordering. We do not claim that, in actual designing, there are two clearly distinct stages, the first of which ends in the choice of a use plan, and the second of which takes this plan as fixed input and ends in the description of a new object. Just as the various steps in our reconstruction are, in reality, connected by iteration and feedback loops, actual product designing may not be sharply distinguished from the choice of a use plan. Yet the activities are conceptually distinguishable, and this distinction matters.

Once again, our reconstructions may not describe actual designing. The different steps in our reconstructions are discerned and ordered on the basis of conceptual criteria such as rationality, and ignore the problems that designers may encounter in actual practice: it, for instance, does not incorporate that the knowledge basis of the designer evolves while he is designing, and it does not accommodate the realistic possibility that designers err and temporarily develop a use plan or description of an object that is later on rejected. The different steps in our description are not steps that designers should make in a temporal order. Our characterisations do not translate directly into a design methodology. Still, our reconstruction may be useful for design methodologists. It describes how designers can or should justify (in part) the results of their work, once they are done. Designers may then ignore that they acquire knowledge during the actual process of designing, or that they pursued red herrings, and instead give a rational explanation of their work along the lines of our reconstruction (see, e.g., Ridder, de (2006)). It may, furthermore, be of value to evaluate proposals for design methods. In our reconstruction a clear distinction is made between beliefs about goals and plans to achieve those goals, and beliefs about capacities of objects and the contributions to the actions of the plans. Thus, we keep apart intentional and structural descriptions of artefact using and designing. Design methodologists are prone to mix up these descriptions and describe (product) designing directly as a process in which purposes of agents are translated into structural descriptions of objects. Analysing the temporal steps by which designers are supposed to achieve this translation by means of the distinction that is part of our reconstruction, can therefore shed further light on, for instance, the types of knowledge — scientific knowledge and/or knowledge about the actions of users — designers need for making these steps.[28]

2.6 Standards for use plans

The previous reconstructions of using and designing provide the background for our theory of technical functions, but they also provide an evaluative framework. In this section and the next, we show that use plans provide standards for evaluating use and design in a natural, context-sensitive and conceptually rich way.[29]

The cornerstone of the evaluative framework is that use and design can be assessed on the basis of the quality of the plan that is executed or constructed, relative to the circumstances in which it is executed or constructed. Throughout this book, the relevant quality of the plan is indicated as *(practical) rationality.* This is, to a large extent, a technical term. In ordinary language, plans may

[28]See, for instance, the analyses in (Dorst and Vermaas 2005) and (Vermaas and Dorst 2007) of the design methodology advanced by Gero (1990) and Rosenman and Gero (1998).

[29]Houkes and Vermaas (2004, §§2–3).

carry such normative labels as 'realistic', 'sound', 'feasible', impracticable' or plain 'good', but one seldom hears a plan recommended as rational or discredited as irrational. Our notion of plan rationality is intended to include most members of this family of positive labels, with effectiveness as the core value.[30] Similarly, irrationality is meant to capture the negative labels, centred on 'ineffective' or 'inappropriate'.[31] Plans are, in other words, recommendable if agents are likely to realise the corresponding goal states by means of them, and not recommendable if they make goal-realisation unlikely.[32]

Although effectiveness is a central standard for evaluating plans, applying it is not always easy. For one thing, plans are typically evaluated with regard to their *comparative* effectiveness, where the reference class is determined by the circumstances and skills of an agent. One may deliberate whether to take the car or train to work, unless one does not have a driving licence; methods for curing diseases should improve upon existing methods in terms of cure rate or side effects; and household items are often sold as less time-consuming, costly or waste-producing.

Assessing effectiveness is context-sensitive in more ways. It depends, for instance, on what the agent regards as a satisfactory result. If one just wants a potable cup of tea and has little time or money to spare, the camping-ground plan described in section 2.2 may suffice. If one's standards are higher, this plan may be regarded as ineffective. Effectiveness also depends on the agent's ability to recognise when the goal has been achieved. Many people judge detergents by their capacity to create a fresh-smelling foam, although this capacity is only loosely related to their capacity to dissolve grease. Having the choice between two equally effective detergents, only one of which creates foam, the average consumer would prefer the foam-creating one. This preference may, strictly speaking, be irrational, but there is little point in discrediting it. Hence, in

[30] One might wonder whether effectiveness, or instrumental rationality, is a specification of practical rationality or a definition. See Dreier (2001) for a powerful argument in favour of the definitional horn of this dilemma.

[31] There are many evaluative terms that do not belong to this extended family. Calling a plan 'ruthless' is, for instance, surely evaluative, but this type of evaluation cannot be grouped under the heading of rationality: the ancient Pharaohs' plans to build pyramids by means of slave labour was ruthless, but there is little doubt concerning their effectiveness, whereas Mao Zedong's attempt to reform agricultural China was perhaps even more ruthless and a dismal failure.

[32] This effectiveness standard for plans may be so fundamental that the very notion of 'plan' incorporates a measure of effectiveness. Sailing to the moon by paper airplane, or obtaining a Ph.D. by running around in circles are not naturally described as irrational plans; rather, one would say that these are not plans at all. There must be some method in the madness to speak of plans. We remove these evaluative connotations from the notion of 'plan', and call every goal-directed series of considered actions a plan. We treat, perhaps somewhat artificially, the evaluation of plans as external to them instead of partly intrinsic.

Still, the notion of 'goal-directedness' used in our characterisation of plans may be taken as crypto-normative. It seems to conceal an implicit standard of belief coherence, as introduced later on in the main text: if the agent constructing or executing the series of considered actions believes that there is no chance whatsoever that they result in the goal state, there seems to be no reason to collect these actions in a plan; they are an arbitrary series of actions, which may be intentional individually, but not collectively — the practical equivalent of cut-up prose or experimental poetry.

many cases a plan may be judged effective if the agent reasonably believes that its goal can be realised by executing it and that there is no better (i.e., more effective) alternative.

We leave aside all these complications, and suppose that plans can be evaluated in terms of their effectiveness, most likely in a context-sensitive way. Besides effectiveness, we sometimes employ some other standards for use plans: goal consistency, means-end consistency and especially belief consistency. These have been proposed in the literature as standards for or demands on plans in general.[33] Below, we briefly introduce each of these standards and give some examples of their application to use plans.

Goal consistency

Some plans are directed at multiple goals, such as the plan to post a letter while taking an evening walk. Other plans are intended to realise subsidiary goals besides their central, defining goal. The main goal of driving a car is probably transportation, but one may want to hear the latest news while driving home from work, realising two goals simultaneously; or one may want to drive safely, realising the main goal in a specific manner.

If executing a plan serves more than one goal, the plan must be goal-consistent: the agent executing the plan must reasonably believe that both goals can be realised by executing the plan. If the agent does not have this belief, or if it is not reasonable, the plan is goal-inconsistent.[34] A plan to hear the news while driving home is goal-inconsistent if the driver knows that his car does not have a radio or other audiovisual equipment on board; but if, unknown to the driver, the radio has broken down since the last time that he used his car, planning to use the radio while driving is still rational — until the driver has revised his beliefs in a reasonable manner.

Some authors have pointed out that virtually all plans are embedded in a hierarchical structure.[35] This is especially clear for use plans, and our reconstructions have brought out how a hierarchy of plans may be constructed. Furthermore, artefact use seldom serves only its immediate goal. We drive a car primarily for transportation, but our way of doing so reflects various different, more encompassing goals. Safety, for instance, is not a matter exclusive to private transportation, but often reflects a general concern for one's own health and that of others. These general concerns may also determine which artefact is used, or whether an artefact is used at all. To minimise fuel consumption, one may refrain from using the air-conditioning in one's car; for safety reasons, one may decide against bungee jumping. Executing a specific use plan may be deemed inconsistent on the basis of this hierarchy of plans; the many inconsistencies with respect to sustainable behaviour show how subtle and controversial such evaluations of plans may be.

[33] Bratman (1987, §3.2).

[34] Bratman (1987, p. 31) refers to this standard as internal consistency. We have chosen to call it 'goal consistency' since plans have more internal structure than just their goal-directedness.

[35] Bratman (1987, p. 29); Pollock (1995, p. 233; §7.2).

As a first approximation, we ignore this hierarchical embedding of use plans and focus on the effectiveness of a use plan in bringing about its characteristic goal. We regard safety and sustainability as secondary goals, which serve to select one plan out of a class of equally effective ones. In this way, we avoid the conclusion that safety, sustainability, or — even more generally — leading the good life are the functions of virtually all artefacts.

Means-end consistency

The effectiveness of many use plans depends on the availability of various means. Using a coffee percolator to make coffee is ineffective if one does not have a coffee percolator, but also requires coffee and water. Means-end consistency includes both the availability of the artefact itself and of auxiliary items; our use-plan approach does not distinguish between the two. A plan is means-end consistent if and only if the agent executing it reasonably believes that both the artefact and auxiliary items are available to her.

Means-end consistency partly explains the typical dynamics of making plans. The plan to drive a car cannot be executed successfully if, among other things, the tank is empty, the tire pressure is too low, or the windows are smeared with mud. Hence, a driver has to stop at a gas station, garage or carwash every now and then. Many of these activities come with subplans of their own; Dutch gas stations, for instance, have a rather fixed routine of parking the car at a pump, switching off the engine, etc. that drivers are supposed to follow. In addition, drivers may develop ways of dealing with these auxiliary goals by constructing subplans that realise them. Someone may, for instance, choose to stow a jerrycan filled with gasoline in the trunk or to fill the tank whenever the fuel meter says that it is three quarters empty. These dynamics of deliberation, by which plans lead to subplans, is driven by considerations of rationality: a plan is means-end-inconsistent if the user knows that auxiliary items are unavailable; hence, someone executing the plan should check and ensure the availability of these items, on pain of irrationality.

Belief-consistency

Plans are deliberate efforts to change the world. As such, they are based on beliefs about the world, ourselves, and the effects of the actions. On this belief base of plans, a rather general standard may be based, which has been formulated as, e.g., '[...] it should be possible for my entire plan to be successfully executed given that my beliefs are true.'[36]

We impose a stricter belief-consistency standard than is sometimes found in the literature, because this stricter standard is more appropriate for our theory of technical functions. Suppose that someone uses a Styrofoam cup as a device for communicating with a Mars rover, placing it against his mouth and ear, and talking and listening intently. When interviewed, this person denies that he is involved in pretend-play and asserts that the cup is a communication

[36]Bratman (1987, p. 31).

device when used in the way described. Since the plan is consistent with this person's beliefs, it satisfies the standard of belief-consistency as it is phrased in the previous paragraph. However, use of Styrofoam cups as interplanetary communication devices is so ineffective that it may be discredited as irrational. This can be arranged by requiring, not just that the use plan be based on actual beliefs, but on *justified* or *reasonable* actual beliefs. We do not go as far as requiring *true* beliefs, or knowledge. The reason is that even seemingly flawless deliberating practices are not guarantee to lead to effective plans. Someone may, for instance, justifiably believe that he can cycle from Amsterdam's Central Station to the Van Gogh Museum via the Dam square, only to find that the Dam square is blocked because of (unannounced) construction work. Intuitively, there is little wrong with the process that led to this travelling plan, except that the resulting plan is unsuccessful. We therefore chose to evaluate plans that are based on justified, but false beliefs as rational.

2.7 Evaluating artefact use and design

Since use can be analysed as the execution of a use plan, the standards on use plans can be translated more or less directly into an assessment of artefact use. In comparison, deriving standards for design is more complicated, but also potentially more rewarding. We now discuss the evaluative frameworks for use and design in turn.

Rational and proper use

As detailed above, plans can be evaluated in terms of their rationality. Therefore, if use can be described as the execution of a use plan, artefact use is rational if and only if it is the execution of a rational use plan; otherwise, it is irrational. Standing on a chair to change a light bulb is rational, provided, among other things, that one justifiably believes that the chair can hold one's weight (i.e., that the executed use plan is effective) and that there are no artefacts available that are more appropriate to stand on (i.e., that the executed use plan is efficient). Thus, the beliefs and specific circumstances of individual users, e.g., concerning their weight, determine whether artefact use, considered as a particular event, is rational or irrational.

This distinction between rational and irrational use is not the same as that between proper and improper use. Use is often assessed in the latter terms. As stated in the introductory chapter, for instance, the warranties of many artefacts contain a void clause for cases of improper use. This distinction between proper and improper use can also be analysed in terms of use plans: artefact use is proper if and only if it is the execution of a use plan that is deemed acceptable within a certain community. This acceptance or privileging of a use plan may have nothing to do with rationality. There are reasons to suppose (and hope) that rational and proper use overlap. But, as a first approximation, we treat the notion of 'improper use' as an expression of social disapproval, signalling the execution of a use plan different from the one that is socially expected or proper under the circumstances. For determining this, the difference conditions listed in section 2.2 apply.

Some examples may clarify this descriptive notion of improper use. Bed sheets are a useful and stereotypical tool to escape from a prison cell. In fact, using sheets may be the most appropriate plan in a prisoner's circumstances, since janitors go to great lengths to ensure that prisoners have no means to escape from their cells. Hence, using a string of sheets to escape from prison may be rational. Still, it is an example of improper use, for the simple reason that escaping from prison is a crime. If the prisoner would use a ladder made of diligently collected drinking straws, this would be equally improper, but also irrational. Conversely, using a car for private transportation is proper, since it is socially accepted (at least if the driver is sober); using it to drive three blocks in rush-hour traffic is irrational (at least if the driver can walk), yet proper.

For the moment, this descriptive reading of the proper-improper distinction suffices; we forestall further discussion until section 4.3 and chapter 5, when developing a more detailed understanding of the distinction will have turned out to be important to our theory of artefact functions.

Good designing

On the basis of our use-plan analysis of designing, we can also present an evaluative framework for this activity. The reconstruction presented in section 2.4 brings out the plan executed in designing, i.e., the series of considered designing actions, each with their belief base. Hence, the standards for plans, presented in section 2.6, may be applied to designing in the same way as they are applied to using. Like using, designing is rational if and only if it executes a design plan that is effective, goal-consistent, means-end consistent and belief-consistent. We discuss each of these standards in turn, to see how they can be used to assess designing.

First, effectiveness. Since the goal of this design plan is to contribute to the goal realisation of other agents, the effectiveness of designing is determined by the extent to which the designer succeeds in contributing. For all types of designing, the contribution made to goal realisation consists of a use plan, execution of which is supposed to lead to realising the goal in question. Therefore, effectiveness of the design process requires effectiveness of the use plan that is constructed in it. If a designer constructs an irrational use plan and does not aid users in realising their goals, he has not achieved his own goal and designing is irrational.

This connection between the effectiveness of the use plan and the rationality of designing is the backbone of our evaluative framework for designing. Plan effectiveness is not, however, sufficient for good designing. To succeed in aiding users, the designer should also effectively communicate the effective plan, including conditions for its effectiveness, to the users. Moreover, if the designer adjusts the goal state to be realised by executing the use plan, this adjustment ought to be appropriate: helping users to realise their goal of making toast by providing them with an effective plan for making tea is bad designing despite the effectiveness of the constructed use plan. Hence, effectiveness affects all three stages of designing that we distinguish in our reconstruction, as well as the additional stages introduced in product designing. This emphasis on effectiveness leaves open the possibility that (product) designers ground their work in models that are known, also by the designers themselves, to be incorrect. As long as their belief in the effectiveness of the communicated use plan is justified, e.g., on the basis of extensive tests, such designing can be evaluated as rational. This provides an additional reason for requiring that the beliefs that feature in planning are justified, but not necessarily true.

Let us now turn to each of the subsidiary standards. There are various ways in which designing can fail to be goal-consistent, which should not be confused with the designer's belief that *executing* the plan is goal consistent for the users. A designer could wish to aid two groups of prospective users who have slightly different goals. Then, in aiding one group, the designer may fail to aid the other; or even worse: in constructing a plan or product to aid both, he may succeed in aiding neither. Consider, for instance, hybrid artefacts, which are designed to be used for various purposes, typically by making some slight modifications such as replacing components. Food processors are a relatively successful example, but there are many less successful cases.[37] Designing can also be goal inconsistent if the activity is directed towards some goals besides the central goal of aiding the user. Although we regard goal contribution as the central aim of designing, actual designing is constrained by many other considerations, such as cost effectiveness, time constraints, the desire to patent new products before competitors do, etc. These additional goals may lead to goal-inconsistency if they interfere with the aim of providing users with effective ways of realising their goals.

Means-end consistency provides designing with an internal dynamics that strongly resembles the dynamics of using; thus, a very brief indication of its application to designing may suffice. Just as using an artefact involves searching necessary auxiliaries and planning subsidiary actions, rational designing requires one to seek out the means for taking all actions included in the design plan and to make subsidiary plans for using these means. For instance, means-end consistency requires the designer to obtain appropriate means for communicating the use plan to prospective users.

Finally, the design plan has a broad belief base, which overlaps only in part with that for use plans. This means that the standard of belief consistency has a broad impact on designing. The designer should, for instance, believe that he has the skills needed to aid prospective users in realising their goals by means of the use plan. Although an architect could conceivably design a complicated electrical appliance, he is primarily equipped to design buildings. Other parts

[37]One example of a failed hybrid artefact is the Piccolo, a multifunctional household appliance manufactured in the 1950s. A predecessor to today's food processors *and* vacuum cleaners, the Piccolo could also be used to spray paint and pesticides, grind coffee beans, drill small holes and clean carpets. An equally unsuccessful rival device, the ominously named Kobold, could be used as both a vacuum cleaner and hair dryer, with predictably unfortunate results.

of the belief base, made explicit in the parenthesised 'belief steps' in our reconstruction, concern the constructed use plan, the objects manipulated in it, and the circumstances in which the plan is executed. To acquire these beliefs, the designer has to put himself into the users' shoes, in order to imagine what would be the best thing to do in the users' situation. This requires information about the typical physical circumstances in which users seek to realise certain goals, about the skills that they can appeal to, about the availability of alternative means, etc. Hence, by way of belief consistency rational design requires justifiable beliefs about, among other things, the users' skills, circumstances and available artefacts, just as rational use does.

Even with regard to those beliefs that concern the plan and its execution, belief consistency is more demanding for designing than it is for using. Unlike the user, the designer has to consider all prospective executings and executors of the use plan that he has constructed. Since these potential users may, for instance, have different skills and resources, aiding one group might actually decrease the chances of aiding the other group, meaning that the designer cannot consistently have skill-compatibility beliefs regarding both groups. If a designer wants to make his design as *inclusive* or *universal* as possible,[38] belief consistency makes good designing a considerable challenge.

[38] Inclusive design, which is also called 'universal design' or 'design for all', is an increasingly important topic for product designers, who are faced with a rapidly aging population in most of their existing markets. See Goldsmith (1997), Clarkson, Coleman, Keates, and Lebbon (2003) and Clarkson and Keates (2003) for some introductions to problems and approaches within this field.

Chapter 3

Function theories

In the previous chapter we analysed the using and designing of artefacts in terms of the goals, beliefs and actions of the agents involved. In this chapter we switch to a perspective on artefact teleology that is more common in philosophy and engineering. This object-oriented perspective brings us to our central project of developing an adequate theory of technical functions.

In this chapter, we start the project by reviewing the various function theories for artefacts that are currently discussed in the literature. This review serves two purposes. First, it establishes that existing theories about functional descriptions of artefacts are not suitable for the aims we have in mind — we argue that the currently available function theories do not meet the four desiderata that we gave in chapter 1. Second, the review is instrumental to developing our own function theory, which does *meet the desiderata. We organise the review by considering three basic archetypical function theories — intentional, causal-role and evolutionist function theories — that can be taken as spanning up the spectrum of currently existing function theories. Our proposal, which we present in the next chapter, is then constructed employing the useful material in those three basic theories.*

W. Houkes, P.E. Vermaas, *Technical Functions*, Philosophy of Engineering and Technology 1, DOI 10.1007/978-90-481-3900-2_3,

3.1 Function theories for technical artefacts

A philosophy of artefacts that remains silent on technical functions would be incomplete. In the preceding chapter we have analysed artefact using and designing as actions, connecting them to the goals, beliefs and actions of agents. By making this connection, we offered an agent-centred, action-theoretical perspective on artefact using and designing. But a more object-oriented perspective is also possible: using can be described as the employment of objects for performing specific functions; designing can be taken as the selection of objects that can perform these functions. When considering artefact use this perspective might be of secondary value: agents who use artefacts do so primarily to obtain goals, and do not typically describe this use in terms of the functions of these artefacts (Houkes and Vermaas 2004). We therefore chose not to include the concept of function in our action-theoretical reconstruction of using. However, a similar argument for designing seems untenable. In design methodology, artefact functions are typically mentioned explicitly: reasoning about functional descriptions is an important part of engineering designing,[1] and some methodologists, such as Gero (1990) and Roozenburg and Eekels (1995), even define designing as an activity in which required functions are transformed into (structural) descriptions of objects that can perform these functions. It thus seems reasonable to assume that, unlike users, designers describe their own activities partly in terms of functions.[2]

Another reason for turning our attention to technical functions is that there is an overall tendency in philosophy to consider technical artefacts as primarily, even essentially, objects with functions. On this orthodox functional approach, it becomes a central task for philosophers who consider artefacts to give a theory of functional descriptions. As is clear from our introductory chapter, we take this task seriously, and thus turn our attention from an agent-oriented description, in terms of actions, to an object-oriented description, in terms of functions.

There are a number of proposals for *function theories* on offer. But despite the central role that some philosophers grant functions in the understanding of artefacts, the current debate on functions is only in a subsidiary way concerned with artefact functions. If this debate is taken as one that originated in the work of Hempel and Nagel on functional explanations and that transformed via authors such as Ruse, Cummins and Wright partly into a debate about the concept of function itself,[3] then most contributions focus on functions in the biological domain. Social functions, functions in psychology and, increasingly, functions in linguistics also get their fair share of attention, but technical functions remain far from the limelight. If they are mentioned and discussed in the

[1] The field of engineering design methodology is one in which consensus is lacking and possibly even a field in which consensus is not aimed at (Vermaas 2009a). Yet, in many of the methodologies proposed for designing, functional reasoning is a key element. See, for instance, (Erden et al. 2008) for a survey of the modelling of functions in 18 different methodologies.

[2] We have not yet included explicit references to functions in our reconstruction of designing. In the next chapter we, however, identify functions of artefacts with the capacities that are mentioned in step D.7 of this reconstruction (see table 2.2 on page 29).

[3] See, e.g., McLaughlin (2001, part II).

literature, it is usually in passing. They are dealt with quickly and taken as an unproblematic, philosophically insignificant case.[4] Technical functions may, at most, act as a standard against which other functional descriptions can be measured — the contrast between functional descriptions in technology and in biology may, for instance, be used to show what is problematic about the latter.

We have no quarrel with this attention for functional descriptions outside of the domain of artefacts. Given the teleological overtones of functional descriptions of items, philosophers of science are understandably keen on explaining the use of functional descriptions in biology. Functional descriptions in cognitive science and the social sciences have become increasingly important in the philosophy of mind as a possible route towards naturalising intentionality. The analysis of technical functions is not connected to such grand philosophical programs or to familiar conceptual problems in well-studied academic disciplines. 'Function' is, however, a central concept in common-sense and engineering descriptions of artefacts.

Because of the uneven spread of attention for various functional discourses, the intricacies of functional descriptions in the technical domain are ignored. Authors typically develop a theory for biological or social functions, and present their proposals as general theories applicable to artefacts as well, but are not too specific about how their theories apply to those artefacts. This may result in proposals that are well-formulated, sophisticated and distinct in their application to the biological domain, but vague, naive and indiscriminate once they are applied to the technical domain. Searle (1995), for instance, gives a characterisation of the ability of agents to assign functions to items as part of his construction of social reality, and also extends this analysis to artefacts. But, as we show in the next section, there are at least two distinct ways of understanding this application to function assignments to artefacts. Bigelow and Pargetter (1987), to give another example, have presented a forward-looking propensity function theory, which in the domain of biology is clearly contrasted to the backward-looking etiological function theory of, say, Neander (1991a, 1991b). This contrast, however, fades away in the technical domain: there, these authors share a naive 'designer and user intentions determine functions'-position. There are, of course, exceptions to this rule of neglect. Cummins' definition of functions, in his seminal (1975) paper, is straightforwardly applicable to especially components of artefacts. And Preston (1998a, 1998b, 2000) is one of the few authors who explicitly and extensively publishes on artefact functions.

In this chapter we review the function theories currently discussed in the philosophical literature. We indicate which proposals for analysing functional descriptions of technical artefacts are already available, and we argue that these proposals are not adequate to the domain of technical artefacts: the theories cannot accommodate the phenomenology of artefacts because they do not satisfy the *proper-accidental*, *malfunctioning*, *support* and *innovation* desiderata (see table 1.1 on page 5). In our critical analysis, we also seek to isolate useful elements of existing function theories. These elements are used in the next

[4]See, e.g., McLaughlin (2001, p. 61) and Perlman (2004, pp. 31–46).

chapter to develop an alternative function theory, which *does* meet the desiderata. This alternative is built on our analysis of artefacts in terms of use plans, demonstrating the versatility of our use-plan approach in dealing with a number of subtleties particular to functional descriptions of technical artefacts.

Our review

Before embarking on our review, some remarks about its organisation may prove helpful.

First, the sketchiness of most existing theories of technical functions both allows and forces us to organise our review of function theories, not as a sequential assessment of refined proposals, but as a more systematic and global one. We define three basic function theories that can be taken as archetypical to currently existing theories, and then assess the whole spectrum of function theories that can be construed by means of these three basic theories. These archetypes are the *intentional*, the *causal-role* and the *evolutionist* function theories. Each of these three theories fails to meet at least one of the desiderata of malfunctioning, support and innovation, as we shall show. This result may lead one to think that the basic theories might repair each other's shortcomings, suggesting that a combination of basic theories might fare better than each individual theory. Yet an analysis of relevant senses of 'combining', in section 3.5, proves this wrong.

We have chosen our basic theories partly for the resemblance that they, and their combinations, bear to existing function theories. Our template criticisms are therefore meant to apply to all theories, existent or non-existent, that are elements of the spectrum spanned by our three basic theories. This type of criticism should not be taken as an unconditional and specific critique of existing theories. Two of the three basic theories merely *resemble* existing proposals. Some proposals may therefore fall outside the spectrum covered in this chapter, or may be (re-)interpreted as such. Our approach is designed to explore and illustrate systematically the difficulties faced by function theories in the technical domain; it is not designed for polemic purposes. Therefore, we avoid lenghty interpretative analyses of the existing proposals; given the general lack of well-articulated formulations noted above, a definite assessment of existing theories is nearly impossible anyhow. We do consider a number of the more known proposals for function theories, but the discussion will be sketchy. Readers interested primarily in the development of our philosophy of artefacts may consider reading only the first parts of the next three sections, in order to get acquainted with the basic function theories, and skip the remainder of this chapter.

Second, throughout our review, we emphasise a distinction between two types of function theories. The first type characterises functions as *properties* that items may have, independently of the beliefs and actions of agents; these theories are analyses of the claim 'x has the function to ϕ'. The other type analyses the rational or justifiable *ascription* of functions to items by agents, depending on their beliefs and actions; these theories address the claim 'agent

a (justifiably) ascribes to x the function to ϕ'. In the literature, this distinction is left implicit. The reason may be that there is little sense in making it for analysing the biological domain: it seems natural or even necessary to analyse functional descriptions along the first, agent-independent line. For artefacts, this is not a given, since functional descriptions of artefacts seem more intimately connected to the beliefs and actions of agents.

Third, we choose to call our third basic theory the *evolutionist* function theory rather than the *etiological* theory. Our evolutionist theory is, admittedly, similar to Millikan's (1984, 1993) etiological theory of direct proper functions and Neander's (1991a, 1991b) etiological theory of biological functions. But both authors have added elements to their theories that cannot be called evolutionary in our sense: Millikan's derived proper functions, and Neander's technical functions. We therefore introduced our 'evolutionist' label in order to characterise part of Millikan's etiological theory as an evolutionist theory and to reveal that Neander's etiological theory on the technical domain is rather an *intentional* function theory.[5] Our review of existing functions theories thus differs from the usual ones. It identifies problems and distinctions that are particular to the technical domain, and it leads to a different categorisation of existing function theories than the one that is generated if one considers these theories from a biological perspective.

A final remark is that we assess the spectrum of function theories spanned by the three basic theories by considering whether they meet the malfunctioning, support and innovation desiderata only. This test is already sufficient for showing that these function theories cannot be taken as theories of artefacts that satisfy all our desiderata. It would therefore be redundant to argue that some of these function theories do not meet the proper-accidental desideratum either. Moreover, and this brings us to our second aim, the review is meant to collect the means for developing our own proposal to a function theory. As noted above, we need a function theory that, against the background of our use-plan approach, leads to a theory of artefacts that does meet all desiderata. The use-plan approach itself already provides the means to meet the proper-accidental desideratum, in the distinction between proper and rational use (see section 2.7): proper functions can in any theory be taken as those functions the theory describes when artefacts are considered in the context of proper use; accidental functions can be taken as the remaining functions described by that theory (we take this route in section 4.3, when arguing that our proposal is meeting the proper-accidental desideratum). We are therefore now mainly interested in function theories that can handle the malfunctioning, support and innovation desiderata.

[5] In their taxonomy of biological functions Walsh and Ariew (1996) also distinguish evolutionary and etiological theories, but take the latter as special cases of the former. Walsh and Ariew's evolutionary theories include, for instance, the propensity theory by Bigelow and Pargetter (1987). Our focus on technical functions leads us in a different direction. The connection between etiological theories and evolutionary theory becomes less clear on the technical domain. And the propensity theory by Bigelow and Pargetter seems to be an intentional function theory when applied to artefacts.

3.2 The intentional function theory

Our first basic theory is called the *intentional function theory* or, more briefly, the *I-theory*. On this theory, the intentions, beliefs and actions of agents determine the functional descriptions of artefacts. One way to specify this is as follows: an agent designs or constructs an artefact with a specific goal g or capacity 'to ϕ' in mind, and by those actions the artefact may be functionally described as an item 'for realising g' or for 'ϕ-ing', respectively. Alternatively, the fact that an agent uses the artefact for realising g or for ϕ-ing, or merely takes the artefact as something that can be used for these reasons, may be taken as sufficient ground for describing the artefact as an item 'for realising g' or for 'ϕ-ing', respectively. Formulating this in slightly more abstract terms, one can say that, in the I-theory, the agent ascribes a function to an artefact by embedding it in a system of means and ends — by designing or using it, or simply by reflecting on it. That is, the agent takes the item to be contributing to goals, and describes the item itself and its constituents functionally on the basis of these goals. This process of 'means-end embedding' can be interested or disinterested: the agent can ascribe 'I-functions' relative to his own goals or relative to goals in which he takes no immediate practical interest.

To keep our characterisation of this basic theory as broad as possible, we deliberately leave it open whether, on it, the functions ascribed to artefacts correspond to intended goals g, i.e., to the states of affairs that are supposed to be obtained by manipulating the artefacts, or to capacities 'to ϕ', i.e., to the physicochemical dispositions of the artefacts by which they contribute to realising these goals. Various positions seem possible. As we show below, Neander's (1991a, 1991b) theory can be taken as an I-theory in which functions are purposes, the theory of Bigelow and Pargetter (1987) can be taken as one in which functions correspond to capacities, whereas McLaughlin (2001, p. 52) rejects the distinction. Moreover, it is difficult to determine the position of authors who do not explicitly take one, since most formulations and examples conceal ambiguities. If, for instance, a light bulb is ascribed the function of illumination, one can interpret this as the ascription of an intended state of affairs — a state in which a sufficient amount of light falls on objects close to the bulb — or as the ascription of a capacity — the bulb's capacity to emit light.

Another issue that is left implicit concerns the distinction between theories that treat functions as *properties* of items and theories about the *ascription* of functions to those items by agents. The intentional function theory seems primarily of the latter type, since it analyses functional descriptions of artefacts in terms of the actions and beliefs of agents. Nevertheless, the terminology of some existing function theories that resemble the I-theory is suggestive of the functions-as-properties approach; Bigelow and Pargetter (1987) is again an illustration. To encompass all existing proposals, we take the I-theory as one that can, at least in principle, be of both types.

This being said, the intentional theory has some marked advantages in the technical domain. Most importantly, it stays true to the intuitive connection between intentional actions and function ascriptions to artefacts. Furthermore, it has a broad scope. It applies to items that are in working order, but also to malfunctioning items: after all, one may regard something as a means to an end even if it cannot serve as such. The I-theory also allows for a variety of function ascriptions to the same artefact in different contexts: agents can embed any item in various means-end systems. Finally, I-functions may be ascribed to both traditional and highly innovative artefacts. Summing this up in terms of the three remaining desiderata, it takes little effort to see that the intentional function theory satisfies the desiderata of malfunctioning and innovation.

However, looking at the I-theory from this perspective also brings to light that it fails to meet the desideratum of support. If beliefs about an artefact are by definition sufficient for embedding the artefact in any means-end system, the intentional function theory does not require any support for a function ascription. As such, the theory allows all kinds of intuitively incorrect function ascriptions. The ascription of the function of underground travelling to an ordinary car is perfectly acceptable as an I-function ascription, as soon as a designer or user takes the car as a means to this end; so are all other function ascriptions in science fiction literature[6] and crackpot technology. More generally, the I-theory has such a broad scope because it suffers from a proliferation-of-functions-problem. In this theory, function ascriptions require no support at all, and all kinds of intuitively secondary motives are the basis for additional function ascriptions. When, say, a chemical engineer of some company designs a new detergent, lots of goals other than removing stains from laundry may play a role: the directors of the company may want to get a share of the detergent market and break the monopoly of a rival firm, or the engineer may have the goal of earning enough money to pay his or her landlord. These desires are sufficient to embed the detergent in variant means-end systems; relative to these systems, the detergent may be ascribed the function to weaken the market position of the competitor company and to secure the engineer's housing. As it stands, the I-theory does not have the resources to distinguish such additional roles played by an artefact from its technical functions.

Existing intentional function theories

The intuition that technical functions are connected to intentions, beliefs and actions can be found throughout the literature and is incorporated in existing function theories. Therefore, numerous proposals resemble the intentional function theory to some degree. We do not list them all, but present four examples, both to show how our global review applies to specific existing theories and to illustrate some of the distinctions we have made.

There is only one current theory, Neander's (1991a, 1991b), which really fits the definition of the I-theory. Other theories merely resemble it, and also contain elements of the two other basic theories, defined below. The theory by

[6] A beautiful example is provided by a machine, described by Philip K. Dick (1957), which produces coffee and other beverages through a mechanism based on miraculous multiplication of substances — which happens to work in the particular imaginary world described. If there is no such thing as divine intervention, or if it cannot be employed to produce coffee, one would probably hesitate to describe the machine as one that has the function to make coffee.

Bigelow and Pargetter (1987), for instance, adds evolutionist elements to a predominantly intentional approach towards technical functions; and McLaughlin (2001) adds elements that resemble the causal-role function theory. Other theories fit the theory on some interpretations, but resemble another basic theory if interpreted differently: Searle's (1995) theory, which lacks a well-articulated application to the technical domain, provides an example.

Neander

Neander is well known for her seminal work on biological functions, but her etiological theory is in fact far more general. It is based on a single, general definition, which says that 'the proper function of a trait is to do whatever it was selected for'.[7] When Neander specifies this definition for biological items, the term 'selection' refers to natural selection as defined in neo-Darwinian evolutionary theory.[8] The resulting theory falls under the heading of an evolutionist function theory, as defined in section 3.4. When Neander turns to the technical domain, the connection to evolutionary theory is lost. Instead of defining functions of artefacts relative to some evolutionary account for artefacts — a possibility that she briefly notes[9] — she chooses to let 'selection' refer to intentional selection by agents. This process does not operate on existing 'lineages' of reproduced items, slowly favouring one of those lineages over the others; instead, it is a one-shot process that operates on the level of individual artefacts over short periods of time. In this way, Neander arrives at characterisations of technical functions such as:[10]

> I suggest that the function of an artifact is the purpose or end for which it was designed, made, or (minimally) put in place or retained by an agent [...]
> [u]nique inventions, like the additions to James Bond's brief case, can have proper functions peculiar to them because they can be individually selected for particular effects. [...] It is enough, in the case of intentional selection, if the designer believes or hopes that the artifact will have the desired effect and selects it for that purpose.

For Neander, (proper) functions of an artefact thus correspond to goals and can be ascribed on the basis of beliefs of designers and other agents who are involved with the artefact. This immediately leads to the proliferation problem that plagues purebred I-theories: if an engineer hopes that a car that he has designed will be able to move below ground level, or if an agent bought a vehicle for that purpose and keeps it in his garage, they may ascribe to that vehicle the function of subterranean transportation.

[7] Neander (1991a, p. 173).
[8] Neander's definition of proper functions of biological items can be found in (1991a, p. 174).
[9] Neander (1991b, n. 11).
[10] Neander (1991b, p. 462).

Bigelow and Pargetter

In the forward-looking propensity function theory of Bigelow and Pargetter, an item 'has a certain function when it has a propensity for selection in virtue of that [item]'s having the relevant effects'.[11] This is presented as a general, overarching characterisation of functions, which is developed in slightly different ways for the biological and technical domains — two domains that are considered explicitly in the paper. In the biological domain, propensities are conferred by actual capacities of the item to which the function is ascribed, and the relevant propensity is survival-enhancing,[12] i.e., the function confers a propensity to be favoured by natural selection in a specific environment. For the technical domain, propensities are not conferred by actual capacities of an item, but by *representations* of capacities. Presumably, this means that technical items have functions when their representations have relevant effects, and thereby confer to the artefact of which the item is a part a propensity for being selected. These representations may be involved in the process of designing the artefacts, say, by engineers who have explicit knowledge about the capacities for which they reproduce artefacts. To account for items that are reproduced by artisans, who may be largely ignorant about the relevant capacities,[13] Bigelow and Pargetter also allow for representations of capacities at the time of selection.[14]

This propensity theory for technical functions is sketchily presented in a mere paragraph. However, its emphasis on representations makes it a near-perfect fit for the intentional function theory. If an agent believes that an item that is part of an artefact has a capacity to ϕ, i.e., represents it as having this capacity, and therefore selects the artefact, the item has a function to ϕ. That these capacities may be represented at other moments than the first design of the first prototype widens the scope of this function theory even further. Taken literally, Bigelow and Pargetter claim that everyone selecting the artefact for some purpose, because of some representation of its capacities, ascribes a function to it.[15] Since the relevant notion of 'selection' is indeterminate and may, like Neander's notion, comprise maintenance or merely being put in place, the propensity theory is vulnerable to our template criticism of the I-theory. When an agent represents an odd-looking vehicle as having the capacity of subterranean transportation and stores it in a garage, or even exhibits it in a museum because of this representation/belief, the vehicle has a function of subterranean transportation, which lacks all support.[16]

[11] Bigelow and Pargetter (1987, p. 194).

[12] Bigelow and Pargetter (1987, p. 192).

[13] Bigelow and Pargetter give the example of copying hammers: artisans know that hammers work well for banging in nails, but need not be aware that the copied shape of the hammer balances the tool (1987, pp. 185–186).

[14] Bigelow and Pargetter (1987, p. 194).

[15] One may think that in de-emphasising the designer's representations, Bigelow and Pargetter shift away from an intentional function theory to an evolutionist function theory of the type presented in section 3.4. However, unlike evolutionist theories, the propensity theory of technical functions ascribes no role to the physical make-up of the item or its ancestors: the representations do all the work.

[16] One may respond that, like for the biological domain, Bigelow and Pargetter (1987, p. 194) only claim that *a* propensity theory offers the most promising theory of technical functions, and not that the sketch offered in their paper is *the* correct propensity theory. Indeed, our

McLaughlin

Another function theory that strongly resembles the I-theory is formulated by McLaughlin (2001, ch. 3).[17]He writes that:[18]

> In the case of the functions of whole artifacts the determination of their functions or purpose is completely external. It lies in the actual intentions of the designer, manufacturer, user, etc., however socially determined these intentions may in fact be. Such functions or purposes can be changed by a change of mind, and we can use the terms *purpose* and *function* interchangeably.

For McLaughlin, these actual intentions of agents are not only necessary[19] but also nearly sufficient for conferring functions. Designers or manufacturers need not have created the physical structure of the artefact, and users need not have evidence that the artefact can indeed perform the function. 'Virtual design', meaning that the designer decides to leave the artefact as it is, or 'virtual intervention', meaning that the user decides to (eventually) use the artefact, is enough. The only constraint McLaughlin mentions is that this virtual design or intervention is 'realistically possible'. An agent 'cannot convert the Alps into an artificial ski slope simply by "welcoming" the collision of the two subcontinental plates and being willing if necessary to give them a push'. And preferring a log to have fallen across a creek is insufficient to ascribe to it the function of being a bridge if these preferences have no conceivable effect on the log.[20] This constraint of realistic possibility may be seen as adding a structural twist to McLaughlin's theory. However, the theory merely puts this constraint on the function-conferring process, which might still be one of virtual design or virtual intervention. Unlike the causal-role function theory introduced in the next section, McLaughlin's theory does not require support for the claim that the artefact can perform the function. For this reason, the theory is vulnerable to our template criticism: an agent can still ascribe the function of subterranean transportation to an ordinary car by a mere act of will, and detergents still acquire the undesired function of paying next month's rent of the designer's apartment.

---

own proposal for a theory of technical functions shares the emphasis on beliefs and/or representations with the propensity theory. However, the intricacies of our own theory, presented in chapter 4, show that merely pointing out that artefact selection 'clearly involves representations' (Bigelow and Pargetter 1987, p. 194) is still a long way from offering an adequate theory of technical functions.

[17]We only consider McLaughlin's function theory for artefacts as a whole. His description of the functions of *components* of artefacts puts less emphasis on the intentions of agents, but introduces structural elements as part of the causal-role function theory defined in the next section.

[18]McLaughlin (2001, p. 52; original emphasis).

[19]McLaughlin formulates 'Sorabji's Rule', after Sorabji (1964), to express the necessity of agent's intentions: 'Function conferring must involve some act of the will and the intellect, or a pro-attitude and a belief [...] [I]t would seem that some valuational component, however minimal, is always involved in conferring an artifactual function on something.' (McLaughlin 2001, p. 45).

[20]McLaughlin (2001, pp. 45–46).

These three examples illustrate how pronounced differences between function theories in the biological domain disappear when the theories are applied to the technical domain. In Neander's etiological theory, biological functions of an item refer to *capacities of the ancestors* of the item that contributed to the natural selection of those ancestors; Bigelow and Pargetter contrast their theory with etiological theories such as Neander's and let biological functions of an item refer to *propensities of the item* for being favoured by natural selection; and McLaughlin[21] rejects the role of natural selection in these two theories and lets functions refer to those capacities of the item that contribute to the *self-reproduction* of the item's encompassing organism. But in their treatment of technical functions, these rivals team up in an intentional approach. These three theories also cover the different positions on the issue of whether functions correspond to goals or capacities. And the formulations that, in particular, Bigelow and Pargetter choose show that a function theory that comes close to the I-theory can still be a theory about functions as properties.[22]

Searle

Our final example, Searle's[23]function theory, illustrates the difficulties one may encounter when reviewing the application of existing function theories to artefacts. Searle's theory covers biological, technical and social function ascriptions and takes functions not as intrinsic to the functional items but as features that agents ascribe relative to values they impose on the items concerned. For artefacts these values originate in the practical interests of agents, and for biological items they originate in a theory. Searle does not give an explicit definition but gives two, what he calls, 'central conditions' for making the statement that 'the function of X is to Y':[24]

> 1. Whenever the function of X is to Y, X and Y are parts of a *system* where the system is in part defined by *purposes*, *goals*, and *values* generally. [...]
> 2. Whenever the function of X is to Y, then X is *supposed to* cause or otherwise result in Y. This normative component in functions cannot be reduced to causation alone, to what in fact happens as a result of X, because X can have the function of Y-ing even in cases where X fails to bring about Y all or even most of the time. Thus the function of safety valves is to prevent explosions, and this is true even for valves that are so badly made that they in fact fail to prevent explosions, i.e., they *malfunction*.

The system introduced in the first condition should be taken as a system in a broad sense of the word. It contains X but is not just some composite material

[21] McLaughlin (2001, ch. 8).

[22] Bigelow and Pargetter capture their general function theory in a single sentence: 'So a character or structure *has a certain function* when it has a propensity for selection in virtue of that character or structure's having the relevant effects.' (1987, p. 194; emphasis added).

[23] Searle (1995, pp. 13–23, 38–51 and 122–124).

[24] Searle (1995, p. 19; original emphasis).

system since it also contains the function Y of X, and is in part defined teleologically; when X is, say, a chair, the system referred to is one aimed at supporting seated people and containing the chair and its function of supporting buttocks and backs of humans.

Moreover, Searle categorises functions of artefacts as 'causal agentive functions', meaning that they are ascribed relative to the practical interests of conscious agents, and that an artefact performs them solely in virtue of the object's intrinsic physical features.[25] It is unclear how this second clause should be combined with the two conditions given by Searle. One may take it as an additional condition, which then could have the form:

> 3. Whenever the function of an artefact X is to Y, then X can cause or otherwise result in Y in virtue of X's physicochemical structure.

But this way of understanding Searle's function theory for artefacts seems infelicitous: such a third condition would obliterate the normative aspect of function ascriptions, which Searle incorporates in his second condition. Consider, for instance, one of the badly made safety valves mentioned in the quote. When such a valve satisfies the third condition, it can prevent explosions in virtue of its physicochemical structure. This means that it can *always* prevent explosions and thus rules out the possibility of temporal or permanent malfunctioning. An alternative way to combine the clause with the two conditions is to incorporate it in the second one. This then may take the form:

> 2*. Whenever the function of X is to Y, then X is supposed to cause or otherwise result in Y in virtue of X's physicochemical structure.

Again, it is doubtful whether this interpretation captures what Searle has in mind. The reason is that this alternative second condition considerably weakens the connection between the performance of the function Y and the physicochemical structure of X. On it, merely *supposing* that an artefact has the right physicochemical structure to perform this function is sufficient to ascribe a function to it. If such a supposition need not be justified or made plausible, then it hardly implies that X can perform Y solely in virtue of its physical make up.

Further analysis and interpretation may resolve this issue of combining Searle's central conditions with his definition of causal agentive functions. Our point here is that this analysis is needed; a first reading of Searle's texts about function ascriptions to artefacts leaves room for at least two different, equally inadequate ways of understanding his theory.[26] On the first understanding, the theory is an instance of the causal-role function theory (see the next section)

[25] Searle (1995, pp. 38–43 and 123–124). Functions of artefacts may on Searle's categorisation also be 'status functions', i.e., functions that are ascribed relative to practical interests of agents, which artefacts can perform only by way of collective acceptance by those agents. Such social functions fall outside the scope of this book.

[26] See also Kroes (2003) for a (critical) analysis of how to understand Searle's theory of artefact functions.

and fails to satisfy the malfunctioning desideratum. On the second, it is an intentional function theory — functions are ascribed on the basis of suppositions and goals of agents — that fails the support desideratum.

3.3 Cummins' causal-role theory of functions

The second basic function theory is immediately imported from the literature: it is the well-known function theory of Robert Cummins (1975). Here, we call it the *causal-role function theory* or, for short, *C-theory*. It captures another strong intuition about functions, namely that functions of items are related to the causal roles these items have in larger composite systems.

According to Cummins function ascriptions arise in the context of explanations of capacities of systems.[27] He distinguishes two strategies for such explanations. The first is the subsumption strategy, which consists of subsuming a capacity to Φ of a system s under one or more general laws. Cummins illustrates the subsumption strategy with Archimedes' principle of the capacity of (some) objects to rise in water of their own accord. The second strategy is analytic. On it, a capacity to Φ of a system s is analysed into a number of other capacities ϕ_1, ϕ_2, ϕ_3, ..., had by that system or by parts of the system. Cummins illustrates this strategy with the explanation of the capacity of an assembly-line to produce certain products, by means of the capacities to perform certain tasks, possessed by the workers and the machines that are part of the line. These two strategies may be combined: an explanation of a capacity of a system may start analytically and end with employing the subsumption strategy.

A capacity to ϕ of a system x can be called a function of x relative to an explanation of a capacity to Φ of a system s that is identical to or encompasses x, if this capacity to ϕ is part of an analytic explanation of Φ. More precisely, Cummins' definition of a function ascription reads:[28]

> x functions as a ϕ in s (or: the function of x in s is to ϕ) relative to an analytical account A of s's capacity to Φ just in case x is capable of ϕ-ing in s and A appropriately and adequately accounts for s's capacity to Φ by, in part, appealing to the capacity of x to ϕ in s.

The analytical account A consists of the rules or theories that drive the analytic explanation of Φ in terms of, among others, x's capacity to ϕ. Cummins illustrates his definition with the ascription of the function 'to pump' to a heart. This ascription satisfies the definition when the heart is considered against the background of an analysis of the capacity of the circulatory system to transport nutrients, oxygen, etc., because a living heart (x), part of the circulatory system (s), is capable of pumping (ϕ-ing), and this capacity occurs in the explanation by means of a biological account (A) of the capacity of the circulatory system

[27] Cummins (1975, pp. 757–763). Initially Cummins speaks about explanations of *dispositions* of systems. But while developing his theory, he changes terminology and calls these dispositions *capacities*.

[28] Cummins (1975, p. 762). We adjusted the notation to the one in this book.

to transport nutrients, oxygen, etc. (to Φ). In this illustration, x and s are material objects and the respective capacities to ϕ and to Φ can be taken as physical dispositions of those objects. But Cummins applies his theory much more broadly. Another illustration he gives are the functions of workers and machines that are part of an assembly line. The entity x is then not just a physical object but may be an intentionally acting agent. Likewise, the system s encompasses not just material objects, but also workers and all kinds of procedures involving the objects and workers. The capacity to ϕ may refer to the capacity of workers to perform certain actions. And the capacity to Φ refers now to the capacity of the assembly line to produce certain products. These capacities cannot be seen as purely physical dispositions.

More generally, Cummins' definition may be construed in the following way. The encompassing system s is a system in a broad sense of the word: it may include only material objects, but also intentional agents and procedures. The capacity to Φ is the capacity of s that is to be explained, A is the analytical account on which the analytic explanation of Φ is based, and to ϕ is the capacity of x that occurs in the explanation of Φ. The concept 'capacity' may also be construed broadly. A capacity may be a physical disposition, such as the capacity of a heart to pump blood, or it may be a more intentional capacity, such as the capacity of a worker at the assembly line to check sheets of tin for mechanical damage. We, however, assume that if x is an artefact, then to ϕ is a physicochemical capacity.

Cummins' function theory is often called a causal-role theory because the functions of systems refer by his definition to the causal roles these systems have in larger systems; if a capacity is a function of x, then on Cummins' definition, that means that x contributes via this capacity to a capacity of a larger system s.

In the C-theory, functions of artefacts unambiguously refer to capacities of the artefacts and not to goals. And through its reference to an 'adequate and appropriate' account A, it can be argued that the C-theory primarily addresses the conditions for justified *function ascriptions*. But in most discussions, reference to A is omitted and Cummins' functional analysis is presented as a theory of functions *as properties* rather than of function ascriptions.[29] Presented in this way, the theory says that functions correspond to actual capacities that causally contribute to capacities of encompassing systems. This re-reading of Cummins' functional analysis implicitly requires that a correct account A underlies function ascriptions.

An advantage of the causal-role function theory is that it has an in-built guarantee that functional items can exercise the capacities corresponding to their functions. If an artefact is ascribed a C-function to ϕ, then Cummins' definition entails that the artefact 'is capable of ϕ-ing'. Another advantage of the theory is its broad scope. Both traditional and highly innovative compo-

[29]One example of this re-reading of Cummins' theory can be found in (Ariew and Perlman 2002, p. 1): 'The statement 'hearts function to pump blood' indicates a causal role the organ plays within the circulatory system of vertebrates [...] [H]uman hearts contribute to all sorts of 'systems' in virtue of their effects [...]'. No reference is made to analytic accounts.

nents can have C-functions. It also allows for multiple function ascriptions to items. A metal pipe, for instance, can truthfully be ascribed the C-function to transport fluids as part of an installation for synthesising chemical compounds, and can be ascribed the C-function to increase structural integrity as part of the building in which this installation is standing. Finally, causal roles of components of which users and possibly even designers are unaware can be taken as C-functions in retrospect. If for instance an electrical system is found to have caused an explosion in a chemical installation, say because it short-circuited and overheated, 'to detonate' can be ascribed to the system as a C-function. In terms of the three desiderata that we use in our review, the causal-role theory satisfies the desiderata of support and innovation.

A well-known problem of the causal-role function theory is that it cannot account for malfunctioning. The definition of function ascriptions to items entails, as noted above, that these items have the capacity corresponding to that function. Consider, for instance, a broken-down television set. It makes no sense to claim that this set is C-malfunctioning: if it does not have the capacity to show broadcasts, it cannot be ascribed the C-function of showing broadcasts in the first place. The C-theory thus fails to meet the malfunctioning desideratum. Moreover, it suffers, just like the intentional function theory, from a proliferation problem. Artefacts play many causal roles in many encompassing systems; all those roles correspond to functions, which results in many counterintuitive function ascriptions. Everything under the sun, for instance, contributes to casting its own shadow, but, intuitively, only a few things have the function to do so. And any building is part of a system composed of items which contain an even number of screws (or of the one composed of items with an odd number) and contributes to the weight of that compound system, although it seems odd to call this contribution a function.

Other current causal-role function theories

The idea that functions of artefacts are connected to their causal roles can be found regularly in the literature. Searle's (1995) function theory, which we discussed already, can be taken as one that incorporates this idea. For Searle, a function of an item is related to what the item causes or results in (condition 2, as we cited it on page 55) within a larger system of which the item is a part (condition 1). Especially if one interprets Searle's requirement that artefacts perform their functions solely in virtue of their intrinsic physical features (that is, with condition 3), Searle's function theory can be taken as an instance of the C-theory. Artefacts have, in that case, the capacities corresponding to their functions as actual physicochemical capacities. Artefacts are, moreover, part of larger systems that, on Searle's theory, are in part teleologically defined by the 'purposes, goals and values generally' agents impose on them. If one now assumes that artefacts contribute, or are supposed to contribute, to attaining these purposes, goals or values via the capacities corresponding to their functions,[30] then the capacity to Φ that is to be explained, would be the capacity

[30] Searle does not explicitly articulate this assumption that functional items contribute via

of the larger system to realise the purposes, goals or values that agents impose on that system. On this interpretation, Searle actually avoids the proliferation problem of nonsensical function ascriptions: only those causal roles of artefacts that contribute to purposes, goals or values that agents impose on larger systems become functions, which seems to be correct. But the malfunctioning desideratum is still not met.

There are a number of other current function theories that incorporate the causal-role function theory or substantial elements thereof. These theories are not pure C-theories like Neander's (1991a, 1991b) theory is a pure I-theory; rather, they combine the C-theory with the other two basic function theories. Examples are the function theories by Kitcher (1993), by Preston (1998b) and by Paul Sheldon Davies (2000, 2001). We forestall discussion of these combinatory approaches to section 3.5.

3.4 The evolutionist function theory

The final basic theory is the *evolutionist function theory*, abbreviated as the *E-theory*. This theory is similar to parts of the etiological theories proposed by Millikan (1984, 1993) and Neander (1991a, 1991b). But we chose to call it 'evolutionist' rather than 'etiological', to make clear that it resembles only parts and not the whole of the etiological theories. In section 3.2, we showed that Neander's theory is an intentional function theory in the technical domain; in section 3.5, we show that Millikan's theory contains intentional elements too. One can understand the presence of the intentional element in Neander's theory (and the fact that it surfaces in a review of function theories for the technical domain) by noting that her general definition — 'the proper function of a trait is to do whatever it was selected for' — harbours an ambiguity. The key-term 'selection' can be taken as referring to a *long-term process* in which the items that are selected survive numerous cycles of sortings from competitor items. And this term can refer to a *one-shot process* in which there is just one event in which the selected item is picked out. For the biological domain, Neander opts for the long-term meaning by appealing to the evolutionary notion of natural selection; for the technical domain, she opts for the one-shot meaning by taking selection as intentional selection by agents. It is, however, also possible to stick to the long-term meaning and define functions of artefacts on the basis of more extensive selection histories of the artefacts. The reproduction of artefacts by artisans is an example of such long-term selection. Moreover, evolutionary accounts for technology have been proposed, which describe the development of artefacts as long-term selection processes, e.g., Basalla (1988), Mokyr (1996, 2000) and Aunger (2002). These proposals have not yet led to a theory that is as generally accepted as neo-Darwinian evolutionary theory.[31] Yet they do

their functions to the purposes, goals and values of the encompassing system.

[31]This lack of consensus is clearly visible in (Ziman 2000), a volume that brought together work on evolutionary theories in the technical domain, including Mokyr (2000), Constant (2000) and Fleck (2000), and contains quite diverse proposals (Vermaas 2002). See also

show that it is possible to define an evolutionary framework for technology and to define the functions of artefacts on the basis of long-term selection. In her etiological theory, Millikan (1984, 1993) explicitly distinguishes these long-term and one-shot processes and defines two separate types of functions for both biological items and artefacts. The first type are called 'direct proper functions' and are defined relative to 'reproductively established families' of items, where 'reproduction' is a long-term process (we consider these functions in the second half of this section). The second type are called 'derived proper functions' and are defined for individually produced items, where 'production' is a one-shot process (we consider those in the next section). Our evolutionist function theory is abstracted from the long-term selection/reproduction part of etiological theories and is not meant to encompass the 'one-shot elements'.[32]

Against this background, the evolutionist function theory applies to any artefact x that has a long-term reproduction history, that is, to any artefact x that can be taken as a successor in a series p, p', p'', ... of predecessor artefacts. A capacity to ϕ counts as an evolutionist function of an artefact x if and only if that capacity contributed positively to the reproduction of its predecessors and the current artefact x.

We deliberately left this definition somewhat vague. As said, there is not one well-established evolutionary account for the development of artefacts. Consequently, there is not one canonical way of filling in the details. A simple description of a long-term reproduction history of an artefact x is one in which it and all its predecessors p, p', p'', ... actually have the capacity to ϕ and all items form a sequence of exact copies: x is a copy of its predecessor p; this predecessor p is in turn a copy of its predecessor p', which is a copy of p'', etc. A more interesting description is one in which there is some room for error in the copying process and where it is left open whether x itself has the capacity to ϕ; in this case x can malfunction by not having the capacity to ϕ. On both descriptions, the reproduction history can be captured by the following scheme:

$$\ldots \quad \rightarrow \quad p'' \quad \rightarrow \quad p' \quad \rightarrow \quad p \quad \rightarrow \quad x$$

An example that fits this scheme could be Bigelow and Pargetter's hammer (see footnote 13 on page 53) that is copied over the centuries by artisans.

More sophisticated descriptions of the reproduction history can be envisaged. Only very few artefacts copy themselves; virtually all artefacts are copied by agents such as designers, engineers or artisans. These agents typically decide to copy artefacts for the capacity to ϕ. This introduces representations R_p, $R_{p'}$, $R_{p''}$, ... of the artefacts and of their capacities by agents, and the possibility that the predecessor artefacts p, p', ... do not have the capacity to ϕ; the agents need only be *convinced* that the predecessors have this capacity. A scheme for such a reproduction history is:

Lewens (2004, ch. 7) for a discussion of a possible evolutionary theory of artefacts, and see Brey (2008) for an analysis of the different senses in which the proposals by Basalla (1988), Mokyr (1996, 2000) and Aunger (2002) can be taken as *evolutionary*.

[32] In (Vermaas and Houkes 2003) we expanded on this distinction between the long-term and one-shot notions of etiological functions.

$$\ldots \Rightarrow p'' \mapsto R_{p''} \Rightarrow p' \mapsto R_{p'} \Rightarrow p \mapsto R_p \Rightarrow x$$

(We use different types of arrows to indicate that other processes are involved besides the copying process of the first model). Artisan reproduction of simple artefacts, such as hammers, may fit this scheme, as may the present-day production of screwdrivers. In this last case one can construe one series of screwdrivers by one company as a variation/piracy of a predecessor series of screwdrivers produced by another company, etc. These screwdrivers thus have the E-function to turn screws, because predecessor screwdrivers had the capacity of turning screws and this capacity contributed to their production.

For technologically more complicated artefacts one may expect that the agents involved in creating artefacts base their designs not on only the analysis of predecessor artefacts but also on the designs of these predecessors. A scheme that emphasises the role of design descriptions is one in which the representation of an artefact is primarily derived from the representation of the predecessor of the artefact:[33]

$$\begin{array}{ccccccccc} & & p'' & & p' & & p & & x \\ & & \Uparrow & & \Uparrow & & \Uparrow & & \Uparrow \\ \ldots & \rightsquigarrow & R_{p''} & \rightsquigarrow & R_{p'} & \rightsquigarrow & R_p & \rightsquigarrow & R_x \end{array}$$

An example may be the development of cars, where the new model is developed on the basis of the design of the previous model.

This last scheme seems to introduce a technological analogue of the biological distinction between the genotype and phenotype of items, and thus seems to bring us to familiar evolutionary grounds. It may, however, be more fair to say that if there exists an evolutionary theory for artefacts, it will be more complicated than the schemes presented. Some authors single out a genetic code for artefacts, for instance by using Dawkins' (1976) notion of memes, but again there is no consensus about what to take as artefact-genes.[34] Furthermore, it seems that an evolutionary theory for artefacts will have to allow for some feedback mechanism between the artefacts and their representations and thus integrates the second and third schemes.[35]

In the domain of biology, the E-theory seems to analyse 'having a function' without giving explicit conditions for function ascriptions. Yet in biology E-functions are ascribed on the basis of a specific, allegedly true theory, namely

[33]Strictly speaking, a history that satisfies this scheme or the previous one need not be one in which the artefact x and its predecessors structurally resemble one another. To restrict these schemes to reproduction histories, extra constraints should be introduced. One possibility is to require that the different couples p–p', p'–p'', etc., of predecessors resemble each other (the couple x–p can be exempted from this constraint in order to allow for a malformed end result of the history).

[34]Aunger (2002) chooses *memes* as the technical analogues of genes, Mokyr (1996, 2000) takes *useful knowledge* as such genes, Constant (2000) opts for *fabrication techniques*, *designs*, *technological knowledge* and *scientific knowledge*, and Fleck (2000) settles on '*artefact-activity pairs*'.

[35]Lewens (2004, ch. 7) argues, for instance, that the identification of the genotype of artefacts may differ from context to context.

neo-Darwinian evolutionary theory — reference to which can reasonably be bracketed. Without this bracketing — and in the technical domain, there seems not yet to exist a theory that is widely enough accepted to be bracketed — the evolutionist function theory can be taken as one that ascribes functions relative to an evolutionary account of biological or technical items.

An advantage of the E-theory is that it gives support to the belief that an item has the right physicochemical structure to perform its E-function, while leaving room for failures to perform this function. Consider, for instance, a scheme in which the predecessors p, p', p'', ... have the capacity corresponding to the function ascribed to the artefact x and where it is an open question whether x also has this capacity. A plausible constraint on the notion of 'copying' is that each couple of direct predecessor–successor (i.e., the couples x–p, p–p', p'–p'', etc.) should bear some physical similarities to each other. This constraint, together with the fact that p did have the capacity to ϕ, makes it *prima facie* probable that the reproduced item x also has that capacity. This support is, however, not so strict as to rule out E-malfunctioning. An item can be a malformed reproduction of its predecessors, as long as it bears some similarities to them. Such 'black sheep' or imperfect recent specimens are still ascribed E-functions even though they are unable to perform these functions. This advantage is not built into E-theories by definition; one can envisage schemes for the E-theories in which the balance between support and malfunctioning is lost. If, for instance, too much emphasis is put on the reproduction of representations of (alleged) capacities of artefacts, then a perpetual motion machine constructed, with some variation, in accordance to earlier specimens, has the E-function to generate free energy. But it can be concluded that the evolutionary theory is at least in principle able to meet both the support and malfunctioning desiderata.

A disadvantage of the E-theory is that it presupposes a reproduction history for every item to which E-functions are ascribed. This supposition is not obviously met. A first observation is that even though evolutionary theories are proposed for technology, the project is still in its initial phase of exploring the possibility and tenability of describing the development of artefacts by means of reproduction histories. Only the future can prove whether this project is successful. A second observation is that there exist forms of technological innovation that can in principle not be described in terms of reproduction histories. If, for instance, an artefact is designed that does not resemble previously existing artefacts, this artefact cannot be taken as one that has predecessors. It therefore lacks an E-function. The first aeroplane and the first nuclear plant may be cases in point. Proponents of evolutionary accounts argue that innovation is actually a process that unfolds in a much more gradual way, and present historic analyses that reveal that artefacts that appeared novel to the general public actually evolved via a series of intermediate steps from existing artefacts.[36] But this argument is beside the point, or rather an argument against the possibility of innovation in the considered sense: an artefact that does not resemble existing artefacts still does not fit a reproduction history. One can, alternatively,

[36] E.g., Basalla (1988, ch. 2).

try to accommodate this case by allowing that in a reproduction history the predecessors of an artefact may be ideas of the designers of the artefact. This proposal may save the evolutionary account in the case of innovation, since one can then try to show that the design sketches that engineers have generated provide the 'missing link' between innovations and existing artefacts.[37] But it clearly stretches considerably the notion of copying that underlies the long-term reproduction of artefacts — or even overstretches it. For on this proposal, there need not be physical predecessors p, p', ... of the innovative artefact x, but only representations R_p, $R_{p'}$, ... of those non-existing predecessors. Hence, the functions that are ascribed to the innovative artefact x are ascribed on the basis of (a series of) 'copied' engineering ideas, suggesting that these functions are to be taken as I-functions instead of E-functions. This does not only lead to a reclassification of the theory, but also to a re-assessment of its potential problems. As an I-theory, focussed on representations alone, the current proposal might have the disadvantage of lacking support for function ascriptions.

A second disadvantage is that evolutionist function theories may lead to counterintuitive results. Again, consider innovation, now in the form of the ascription of a novel function to an existing artefact or an artefact similar to existing ones. This novel function differs by definition from the functions the existing artefacts were ascribe up to then. An example may be Aspirin, a drug with an existing function to alleviate pain. At some point, Aspirin was ascribed the novel function to prevent blood clotting in cardiovascular patients. The first nuclear power plant was a modified submarine engine, ascribed the novel function of generating electricity. Such artefacts have reproduction histories, but the E-functions ascribed to them are the existing ones, not the novel ones. So the first nuclear plant still only has the E-function of providing propulsion and all tablets of Aspirin, even those taken daily by cardiovascular patients, only have the E-function to alleviate pain. The evolutionist function theory thus does not meet the innovation desideratum in an intuitively satisfying way: it ascribes functions to innovative items, but at the price of explaining away the innovation.

A final problem is epistemic: the E-theory gives a narrow-minded account of function ascriptions by users. Users should determine the E-functions of an item solely by determining the capacities of predecessor artefacts that contributed to their reproduction. For some artefacts and for some users, this may indeed be the case. But users can also determine the functions of artefacts by examining the artefacts themselves or by being informed about the intentions of the designers of the artefacts.

Current evolutionist function theories

It is a recurrent theme in the literature that artefacts can often be described as copies of earlier artefacts. This phenomenon has led a number of authors to advance an evolutionist function theory. Millikan's theory of direct proper functions is a clear example to which we now turn. And in the next section we

[37] Carlson (2000) gives, for instance, an evolutionary account of design sketches of Edison.

show that authors such as Griffiths (1993), Preston (1998b) and Paul Sheldon Davies (2000, 2001) have combined E-theories with the other basic function theories.

Millikan's direct proper functions

Before discussing Millikan's (1984, 1993) etiological theory, it should be noted that her theory is extensive, and also applies to biology and linguistics. Our focus on artefacts therefore only leads to a partial assessment of Millikan's work. Moreover, Millikan defines several functional concepts, including 'direct proper functions' and 'derived proper functions'. Especially in (Millikan 1999), she presents these as elements of one overarching theory of 'proper functions'. We refrain from giving an overall assessment. In this section, we only consider the 'subtheory' of direct proper functions. In the next section, we switch to derived proper functions. Finally, Millikan presents her functional concepts as stipulations, which do not capture the ordinary meaning of functional discourse.[38] Therefore, we cannot properly evaluate her theory by means of our desiderata. Still, direct proper functions are clear examples of E-functions, so it seems worthwhile to review whether they are (despite their designer's intentions) suitable for our purpose.

Millikan[39] ascribes direct proper functions to items that are members of 'reproductively established families', that is, items that are members of a set of items that are related by 'reproduction'. Reproduction is defined as a direct causal relation between items or their properties, which leads to a counterfactual-supporting similarity. To account for malformed items, Millikan allows a final member that only approximately resembles earlier family members. If it is the case that (*i*) for earlier members of such a family — which take the role of predecessors p, p', etc. — there is a positive correlation between the reproduced similarity (the 'character') and a certain capacity to ϕ of the members, and (*ii*) the existence of a present member x can be explained on the basis of this positive correlation, then that capacity is called a 'direct proper function' of the present member x. Artefacts 'that are not of original design', such as household screwdrivers of which 'the same design has been copied over and over',[40] are examples of items that form such reproductively established families. Mass-produced products that come off an assembly line are, however, not direct reproductions in Millikan's sense. To ascribe direct proper functions to those items, Millikan introduces a distinction between first-order families, the members of which have counterfactual-supporting similarities, and higher-order reproductively established families. Just as the members of a first-order family, those of a higher-order family should have some properties in common, but this similarity need not support counterfactuals. The first-order and higher-order levels are connected as follows: if a set of entities is the product of a set of members of a lower-order family, the direct proper function of which is to produce such entities, they are called a higher-order family. Hence, if a number of

[38]See, e.g., Millikan (1984, p. 18) and, especially, Millikan (1989).
[39]Millikan (1984, ch. 1).
[40]Millikan (1984, pp. 21 and 23).

assembly lines for roofing tiles form a first-order family, the roofing tiles they produce form a higher-order family. The definition of higher-order reproductively established families is more involved than presented here, because Millikan also allows the possibility of production of members of a higher-order family by a single entity with the direct proper function of producing similar items. So, if there is just one assembly line, the produced tiles still form a higher order family. The ascription of direct proper functions to items of those higher-order families is similar: if (i) for earlier members of such a family there is a positive correlation between the reproduced similarity and a certain capacity to ϕ of the members, and (ii) the existence of a present member x can be explained on the basis of this positive correlation, then that capacity is called a 'direct proper function' of the present member x.

This theory of direct proper functions presupposes that artefacts are elements of reproductively established families. As argued above, this rules out the ascription of a proper function or of an intuitively correct proper function to innovative artefacts such as the first aeroplane and the first nuclear plant. Hence, if this theory were meant as capturing the ordinary meaning of function ascriptions, it would not meet the innovation desideratum. This conclusion does not apply to Millikan's theory *per se*, since it, as said, also contains the concept of a derived proper function.

3.5 Combining the basic theories

Some currently existing function theories do not resemble a single basic function theory. Some proposals *combine* elements from the basic theories, resembling more than one of them. We therefore continue our global review by considering, in general, function theories that can be interpreted as combinations of the I-, C- and E-theories.

One might suspect that such hybrid theories might succeed where the purebred theories fail. The three basic function theories all fail to meet at least one of the three desiderata that we consider in this chapter. Yet the basic theories appear to solve each other's problems. In different ways, the C-theory and E-theory provide the support that the I-theory is lacking. The C-theory and I-theory broaden the narrow focus of the E-theory on reproduction histories. I-functions and E-functions can be ascribed to malfunctioning items, unlike C-functions, and the E-theory lacks the malfunctioning items that plague the I-theory. C-functions and I-functions can, finally, be ascribed to innovative items, which does not hold for E-functions. The three basic theories may thus be taken as complementary; a function theory that combines two or all of these basic theories may satisfy all three desiderata. In this section, we argue that straightforward combinations of the basic theories do not fulfil this promise.

If one considers the possibility of combining the I-, C- and E-theories, a first preliminary question is whether it is even possible to combine them. For if one of these theories can be taken as an instance of (a combination of) the others, some combinations yield no results. This is not just a hypothetical scenario. If our I-theory is construed as liberal as possible, every functional description becomes an I-function ascription: after all, functional descriptions require agents and their representations. To avoid this inflation, we take the I-theory as one that relates functional descriptions of artefacts only to the intentions and actions of designers and users of the artefacts. Functional descriptions by outside observers need not be I-function ascriptions in this narrower sense. Then it is possible to show that the three basic function theories are independent of one another: one can come up with function ascriptions that satisfy only one, only two or all theories, in every possible combination (see the box on the following page).

A second, more pertinent preliminary question is: how can function theories be combined? There are two straightforward strategies, namely disjunction and conjunction. A *disjunctive combination* of two or more function theories allows all function ascriptions allowed by one or more of the combined theories. On the *conjunction strategy*, one accepts only those function ascriptions that are allowed by both of the combined theories.

Below we show that a number of current function theories for the technical domain resemble such disjunctions or conjunctions of our three basis function theories. And we show that these function theories still do not satisfy the three desiderata. More generally it can be argued that both the disjunction and the conjunction strategies will not do as methods to overcome the problems that we identified for the three basic theories.

The general advantage of the disjunction strategy is that the combination need not suffer from the limitations in scope of its constituents. So, a theory that is a disjunction of the E-theory with one of the other basic theories can ascribe functions to the innovative artefacts to which the E-theory itself cannot ascribe functions. But an item that falls outside the scopes of both disjunctively combined theories is still not ascribed a function. Hence a 'disjunctive CE-theory' still cannot deal with malfunctioning innovative artefacts. Other drawbacks are a proliferation of function ascriptions and a loss of conceptual unity. Disjunctive theories introduce an *intra-domain pluralism*, which should be distinguished from *inter-domain pluralism*.[41] According to intra-domain pluralist theories, artefacts have two or possible three types of functions.

The advantage of the conjunction strategy is that the combination inherits all benefits of its constituents. A conjunctive IE-theory, for instance, reproduces the intuitive relation between the intentions and actions of agents and

[41] One could argue that Neander's (1991a, 1991b) function theory is an inter-domain pluralist theory. Her function theory applies to the biological and technical domains but ascribes different types of functions to items in these domains: E-functions to biological items and I-functions to artefacts. But Neander's theory does not lead to intra-domain pluralism: to items of each domain this theory ascribes but one type of functions. Preston's (1998b) theory (see this section) is not an inter-domain pluralist theory because she proposes the same disjunctive CE-theory for both biology and technology; it is an intra-domain pluralist on both domains, since Preston allows the ascription of C-functions and of E-functions to biological items and to artefacts. Finally, Mahner and Bunge (2001) seem to offer a combination of inter- and intra-domain pluralism, in which five different notions of function are combined in different ways in, on the one hand, the biological domain and, on the other hand, psychology, social science and technology.

Independence of the basic function theories:

Assume that a novel vehicle for subterranean transportation does not have a reproduction history and does not actually work. If one nevertheless ascribes it the function of subterranean transportation because it was designed for that capacity, this function is an I-function, but not a C- or E-function. The initial discovery that Aspirin prevents blood clots counts as an ascription of an I- and C-function, but not of an E-function. Similarly one can come up with function ascriptions that satisfy one, two or all basic function theories, in every possible combination:

I,¬C,¬E:	This novel vehicle has subterranean transportation as a function.
C,¬I,¬E:	This short-circuited electrical system had the function of detonating the explosion.
E,¬I,¬C:	These souvenir clogs are for walking.
I,C,¬E:	Aspirin has the function to prevent blood clots.
I,E,¬C:	This water has the function to shorten nylon rope.
C,E,¬I:	The vanadium in the steel of Damascus swords has the function to produce the typical patterns on their blades.
I,E,C:	Axes have the function to cut wood.

Three notes:
First, we assume that souvenir-clogs do not have the functions clogs originally had.
Second, if one puts water on old-fashioned hemp-fibre rope, the rope shrinks a bit. An agent who often used this trick may ascribe 'shortening rope' as an E-function to the water, also when applying the trick to nylon rope. We assume that water does not shorten nylon rope.
Finally, Damascus swords were made in Damascus from 'wootz steel' produced in India. The bladesmiths of Damascus produced blades from this steel with distinct 'damascene' surface patterns. In the 18th century this technique became a lost art. Recent research (Verhoeven et al. 1998) suggests that the trace element vanadium in wootz steel contributed to the formation of these surface patterns (together with the bladesmiths' techniques, of course). This suggestion leads to the hypothesis that the art of producing Damascus swords was lost because the steel that was used stopped to contain the right amounts of this element (say, because the old Indian ore body was exhausted and the new ore body did not contain vanadium impurities). On the supposition that the bladesmiths themselves did not know that vanadium contributed throughout the ages to the damascene patterns on their swords, the ascription of the function of contributing to these patterns to vanadium, is the ascription of a C- and E-function, but not of an I-function.

the IE-functions it ascribes to items *and* provides support to the belief that the items have the right physicochemical structure to perform these functions. The disadvantage is that the combination fails whenever any of its component theories fails. A conjunctive IE-theory, for instance, still does not ascribe (the right) functions to innovative artefacts.

The upshot of this discussion is that not only the I-, C- and E-theories fail to meet the desiderata of malfunctioning, support and innovation; all function theories that are elements of the 'space' spanned by disjunctions and conjunctions of these three basis theories fail to meet these desiderata. Hence, if the promise that the I-, C- and E-theories may be combined into an adequate theory is to be fulfilled, a more sophisticated way of combining these basic function theories is required. In the next chapter we show that our action-theoretical use-plan analysis of artefacts gives a background to functional descriptions of artefacts that allows for such a sophisticated and adequate combination. But before turning to our proposal, we end our review by giving some examples of current function theories that resemble disjunctions and conjunctions of pairs of the I-, C- and E-theories.[42]

Philip Kitcher: a disjunctive IC-theory

Philip Kitcher (1993) proposes a function theory that applies to both biology and technology and that, when considered on that last domain, resembles a disjunction of the intentional function theory and of (an instance of) the causal-role function theory. For Kitcher, 'the central common feature of usages of function [...] is that the function of S is what S is designed to do'.[43] He has to devote quite some space to apply this common feature to biology. For artefacts, the application is comparatively straightforward and leads to an I-like function theory: 'Imagine that you are making a machine. You intend that the machine should do something, and that is the machine's function.'[44] Yet Kitcher is not simply advancing an intentional function theory, for he also recognises that components of artefacts can have functions that are not explicitly related to the intentions of the designers of those artefacts. He therefore continues and writes about the machine you are imagining to be making:[45]

> The machine has a function grounded in our explicit intentions, and its fulfilling that function poses various demands on the parts of which it is composed. You recognize some of these demands and explicitly design parts that can satisfy them. But in other cases [...] you do not see that a demand of a particular type has to be met. Nevertheless, whatever satisfies that demand has the function of so doing. The function here is grounded in the contribution that is

[42] We have not been able to find existing proposals that can be taken as combining all three basic function theories, which conveniently allows us to present ours under the heading of an *ICE*-function theory.
[43] Kitcher (1993, p. 480).
[44] Kitcher (1993, p. 480).
[45] Kitcher (1993, p. 481).

made towards the performance of the whole machine and in the link between the performance and the explicit intentions of the designer.

This second way of ascribing functions to components can be unpacked as a causal-role function theory: it ascribes the function to ϕ to a component x in accordance with Cummins' definition[46] (see page 57), where the containing system s is the artefact of which x is an part, and where the capacity to Φ of that artefact that is to be explained corresponds to the function that the designer explicitly ascribed to the artefact.

As a disjunctive IC-theory, Kitcher's proposal inherits a number of the problems of the I-theory and the C-theory. Function ascriptions to artefacts as a whole can lack support — the machine in Kitcher's example can be designed for subterranean transportation and can then be ascribed this capacity as its function, even if it is physically identical to an ordinary car. Function ascriptions to components that are not explicitly designed for specific capacities do not satisfy the malfunctioning desideratum and these components can, because of the intra-domain pluralism, actually have conflicting functions. If, for instance, a component of some artefact is explicitly designed as an electrical insulator but contributes to the (intended) functioning of the artefact by conducting a current, then the component has both isolating and conducting as its functions.

Krohs: a conjunctive IC-theory

An example of a conjunctive IC-theory — more precisely: a C-theory that, for technology, includes elements of the I-theory — is Ulrich Krohs' (2009) general function theory, proposed for both biological items and artefacts. Krohs first defines a general concept of design as the *type-fixation* of a complex entity, meaning that the components of the entity are part of the entity because of their types and not merely because of their properties (2009, §3). A function to ϕ of a type-fixed component by some design is then a contribution of this component to a capacity to Φ of a system that is the realisation of the design.

Krohs introduces elements of the I-theory to prevent inappropriate C-functional analyses of systems — his examples are functions ascribed to physico-chemical systems like electrons in atoms and to clouds in the hydrological cycle. In particular, the process by which the systems come about, i.e., their ontogeny, should be one of 'design', i.e., the process should be type-fixing. For artefacts, this design is the blueprint, in which components such as engines, axes and wheels are described and thus intentionally fixed as engines, axes and wheels of some types. Physical systems lack a process of type fixation; they arise purely on the basis of physical properties of their parts, which consequently do not have functions. Via this design constraint, elements of the I-theory are incorporated and the proliferation problem of the C-theory is avoided. Simultaneously, the proliferation and support problems of the I-theory are avoided by the C-characterisation of a function as the actual contribution of a component to a system.

[46] Kitcher (1993, §V) explicitly integrates Cummins' (1975) theory when he discusses the application of his function theory to biological items.

Like all advocates of C-theories, Krohs can account for the relation between functions and physical properties (roughly, functions are contributions of type-fixed components to capacities of a system); and for technical innovations (as long as they have a type-fixing blueprint). Unlike purebred C-theories, Krohs' theory also handles some cases of malfunctioning: at least for components, type fixation determines a standard for the contributions of components, which they may fail to meet. This means that, in sum, Krohs' IC-theory satisfies the three desiderata central to this chapter.

Still, it inherits a problem shared by its constitutive elements: it is not clear how the theory accounts for the proper-accidental distinction. More specifically, Krohs' theory has no straightforward connection to the use-plan analysis that was earlier found to account for the proper-accidental distinction. The reason is that the I-element of the theory only refers to type-fixation of items in blueprints — in our characterisation: to product-designing. Krohs' general notion of design allegedly encompasses biology, but does not extend to plan designing. Thus, it is insufficiently general to enforce a distinction between proper and accidental functions.

To be fair, Krohs analyses 'having *a* function', which does not call for privileging any system-with-design over another system-with-design. Thus, in one system, milk cans have a function to store milk, while in another, they are to grow flowers in, as long as there are designs for both systems. At the end of his (2009) paper, Krohs remarks that elements from the etiological theory might serve to supplement his analysis and introduce the necessary privileging of systems. Our own proposal in the next chapter vindicates this suggestion. A sophisticated combination of the intentional, causal-role and evolutionist theories, against the background of our use-plan analysis, indeed satisfies all four desiderata.

Millikan: a disjunctive IE-theory

As we indicated above, on Millikan's (1984, 1993) etiological theory, artefacts can be ascribed direct proper functions and derived proper functions. We already argued that direct proper functions resemble evolutionist functions. In a moment, we show that the characterisation of derived proper functions resembles the intentional theory. Millikan's theory can thus, in sum, be taken as a disjunctive IE-function theory. But this position should be qualified,[47] for Millikan presents the characterisations of the two types of functions as subtheories within one overarching theory of proper functions.[48] The (intra-domain) pluralism implied by taking her theory as a disjunctive theory may thus be only apparent (but see Preston (1998b, pp. 225–239)).

In the previous section, we introduced Millikan's direct proper functions. Membership of a reproductively established family is a necessary condition for ascribing such a function to an item. There is no such condition for derived

[47] Lewens' (2004, §5.1 and ch. 7) analysis of function ascriptions to artefacts may be a more faithful example of a pluralistic disjunctive IE-theory.

[48] E.g., Millikan (1999).

proper functions. Millikan introduces this notion with an example from biology.[49] The skin of a chameleon is equipped with a colouring mechanism, i.e., a mechanism that redistributes pigments. This mechanism has a direct proper function: earlier chameleons had a similar mechanism, which was reproduced because of a positive correlation between it and a survival-enhancing performance. This direct proper function is relational, namely to turn the colour of the chameleon's skin into the same colour as its environment. This relational (direct) proper function of the mechanism can be used to define the function of items it produces; in this case, the specific colours that the chameleon's skin acquires. A specific colour need not be a member of a reproductively established family of any order: it might well be the first token ever of this colour in chameleons. Yet it is produced by a device that has a relational proper function. Therefore, Millikan's theory says that the specific colour has the *derived proper function* to hide the chameleon in a specific environment. Derived proper functions are thus ascribed on the basis of an item's causal history, namely its production by an item that has a direct proper function. Moreover, this criterion leaves room for intentions, since there are no constraints on the nature of the production process involved. And the way in which Millikan applies this theory of derived proper functions to artefacts turns it into an intentional function theory for the technical domain.

According to Millikan, 'all tools have as derived proper functions the functions that their designers intended for them'.[50] Any artefact is a product of a system that has been reproduced, namely the intentional system of human agents.[51] If we are to ascribe derived proper functions to artefacts in accordance with this intentional theory, they must stand to the intentional system as a specific colour stands to the chameleon's colouring mechanism. However, this analogy breaks down at a crucial point. The intentional system primarily produces particular desires and beliefs, and artefacts only through these; no such intermediaries exist for the chameleon's colouring mechanism. Moreover, the chameleon's skin mechanism necessarily performs its function by producing a particular colour, but the intentional system can perform its function without producing artefacts, for instance by instigating immediate action. Millikan seems aware of this problem. She argues as follows: particular desires have derived proper functions as products of the intentional system. Their proper function is to get themselves fulfilled. Artefacts are the products of desires that have the proper function to get themselves fulfilled using artefacts. Hence, she claims, artefacts also have derived proper functions:[52]

---

[49] Millikan (1984, ch. 2).

[50] Millikan (1984, p. 49).

[51] In fact, one of the main advantages Millikan claims for her theory of biosemantics is that it naturalises intentionality. She claims that intentionality and deliberation, as constellations of desires and beliefs, arise as products of evolutionary pressure towards true beliefs and realisable desires. We remain agnostic towards this account — which is more important for the unity of Millikan's theory than for its application to the technical domain anyway — and simply study its consequences for artefacts.

[52] Millikan (1999, p. 205).

> [...] if the desire is to produce a certain result by means of making, for example, a certain tool [...] then a derived proper function of the tool (token) [...] is to produce the result. Thus, it happens that artifacts have as derived proper functions the functions intended for them by their makers [...]

Whether one accepts this solution or not, our construal of Millikan's theory as a disjunctive IE-theory suggests that other problems are to be expected, namely at the intersection of problematic cases of the I-theory and of the E-theory. In particular, the theory leads to unsupported function ascriptions to innovative artefacts. On Millikan's theory, the vehicle for underground travelling is not ascribed a direct proper function at all, but it *is* ascribed the derived proper function of subterranean transportation, which seems counterintuitive.

Sperber: a disjunctive IE-theory

Sperber (2007) presents a function theory that applies to both biology and technology, including the cross-section of these domains made up by agricultural crops, domesticated animals and the organisms that are the result of bioengineering. Sperber adopts his general concept of function, called 'teleofunction', from Millikan: 'an effect of type F is a *teleofunction* of items of type A just in case the fact that A items have produced F effects helps explain the fact that A items propagate, i.e. keep being re-produced.'[53] Teleofunctions can be biological or cultural, and for the technical domain the last are the relevant ones. Items that can have *cultural teleofunctions* are what Sperber calls mental representations and public productions: 'Mental representations are constructed within agents by mental processes. By 'public productions', I mean both behaviors (e.g. speech) and traces of behavior (e.g. writings) that can be perceived and therefore serve as input to the mental processes of other agents.'[54] Sperber illustrates cultural teleofunctions with the example of suntans. The mental representations of suntans and the actual suntans are currently reproduced because of their perceived attractiveness, making this attractiveness their cultural teleofunction. Sperber acknowledges that cultural teleofunctions do not capture all functions of artefacts; '[t]he function of an artifact qua artifact' may depend on the artefact 'having been intended by whoever devised the artifact.'[55] To capture these remaining functions Sperber defines the notion of an *artifactual function* of an artefact as the intended effects that explain why the artefact is being produced; the example is a tree leaf that is folded by someone for the purpose of retrieving a ring fallen in the crack between two floorboards. This folded leaf has retrieving the ring as its artifactual function but not necessarily as a teleofunction since the folded leaf need not be an instance of a type that is propagating. Yet Sperber notes that artifactual functions may quickly become cultural teleofunctions when they are being reproduced for their artifactual functions: 'New sugar cubes are being reproduced with the expectation

[53] Sperber (2007, p. 128; original emphasis).
[54] Sperber (2007, p. 128).
[55] Sperber (2007, p. 129).

and intention that, by dissolving, they will sweeten hot drinks (this is their intended effect) because sugar cubes have reliably had this effect in the past (and therefore this is also their teleofunction).'[56]

In our terminology the cultural teleofunctions are instances of E-functions and the artifactual functions are I-functions, turning Sperber's theory into a disjunctive IE-theory. Being an IE-theory, it suffers from the same problem that we discussed when considering Millikan's theory, namely that it leads to unsupported function ascriptions to innovative artefacts.

Griffiths: a conjunctive IE-theory

Griffiths (1993) constructs a function theory for artefacts that can be taken as a conjunctive IE-theory. Griffiths' main focus is biology; for that domain he sets out to integrate the C-theory with an E-theory. But at the end of his paper, he applies his approach to artefacts. He distinguishes an I-like proposal by which 'the functions of artifacts are their intended uses'[57] and an E-like proposal in which the proper functions of artefacts correspond to those capacities for which designers consciously or unconsciously — e.g., as in design by trial and error — select these artefacts to be reproduced. Griffiths then stipulates that the differences between these proposals have to be overcome in order to understand the process of designing artefacts as a selection process, and suggests that 'the function of an artifact is its intended use only because its ability to fulfil its intended use gives it a propensity to be reproduced.'[58] Hence, functions of artefacts are both I- and E-functions.

This conjunction with the E-theory avoids some of the disadvantages of the I-theory: it seems unlikely that on Griffiths' theory detergents of one company can be ascribed the function to weaken the market position of another company. But many disadvantages of the E-theory also pertain to this conjunctive IE-theory: innovative artefacts, for instance, are still not ascribed (the right) functions.

Preston: a disjunctive CE-theory

Preston (1998b, §III) formulates a function theory for both biological and technical items, which she presents as a pluralist disjunction of Millikan's theory of direct proper functions and Cummins' theory. On her terminology, *proper functions* refer to Millikan's direct proper functions, and *system functions* refer to Cummins' functions; on ours, Preston's proposal counts as a disjunctive CE-theory.

Proper functions of artefacts are for Preston 'the functions we are most likely to describe in answer to questions like 'What is that for?' or 'What is that?' '.[59] 'Artifacts get these proper functions by a process analogous in basic respects to the natural-selection process by which biological traits get theirs'[60] and Preston

[56] Sperber (2007, p. 129).
[57] Griffiths (1993, p. 418).
[58] Griffiths (1993, p. 420).
[59] Preston (1998b, p. 243).
[60] Preston (1998b, p. 243).

is indeed staying close to evolutionary theory in explicating her theory. An inventor or designer may produce a new artefact, or an agent may put an existing artefact to a new use. Yet only if this new artefact or new use is successful, the artefact will be reproduced and the capacity for which it is reproduced will become its proper function. So, compared to Millikan, Preston emphasises users at the expense of the designers of artefacts. Users determine which artefacts are successful and their preferences create the technical analogue of the selective environment for biological traits. Designers seem to play a subsidiary role. They generate the artefacts or create enhanced and more complex versions of them (and do so quite effectively by the fact that they proceed intentionally). But their intentions, when designing or reproducing those artefacts, do not (fully, or even partly) determine the proper functions of those artefacts.[61] So, if an artefact has a series of predecessors, and the item and these predecessors are all designed for some desirable capacity to ϕ, then this capacity is a proper function of the artefact only if this capacity made the items successful for users.

Before an artefact has some capacity as a proper function, e.g., as long as the artefact is new and cannot have been reproduced for that capacity, it has the capacity as a system function. Hence, in Preston's theory, system functions of artefacts enter the arena from which some may emerge as proper functions. Preston also uses the concept of system functions for analysing other phenomena concerning artefacts and artefact use. She, for instance, defines the sets of *standardised ongoing exaptations* and *idiosyncratic ongoing exaptations*[62] for capturing that artefacts are often systematically used for specific capacities, even though these capacities are not the ones for which the artefacts are reproduced. Standardised ongoing exaptations are culturally widespread cases of such systematic use. Examples are chairs that are standardly used as steps, and screwdrivers that are standardly used to open paint cans with. Idiosyncratic ongoing exaptations are private creative uses of artefacts. Preston herself, for instance, uses an antique cast-iron doorstop as a bookend and a ten-pound Planter's peanuts tin as a wastebasket. She can characterise these ongoing exaptations as defining system functions of artefacts that are not their proper functions: the exaptations single out capacities of the artefacts that explain their systematic uses, but not capacities for which they are reproduced.

Taken as a disjunctive CE-theory, problems with Preston's proposal are to be expected for cases in which both the C-theory and E-theory fail. Such cases concern innovative but malfunctioning artefacts. This class of artefacts may be numerically small in comparison with other classes of artefacts, but it is of considerable importance conceptually. The production of innovative prototypes, many of which will malfunction, is a standard element of most design processes. Since the aim of this process is to produce a functional object, not being able to ascribe functions to the unsuccessful prototypes seems unacceptable. Preston seems to acknowledge this problem; she, for instance, argues to give up the

[61] Preston argues, for instance, in her (2006), that designer's intentions cannot be taken as sufficient to establishing a proper function of an artefact; see also (Preston 2003).

[62] The term 'exaptation' originates from the literature on evolutionary theory and biological functions (Gould and Vrba 1982), and is here transposed to the technical domain.

innovation desideratum.[63] And she rightfully points out that her theory still allows her to describe these malfunctioning prototypes as artefacts that were designed by agents who had particular purposes in mind for the artefacts.[64]

Paul Sheldon Davies: a conjunctive CE-theory?

The C-theory and the E-theory are the two main function theories in biology.[65] A number of authors have proposed accounts for that domain that combine these two theories. Preston's theory is a disjunctive CE-theory for biology, like Kitcher's (1993). Conjunctive CE-theories can also be found in that domain. Godfrey-Smith (1994) incorporates Cummins-like elements in his evolutionist etiological theory, Buller (1998) argues that etiological theories must be regarded as specifications of Cummins' theory of functional analysis, and Walsh and Ariew (1996) take biological E-functions as instances of C-functions. Likewise, Paul Sheldon Davies (2000, 2001) has argued that etiological functions are in fact a kind of Cummins functions. Most authors primarily develop positions for the biological domain; only Davies claims that his 'natural norms' theory 'covers function attributions in the biological and the non-biological sciences'.[66] Hence, if technology is included in Davies' non-biological sciences, his theory can be taken as an example of a conjunctive CE-function theory for artefacts. It may by now be predictable what disadvantages we perceive for such a theory. From the C-theory, it inherits the problem that it need not satisfy the malfunctioning desideratum;[67] and from the E-theory, it inherits the problem that it cannot ascribe (the right) functions to innovative artefacts.

[63] Preston's (2003) argument consists of making plausible that the four desiderata we introduced in (Vermaas and Houkes 2003) are inconsistent, and that it is best to give up on the innovation desideratum. These four desiderata are essentially identical to the ones we set out in this book; we respond to this inconsistency claim by showing that the proposal in the next chapter satisfies these desiderata.

[64] Preston (2006).

[65] The I-theory also plays a role in biology, for instance, when authors try to unpack biological functions as functions that are ascribed by taking biological items as though they were artefacts; see section 6.4.

[66] Davies (2000, p. 100).

[67] Davies (2000, p. 94) admits this problem, but refuses to regard it as a fatal objection to his theory. He claims that etiological theories did not satisfy the malfunctioning desideratum in the first place, so that there is no harm in subordinating them to Cummins' functional analysis.

Chapter 4

The ICE-function theory

With our use-plan approach to artefact using and designing, and with the review of existing function theories, we have collected the means to formulate a function theory that is adequate to the technical domain. This theory is constructed against the background of, specifically, our use-plan analysis of designing, and incorporates elements from the intentional, causal-role and evolutionist function theories. For this reason, we call this central result of our work the ICE-function theory.

We start by summarising relevant parts of the use-plan approach and identifying what it has in common with the intentional, causal-role and evolutionist function theories. Then, in section 4.2, we present two closely related definitions of justifiable function ascriptions to artefacts, relative to use plans for those artefacts. In section 4.3 we show that these definitions, together with the use-plan approach itself, meet the proper-accidental, support and innovation desiderata listed in the introductory chapter. We also show that the ICE-theory meets the malfunctioning desideratum, albeit in a technical sense, which may be regarded as ultimately unsatisfactory. Therefore, we focus on this final desideratum in the next chapter, where we present how the ICE-function theory and the use-plan approach together can account more broadly for the phenomenon of artefact malfunctioning.

Our proposal does not capture all functional descriptions of artefacts. On the two central definitions, functions may only justifiably be ascribed to artefacts relative to use plans for these artefacts. In practice, agents can give functional descriptions of artefacts without considering such plans, and in some cases reconstructing them as ascribing functions relative to use plans seems contrived. This argument forces us to, in a final step, enrich our function theory with an account of plan-less, functional-role ascriptions to artefacts.

W. Houkes, P.E. Vermaas, *Technical Functions*, Philosophy of Engineering and Technology 1, DOI 10.1007/978-90-481-3900-2_4,

4.1 A use-plan approach towards functions

Having argued that existing function theories and more abstract basic theories do not meet the *proper-accidental*, *malfunctioning*, *support* and *innovation* desiderata, we now present our own proposal for a theory of technical functions. This proposal builds on our use-plan approach, specifically on the reconstruction of designing in section 2.4. On that reconstruction, a designer who develops a use plan, by which users can realise a particular goal, believes that the artefacts manipulated while executing this plan have specific physicochemical capacities that make the execution of the use plan successful. This is the *physical support* belief of the designer (step D.7 in the reconstruction as given in table 2.2 on page 29). On our function theory, the designer ascribes these capacities as functions to the artefacts involved. This makes our proposal primarily an intentional function theory. Function ascriptions to artefacts are determined by the beliefs of the designers of the use plans for the artefacts.[1] It explicitly characterises justifiable *ascriptions* of functions by agents on the basis of their beliefs and actions, instead of defining functions as *properties* that artefacts have independently of beliefs and actions. Yet by building on the use-plan approach, our theory also inherits elements that are distinctive for the causal-role and evolutionist function theories. This makes it a theory that incorporates parts of all three basic function theories. For this reason, we call our proposal the *ICE-function theory*.

Users may, of course, also ascribe functions to artefacts, but this is not a necessary element of their characterisations of artefacts. If they do ascribe functions, we take it that users ascribe as functions those physicochemical capacities that they believe to be responsible for the successful execution of the use plan for the artefacts. These user beliefs may be based on testimony by the designers of the use plan, meaning that function ascriptions by users may indirectly be determined by the beliefs of designers.

In this chapter, we present our function theory. We show how it incorporates and modifies elements from the intentional, causal-role and evolutionist function theories by identifying what these theories and our use-plan approach have in common. We argue that the disadvantages of the basic theories, identified in the previous chapter, may be overcome by appealing to use plans and to our reconstructions of using and designing. This does not immediately show the comparative merits of our proposal. At first, it only leads to a set of more refined characterisations of the roles that agents can play in their involvement with artefacts. It turns out that much work that is usually devoted to developing a theory of functions can instead be done by a theory about function-ascribing agents; this is one reason why we prefer the function-ascription formulation over the function-as-property formulation.

In section 4.2, we give our definitions of function ascriptions to artefacts relative to their use plans. There, we indicate again which elements of the basic theories are adopted — and adapted — in these definitions. Our theory is not

[1] A relation between functional descriptions of artefacts and plans can also be found in Wimsatt (1972, p. 38; table 1).

a straightforward conjunction or disjunction of the intentional, causal-role and evolutionist function theories; such a combination cannot be successful, given the discussion at the end of the previous chapter. In section 4.3 we evaluate our proposal by showing how it captures the function ascriptions of various types of agents and by showing how it meets the four desiderata we adopted for a philosophy of artefacts. Finally, we acknowledge that our definitions do not yet capture all functional descriptions of artefacts. In section 4.4, we enrich our proposal with a second, plan-less ascription of functional roles to artefacts.

First, we combine the results of the two previous chapters by explicitly connecting our use-plan approach to the three basic function theories.

The intentional element in the use-plan approach

In both the use-plan approach and intentional function theories, artefacts are primarily described in a teleological context. Artefacts are items that are considered by agents as means to realise goals. On our broad characterisation, given in section 3.2, intentional function theories are liberal about the purposes that artefacts may serve. This leads to a proliferation of functions and a lack of sensitivity for the limitations set by the support desideratum. The use-plan approach is more restrictive: on it, an artefact is only a means for an end if an agent believes that execution of a use plan for the artefact realises that end. The agent should, in other words, believe that executing a series of actions, including manipulations of the artefact, leads to the goal state. If, for instance, the agent is a user of the artefact, our reconstruction of use requires the agent to have in mind one or more existing use plans for the artefact, and to believe that executing one such plan can bring about a particular goal (steps U.2 and U.3 in the reconstruction of artefact use; see table 2.1 on page 23). If the agent who considers the artefact is a designer of the use plan, or possibly of the artefact itself, then he by definition believes that there is a use plan for the artefact. Yet on our reconstruction, the designer is also required to believe that this plan, when executed, leads to the goal he wants to contribute to (step D.5 in the reconstruction of designing; see table 2.2 on page 29). In section 2.4, we called this belief the *effectiveness belief* about a use plan of an artefact. In this chapter we refer to it more briefly as B_{eff}.

The requirements imposed by the use-plan approach do not completely avoid proliferation, but it puts up significant barriers. In particular, the requirements rule out cases in which agents associate an artefact with a goal without believing that there is a goal-realising use plan for the artefact. Hence, merely retaining a pencil for the purpose of time-travelling without believing that the pencil may be manipulated to enable travelling in time, is insufficient to describe the pencil as a device for time-travelling. Agents may, however, still associate artefacts with idiosyncratic or unsupported goals, as long as they have gone through the trouble of constructing a relevant use plan. Only further constraints rule out such cases as function ascriptions. To find these constraints, we turn to the causal-role and evolutionist elements of the use-plan approach.

The causal-role element in the use-plan approach

In the causal-role function theory, functional descriptions of artefacts should be supported by beliefs that are true or at least justified relative to an account A. In the use-plan approach, we incorporate this support only in the weaker sense of justification. Function-ascribing agents should believe that there is a use plan relative to which a function can be ascribed and they should believe that executing this plan realises the goal. The standard of belief-consistency on plans, discussed in section 2.6, may then be invoked to require that agents justify these beliefs. A designer d of a use plan can trivially justify his belief that a plan exists, but should still be able to justify his belief B_{eff} that the plan leads to its goals, on the basis of some account A. Conversely, a user u may justify his belief in the existence of a use plan and his effectiveness belief B_{eff} on the basis of testimony. A user may also justify B_{eff} on the basis of an account A, but in our function theory we capture this last case in a different way than the testimony-based case.[2] In particular, we introduce different roles besides that of 'designer'. Agents who consider an existing use plan and can justify B_{eff} themselves, are called *justifiers* 'j' of the use plan.[3] Users who are epistemically more passive and rely on testimony are called passive users, denoted with a 'u'. 'User' refers to both justifiers and passive users, unless noted otherwise.

We do not require that B_{eff} and other beliefs are true, but only justified, in order to leave room for unsuccessful executions of the use plan. If the effectiveness belief B_{eff} about use plans should be true, then (correctly) executing a use plan must lead to realisation of the goal state. Thus, there cannot be cases in which the artefact turns out not to work. If, however, B_{eff} is merely justified on the basis of an account or of testimony, use plans may fail and artefacts may malfunction.

The requirement of justification rules out the association of artefacts with unsupported goals. Since designers d and justifiers j have to justify their effectiveness beliefs B_{eff} on the basis of some account A, there must be a collectively accepted measure of support for the belief that an artefact can be used successfully to realise a goal. Since passive users u have to justify their beliefs B_{eff} on the basis of testimony, they have at least indirect support for their effectiveness beliefs. Hence, an owner of a 'Subterra car' cannot ascribe to it a role in subterranean transportation by believing that there is a corresponding use plan for the vehicle, and that this plan works. In the use-plan approach, this agent should justify these beliefs, including an effectiveness belief, by appealing to either testimony or some account.

---

[2]Invoking both testimony and other sources of evidence for the effectiveness belief does not mean that we accept a non-reductionist view of testimony. Testimony might be reducible to, for instance, experiences of successful use acquired by designers while testing the artefact. Thus, we may choose to remain neutral on the much-debated issue of reducibility of testimony; see, e.g., Lipton (1998), Lackey (2003) and Lackey and Sosa (2006) for overviews of and some recent contributions to this debate.

[3]Designers d of a use plan are also agents who justify B_{eff} for this plan. We therefore could take them as justifiers as well. Typically, we take only (non-passive) users to be justifiers.

The evolutionist element in the use-plan approach

The use-plan approach and the evolutionist function theory both offer historical perspectives on artefacts, which provide support for functional claims. In the evolutionist function theory, this historical description of artefacts is grounded on some technological counterpart of neo-Darwinian evolutionary theory. This supports claims about technical functions by referring to the predecessors of the artefact: these have the capacity corresponding to the function, making it likely that the present member of the family also has the capacity. In the use-plan approach, the relevant history of artefacts concerns the *communication* of the use plans in which the artefacts play a role. This communication originates with the designers of the use plans (section 2.4). Communication of the plans provides other agents with support for their beliefs that these use plans exist and that at least some agents — ultimately the designers — have justified effectiveness beliefs B_{eff} about the plans. By accepting these other agents as reliable testifiers, agents may adopt these effectiveness beliefs B_{eff} as well, and justify them on the basis of testimony. Minimally, this history of communication involves the designer of the use plan, who informs one other agent, who is the only prospective user. Alternatively, there may be a whole communication tree, with branches consisting of series of agents passing on the use plan. From the perspective of a user who looks back towards the source of the use plan, this history looks like a *communication chain* of agents who have informed one another about the plans, and who provided or passed on support for each other's effectiveness beliefs.

The requirement that the use plans of artefacts are communicated to users again raises the threshold for associating goals with artefacts. It rules out the association of artefacts with idiosyncratic goals. Designers may have all kinds of private goals when developing new use plans, such as paying next month's rent with the money they earn. But typically, they do not communicate these goals to the users. A new detergent is typically presented to users with a plan for cleaning clothes, not as a means for earning its designers or manufacturers money or as a means for capturing a larger market share.[4]

A detailed discussion of the communication chains for use plans allows us to collect more elements of our use-plan approach that are needed to formulate our theory of function ascriptions.

Communication chains for use plans

We first consider such present-day artefacts as computers and cellular phones. Typically, the communication chain by which a user is informed about the use plans for these artefacts has a well-defined starting point, namely the design

[4]The detergent can, nevertheless, be seen as embedded in a use plan with the goal of capturing a larger market share. Commercial strategists may design such a use plan. This second use plan is typically not communicated to those who eventually use the detergent to clean their laundry, meaning that these agents have no reason — or at least not one based on direct plan communication — to associate the detergent with this commercial goal. If agents who typically use the detergent nevertheless do consider the detergent relative to its capture a larger market share-use plan, they do so as *observers* of the detergent (in section 4.4 we introduce this agent role of observer with respect to artefacts) rather than as its *users*.

process of these artefacts and of their use plans. We require the designers d to have the belief B_{eff} that the use plan leads to its goals, a belief that is justified on the basis of some account A. Furthermore, the designers have to communicate the use plan to prospective users u and in some way or another express their effectiveness belief B_{eff}. By accepting this communication as veridical, other agents may appeal to testimony for the existence of the use plan and for the fact that the designers believe that it works; if, moreover, the other agents accept the designers as reliable testifiers, they may adopt the belief B_{eff} themselves, and may justify it by appealing to testimony. In this case, the other agents play the role of passive users.

The communication chain need not end here. Users typically communicate the use plan to other prospective users,[5] in one of two ways. Firstly, the user may be able to justify the effectiveness belief B_{eff} herself, because she has had hands-on experience that the use plan leads to its goals or because she can present some account A, in analogy to the designer. The user is then a justifier j of the use plan, and can communicate both the use plan and her justified belief B_{eff} to another user u'. By accepting this communication as veridical, the other user u' may, in turn, appeal to testimony for the existence of the use plan and for the fact that the justifier j has the belief B_{eff}; if, moreover, the other user u' accepts j as a reliable testifier, u' may adopt the belief B_{eff} and may justify it by appealing to testimony. Hence, justifiers epistemically screen off designers. Alternatively, the initial user may simply pass on to another user u' the use plan, which was communicated to him by a designer d or an earlier justifier j. In this case, the user u is a passive user and there is no epistemic screening-off.

This model may not apply to highly traditional artefacts such as nails, brooms or flour. Here, the communication chain may not have a well-defined starting point. It seems that the use plans for such artefacts lack true designers, but simply arose in ancient or prehistoric times. Still, these use plans are communicated from one user to another. Moreover, on pain of collective irrationality, some of those users will have had justified effectiveness beliefs B_{eff} for the communicated plans; hands-on experience that a use plan leads to its goals suffices. Therefore, these users may have acted as justifiers j of the plans. These justifiers screen off earlier justifications of the effectiveness beliefs. This means that ignorance about the historical origin of the use plans of traditional artefacts is not a problem for the justification of function ascriptions: for epistemic purposes and purposes of rationality, the communication chain terminates at the most recent justifier, and later users may appeal to their testimony.[6] Hence,

[5]On the use-plan analysis, users are not required to communicate use plans to others. In practice they regularly do. In effect, most people receive the information about the workings of their cups, hammers, dish-washers and cars not from designers but from fellow users.

[6]Similarly, justifiers may compensate for bad designing, or defeat and update justifications by designers. It is possible that designers communicate a use plan without having adequate justification for its effectiveness. Although such designing must be assessed negatively on our reconstruction, this evaluation need not extend to all later use. Users may, perhaps gradually, acquire sufficient evidence for the effectiveness of the use plan and act as justifiers for later users. These justifiers then screen off the incompetence or ignorance of the designers, and the justifiers' testimony grounds rational use.

in the use-plan approach, all artefacts have use plans that are communicated and there are always, on pain of irrational use, some agents in the communication chains who have justified the plan. For use plans that have their origin in designing this chain can be captured as, for instance:

$$d \Rightarrow u \Rightarrow u \Rightarrow j \Rightarrow u \Rightarrow j \Rightarrow u \ldots,$$

Here, the '$d \Rightarrow$' represents the designer d communicating the use plan and testifying for his belief B_{eff} that the plan works, '$j \Rightarrow$' represents a justifier communicating the use plan and giving testimony for its effectiveness, and '$u \Rightarrow$' represents a passive user who may appeal to the testimony of a justifier or designer and who communicates the use plan. For use plans that do not have a clear origin in designing, the communication chain can be captured as, for instance:

$$\ldots u \Rightarrow u \Rightarrow j \Rightarrow u \Rightarrow u \Rightarrow j \Rightarrow u \Rightarrow u \Rightarrow u \ldots.$$

Relative to this analysis of communication chains the roles of designers, justifiers and passive users can be characterised more precisely by the conditions in table 4.1.[7] These conditions are labelled 'I', 'C' and 'E', in order to bring out the similarities with the intentional, causal-role or evolutionist function theories.[8]

This characterisation allows us to draw a conclusion that is useful later on: any user u or justifier j who is part of the communication chain may appeal to testimony by an earlier justifier j' or by a designer d for the existence of the use plan and for the belief B_{eff} that this plan works.

The sketched communication history is not an evolutionary history in the neo-Darwinian sense. There is, for instance, no distinction between a genotype and phenotype, nor do we assume that the communication allows for mutations and selection. For this reason, the function theory proposed in the next section is not an evolutionist function theory; it merely incorporates some elements of that basic theory.

The three roles of designing, justifying and passive using encompass many of the involvements with artefacts, but the typology is not exhaustive: in section 4.4, we introduce two additional roles. Before doing so, we characterise how agents, when playing one of the three roles defined so far, may justifiably ascribe functions to artefacts. This is not an arbitrary way of emptying the stage. As will become clear in the remainder of this chapter, differences in roles may reflect differences in the type of functional descriptions that agents give of artefacts.

[7] The characterisation of justifiers is recursive, and makes sense for communication chains with a well-defined starting point in designing. A designer d then counts as the first justifier. For communication chains with an infinite regress, especially the characterisation of users seems extravagant. One can come up with examples of agents who satisfy two conditions but not the third. But for agents of flesh and blood, satisfying the E-condition typically implies that they satisfy the I- and C-conditions as well.

[8] This labelling should be taken with a grain of salt. The condition that a user or justifier should ground in testimony her belief that the designer or (another) justifier has the belief B_{eff}, originates from the evolutionist function theory. We nevertheless labelled it with a 'C' in order to group all justification conditions together.

An agent is a designer d of the use plan p for an item x, iff:

D. d has developed p with the aim that the execution of the actions of p by other agents leads to the goals of p;
d has intentionally selected x as an item that is to be manipulated as part of the execution of p;
I. d has the belief B_{eff} that the execution of p leads to its goals;
C. d can justify B_{eff} on the basis of an account A; and
E. d communicates p and testifies to his belief B_{eff} to other agents.

An agent is a justifier j of the use plan p for an item x, iff:

I. j has the belief B_{eff} that the execution of p leads to its goals;
j believes that a designer d or another justifier j' of p has B_{eff};
C. j can justify B_{eff} on the basis of an account A;
j can justify on the basis of testimony T that d/j' has B_{eff}; and
E. j received p and T that d/j' has B_{eff}.

An agent is a passive user u of the use plan p for an item x, iff:

I. u has the belief B_{eff} that the execution of p leads to its goals;
u believes that a designer d or a justifier j of p has B_{eff};
C. u can justify B_{eff} on the basis of testimony T;
u can justify on the basis of testimony T that d/j has B_{eff}; and
E. u received p and T that d/j has B_{eff}.

Table 4.1: Roles of agents in use-plan communication chains

4.2 Function ascriptions

In this section, we propose our theory of technical functions or, more precisely, our theory of *justifiable function ascriptions* to artefacts. The definitions of the agent roles of designing, justifying and passive using serve as preliminaries for this theory, because they make clear *who* ascribe functions to artefacts. Our theory offers criteria for function ascriptions, as they are made by agents playing one of the roles.

Functions as physicochemical capacities

Another preliminary issue concerns *what* is ascribed in our theory. What exactly do agents say about artefacts in ascribing a function to ϕ? What is 'to ϕ'? The review given in the previous chapter yielded that, in current theories, a technical function may correspond either to a goal that is intended to be achieved by manipulating the artefact, or to a physicochemical capacity of the artefact by which it contributes to realising this goal. In our theory, we take the latter position: ascribing a function to an artefact implies ascribing to it a physicochemical capacity, such as 'to emit light', 'to remove dirt from laundry' or 'to prevent blood clots'. Thus, throughout this book, the act of ascribing a function is described in full as 'ascribing to an artefact the capacity to ϕ as a function'. If this exact formulation would lead to needlessly complicated or confusing sentences, we use the simplified 'ascribing a function to ϕ'.

One advantage of this position is that we need not describe all cases of unsuccessful artefact use as cases in which artefacts malfunction. Consider, for instance, a laser pen which is designed to emit a green dot of light for letting agents point out items on projected slides during presentations. The use of this laser pen can be unsuccessful when the room in which the presentation takes place is too brightly lit, even though the pen still emits the green dot of light. An accurate description of this lack of success seems to be that the pen performs its function as expected, but that the physical circumstances prevent realisation of the goal of pointing out items on the slides. This description can be accommodated by function theories that differentiate between the goals and functions associated with artefacts, but not by theories in which those goals and functions coincide.

We take capacities to be typically, though not exclusively, dispositions for which, following Stephen Mumford (1998, chs. 3–4), an analysis can be given in terms of subjunctive conditionals that hold true under certain ideal conditions.[9] The capacity corresponding to the laser pen's function is, for instance, more precisely described as 'to emit light when force is applied to the pen's switch'. And the capacity 'to remove dirt from laundry' may be spelled out as, say, 'to dissolve grease when brought into contact with grease and water'. Sometimes, however, capacities that may be ascribed as functions seem to resist this conditional analysis. Take, for instance, the capacity of cobalt-60 to emit gamma radiation, which is used in agricultural breeding to speed up the rate of genetic mutations in crops. This emittance seems a property cobalt-60 simply has, independently of the realisation of any antecedent.

In the conditional analysis of the capacity 'to remove dirt from laundry' the description of the antecedent as 'when brought into contact with grease and water' may be taken as fairly complete. The description of the antecedent 'when force is applied to the pen's switch' is clearly only partial. Even under ideal circumstances, a charged battery is another necessary element for the pen to emit light. We accept that, in the conditional analysis of the physicochemical capacities that correspond to technical functions, the antecedents may be incomplete; we take the remaining relevant antecedents as part of the normal conditions. In this we follow Bell, Snooke, and Price (2005) who, within engineering, model functions in electrical systems by means of *triggers* and *effects* and thus focus pragmatically on only those antecedents and consequences of capacities of systems that are relevant to their use. Triggers can include on/off-positions of switches, since these are manipulated by users, but typically do not include the electrical voltages within the systems, since these voltages are assumed to exist in electrical systems by default.[10]

[9]Because of the inclusion of these ideal conditions, Mumford speaks about an analysis in terms of *conditional conditionals* (1998, p. 88).

[10]It should be acknowledged that not all authors in engineering ignore default antecedents when modelling technical functions. In, for instance, the modelling of functions as proposed by Stone and Wood (2000), the antecedents of the conditionals representing functions contain all material, energy and information flows that are needed to let an artefact produce the

'Folk-physics' descriptions of capacities

The examples of capacities given above bring us to another preliminary issue. In our theory, we require that the functions ascribed to artefacts are *physicochemical* capacities. This does not immediately fit all functional descriptions. 'To emit light' clearly singles out a physicochemical capacity of, say, of a laser pen, but 'to remove dirt from laundry' and 'to prevent blood clots', which may be ascribed to detergents and Aspirin respectively, seem too colloquial and too immediately goal-related to single out physicochemical properties or dispositions of the artefacts concerned. Terms like 'dirt' and 'prevention' refer more or less directly to preferences and goals of agents.

To make the colloquial formulations fit our theory, we take them to be coarse-grained descriptions of capacities, as they are typically given by artefact users. 'To remove dirt from laundry', for instance, describes a physicochemical capacity of the detergent, which is described in a more sophisticated way as 'to bind specific substances and to let it dissolve in water'. Users are typically not informed about or interested in the physical and chemical details that make artefacts appropriate means to their goals. Therefore, they may give coarse-grained or 'folk physics' descriptions of the physicochemical capacities of these artefacts, including those that correspond to their functions. Designers of artefacts or agents who are informed about the physicochemical details, are typically able to characterise a capacity in physicochemical terms. One of the functions of, for instance, Aspirin is characterised in every-day terms as the capacity to prevent blood clots; in physicochemical terms, the same function is 'to acetylate the enzyme cyclo-oxygenase in thrombocytes'. So, if agents ascribe to Aspirin the function 'to prevent blood clots', strictly speaking they ascribe the physicochemical capacity to acetylate the enzyme cyclo-oxygenase in thrombocytes.[11]

Singling out the right capacities

The laser pen example can be used to introduce our theory of justifiable function ascriptions to artefacts. On our use-plan approach, there is a use plan for this pen, which includes manipulations of the pen, and which leads to the goal of pointing out items on projected slides. By our definitions of agent roles, the designers of this use plan, and its possible justifiers, have the effectiveness belief B_{eff} that this plan, when executed, leads to the pointing out of items on slides. Furthermore, these agents can justify B_{eff} on the basis of some account A. Passive users also have the effectiveness belief B_{eff} and believe that designers

relevant consequences. This diversity in positions and approaches within engineering sciences again illustrates the difficulty of finding unambiguous concepts on the basis of an analysis of engineering alone.

[11] Our requirement that the functions ascribed to artefacts are physicochemical capacities has consequences. Recall the example of a detergent that is, in part, designed to let a company capture a larger market share. Suppose one claims that the detergent is 'for being bought by consumers'. Since this is not a physicochemical capacity of the detergent, the claim does not amount to a function ascription on our theory. Still, it may be understood as the ascription of an *economic* or *social* function to the detergent. Our theory does not deal with such functions, but only with the ascription of *technical* functions.

and/or justifiers of the plan have B_{eff}; passive users justify both beliefs by appealing to testimony. Given our resolution of the last preliminary issue, the function ascribed to the laser pen is its physicochemical capacity to emit a green dot of light. Hence, our theory needs to single out this capacity of the pen. The effectiveness belief B_{eff} and its justification provide a natural way to do so.

For a designer, this effectiveness belief is based on beliefs about the physicochemical capacities of the pen. On our reconstruction of the design process (see section 2.4), a designer has the belief that the designed use plan is effective (step D.5 in that reconstruction) and the more specific belief (that is, the *physical support* belief required in step D.7) that this plan is effective because the artefacts have specific physicochemical capacities $\{\phi_{n,1}, \phi_{n,2}, \ldots\}$. The design process is only rational — in particular, belief-consistent — if the designer is able to justify this physical-support belief. Therefore, a designer of the laser pen must be able to justify that it has specific physicochemical capacities, and that these capacities explain why the pen may be used effectively to point out items on projected slides. For the laser-pan it is its physicochemical capacity to emit a green dot of light that features in this physical support belief. Hence, for designers, our theory of justifiable function ascriptions can single out the physicochemical capacities that designers ascribe to artefacts as functions: these are the capacities featured in the physical support beliefs of designers.

Our reconstruction of use (section 2.3) does not provide similar means for singling out physicochemical capacities. Yet the definitions of the justifier and passive-user roles do. A justifier is required to have an effectiveness belief B_{eff} about the use plan concerned, and to justify this belief. This effectiveness belief B_{eff} need not be the physical-support belief that designers have; a justifier may believe on the basis of experience that executing the use plan leads to its goals. Still, we require that, *if a justifier ascribes functions*, her effectiveness belief B_{eff} is based on a justified physical-support belief, like the effectiveness beliefs of designers.[12] Hence, our theory can single out the physicochemical capacities that justifiers ascribe to artefacts in their function ascriptions: again, they are capacities referred to in physical-support beliefs of the justifiers themselves.

Similarly, we require that, if a passive user ascribes functions to artefacts, his effectiveness belief B_{eff} about the use plan of these artefacts is based on a justified physical-support belief, and that his belief that the designers and/or justifiers of the plan have B_{eff} is based on the belief that these designers and/or justifiers have the physical support belief. The passive user may justify all of these beliefs by appealing to testimony. Hence, for passive users, our theory can single out the physicochemical capacities of artefacts that passive users ascribe to artefacts in their function ascriptions: again, these are the capacities that feature in physical-support beliefs. However, passive users typically give coarse-grained descriptions of these capacities, and ground their beliefs in testimony.

[12] The physical-support beliefs of a justifier may be identical to the physical-support beliefs of a designer, but need not be: a justifier may update the physical-support beliefs of the original designer of the plan (see also footnote 6 on page 82).

The definitions of function ascriptions

Because of the way we resolved the preliminary issues, we can give a uniform account of function ascriptions by various agents. If a designer, justifier or passive user ascribes to an artefact a function to ϕ, he believes that the artefact has a physicochemical capacity to ϕ and that this capacity contributes to the effectiveness of the use plan for the artefact; moreover, he can justify this physical support belief and thereby justify the effectiveness belief B_{eff}. To return to the example: if agents ascribe to a laser pen the function to emit a green dot, they justifiably believe that the pen has the capacity to emit a green dot and that this capacity makes a use plan for the pen effective in pointing out items on slides. The means by which the physical-support belief is justified is different for the various agent roles. Designers and justifiers of the pen's use plan primarily rely on some explanatory account A, whereas passive users of the pen rely on testimony. In our definitions of function ascriptions, we therefore differentiate between these agent roles, to indicate more precisely the epistemic sources of function ascriptions. Still, our general characterisation of justifiable function ascriptions already allows the conclusion that all functions are ascribed to artefacts relative to use plans for these artefacts.

Let B_{cap} refer to the *capacity belief* that the artefact has a physicochemical capacity to ϕ, and let B_{con} refer to the *contribution belief* that a physicochemical capacity to ϕ of an artefact contributes to realising the goal of its use plan. In combination, these beliefs constitute part of the physical-support belief about the use plan, i.e., the combination of B_{cap} and B_{con} is the belief that to ϕ is a physicochemical capacity of the artefact that, in part, explains the effectiveness of the use plan. Our definition of a justifiable function ascription by a designer or justifier of a use plan is then:

Function ascriptions by designers or justifiers:

A designer d or justifier j justifiably ascribes the physicochemical capacity to ϕ as a function to an artefact x relative to a use plan p for x, and relative to account A, iff:

I. d/j has the belief B_{cap} that x has the capacity to ϕ;
d/j has the belief B_{con} that p leads to its goals due to, in part, x's capacity to ϕ; and

C. d/j can justify B_{cap} and B_{con} on the basis of A.

The conditions in this definition — we have again labelled them 'I' and 'C' to bring out similarities with the basic function theories — are additional to the use-plan conditions on being a designer or a justifier (see table 4.1). Hence, both designers and justifiers should still communicate a use plan to users u and provide testimony for the belief B_{eff} that this plan works. If the effectiveness belief for which testimony is given includes the capacity and contribution beliefs B_{cap} and B_{con}, and if a user accepts this testimony, the user may believe that the designer or justifier has some beliefs B_{cap} and B_{con}, and the user may justifiably have these beliefs himself. In this way, passive users may spell out their testimonially justified belief B_{eff} that a plan works because of the physicochemical capacities of artefacts. If users spell out this belief, they ascribe a function to the artefact; if they do not, they do not ascribe functions. Our definition of a justifiable function ascription by a passive user reads:

Function ascriptions by passive users:

A passive user u justifiably ascribes the physicochemical capacity to ϕ as a function to an artefact x relative to a use plan p for x, and relative to testimony T, iff:

I. u has the belief B_{cap} that x has the capacity to ϕ;
u has the belief B_{con} that p leads to its goals due to, in part, x's capacity to ϕ;
u believes that a designer d or justifier j of p has B_{cap} and B_{con};

C. u can justify B_{cap} and B_{con} on the basis of T;
u can justify on the basis of T that d/j has B_{cap} and B_{con}; and

E. u received T that d/j has B_{cap} and B_{con}.

These definitions, like the characterisations of the different agent roles, have several elements in common with the intentional, causal-role and evolutionist function theories. Therefore, we call our theory of function ascriptions the ICE-function theory.[13]

Our proposal consists of *two* definitions of justifiable function ascriptions. Yet it does not lead to (intra-domain) pluralism. The two definitions make explicit the differences between function ascriptions by designers and justifiers and by passive users, but they still present a single account for function ascriptions. The differences are in the epistemic details. While these should not be underestimated, they do not divide our theory into two strictly separated parts. In the next section, when we discuss how function ascriptions by users — justifiers as well as passive users — satisfy the ICE-theory, it will become evident that in some cases it can be difficult to decide which of the definitions best suits the ascription. Moreover, the two definitions may be merged into a general one that applies to designers, justifiers and passive users alike. This general definition could, for instance, be generated from the first definition of justifiable function ascriptions by designers and justifiers, if in that first definition the account A is broadened to include testimony that the designer d or a justifier j' had the relevant beliefs B_{cap} and B_{con}. Since this general definition does not bring out the evolutionist element in our theory, and since the distinctions between the roles of function-ascribing agents prove useful to our analysis, we prefer the two-definition (designer/justifier and passive-user) form of the theory.

[13] In (Vermaas and Houkes 2006a) we presented a predecessor of our proposal, which may be taken as a conjunction of improved elements from the intentional, causal-role and evolutionist function theories. The current proposal is not that easily taken as a conjunction of elements of those function theories. The C-conditions are, for instance, rather a disjunction: agents may justify their beliefs by either explanations on the basis of an account A (an element of the causal-role theory) *or* by testimony T by other agents (an element of the evolutionist theory).

4.3 Assessing the function ascriptions

To illustrate our theory, we apply it to various function ascriptions. Firstly, we review function ascriptions by agents in their roles of designing, justifying and using. Then, we evaluate the ICE-function theory on the basis of the desiderata for a theory of artefacts. In the next section, we discuss possible limitations of our proposal by considering functional descriptions by agents other than designers, justifiers and users.

Function ascriptions by designers

In our characterisation of designing, we distinguished various types: product designing was contrasted with designing that consisted solely of constructing and communicating new use plans, expert designing employing scientific and technological knowledge was distinguished from designing on the basis of common-sense and elementary physical knowledge, and amateur designing was set apart from professional designing. The agents involved in these types of designing may all ascribe functions to the artefacts relative to the use plans they developed for these artefacts. On our reconstruction of designing (step D.7 in table 2.2, page 29), these agents have the physical support beliefs B_{cap} that the artefacts $x_1, x_2, \ldots$ have physicochemical capacities $\{\phi_{1,1}, \phi_{1,2}, \ldots\}$, $\{\phi_{2,1}, \phi_{2,2}, \ldots\}$, etc., and they have the contribution beliefs B_{con} that these capacities make successful execution of the developed use plans possible. And on this reconstruction, designing agents are able to justify these beliefs. They can therefore ascribe functions to the artefacts in accordance with our definition of function ascriptions by designers and justifiers. This holds for all types of designing agents; what differs between them is that they may use different accounts to justify their beliefs B_{cap} and B_{con}. In common-sense designing, designers ascribe functions relative to their experience with the artefacts and typically characterise the capacities corresponding to these functions in a colloquial way; expert designers ascribe functions relative to their scientific and technological knowledge and typically give sophisticated, physicochemical descriptions of the capacities corresponding to the functions.

Consider, for instance, an amateur designer who has developed a new use plan for an existing object on the basis of common-sense knowledge. An example is the person who developed the use of brown paper for removing candle wax from clothes; here, the developed use plan has the goal of removing wax from clothes, and consists of the following actions: put the brown paper on top of the wax, briefly place a hot iron on the paper, and then remove the brown paper. The agent who constructed this plan may ascribe to the paper the function to absorb candle wax. This ascription is not just a reformulation of the goal of the use plan: it implies that the use plan is effective because of the paper's absorbance, instead of, say, a chemical reaction between the wax and the paper. Spelling this out, it implies that the designer believes that the paper has the physicochemical capacity to absorb candle wax (B_{cap}) and that this capacity contributes to achieving the goal of wax-free clothes (B_{con}). Moreover, the ascription implies — and we here take a more normative stance — that this agent has evidence that supports these beliefs. This evidence may consist of, say, direct experience that after the ironing the paper contains stains of candle wax, which supports the belief that the paper indeed absorbed the wax, rather than stains of other substances, which would support the alternative belief that the wax was removed by some chemical reaction between the wax and (components of) the paper. Hence, this function ascription satisfies our definition of function ascriptions by designers and justifiers: the agent has the requisite justified beliefs B_{cap} and B_{con}.

By contrast, consider professional designers who have constructed new use plans for existing objects on the basis of detailed scientific or technological knowledge. An example would be the pharmacologists who constructed the new use plan for Aspirin that has the goal of preventing blood clots in cardiac patients. These researchers can ascribe to Aspirin the function to acetylate the enzyme cyclo-oxygenase in thrombocytes relative to this use plan, since they believe that Aspirin has this biochemical capacity (B_{cap}) and that this capacity contributes to preventing blood clots (B_{con}). Moreover, these researchers can explain their effectiveness belief for the new use plan of Aspirin through an account A that consists of scientific and (bio-)technological knowledge, and that explains their beliefs B_{cap} and B_{con}. Hence, this function ascription also satisfies our definition of function ascriptions by designers and justifiers. Function ascriptions by professional designers, who construct new use plans for newly produced artefacts on the basis of expert knowledge, satisfy this definition in an analogous way: for an illustration, just replace the Aspirin in the last example with a hypothetical, newly developed drug that has the same capacity and a similar use plan.

Function ascriptions by passive users and justifiers

The ICE-function theory accommodates user function ascriptions in two ways — through the designer/justifier definition and through the passive-user definition. Users learn about use plans for artefacts from designers or justifiers, via the communication chains discussed earlier. Via these chains, users do not just acquire information about use plans, but also gain (testimonial) evidence for believing that the plans work, and for believing that the artefacts have capacities that explain the effectiveness of the plans. Consider, for instance, a VCR. Users typically learn about the use plan of this appliance through verbal communication in shops or with fellow users, or through texts and drawings in manuals. This communication typically includes at least some colloquial descriptions of the capacity of the VCR that explains the effectiveness of use plans for recording and playing television programmes; in particular, the VCR has the, in this case, electromagnetic and mechanical capacity to store a television signal on tape and the electromagnetic and mechanical capacity to transform the resulting record on the tape back into a television signal. On the basis of these communications, users may ascribe these capacities, or more colloquial descriptions thereof, as functions to the VCR.

These function ascriptions satisfy our definitions in the following ways.

When a user just acquired the VCR and does not have any experience with using it, his function ascription implies that he believes on the basis of the testimony of designers — the product designers at the company that made the VCR — or justifiers — shopkeepers or experienced fellow users — that the VCR can store a television signal and can transform the resulting record back into a signal (B_{cap}), and that these capacities contribute to the recording and playing of television shows (B_{con}). Furthermore, the user should be able to justify his belief that these designers or justifiers themselves believe B_{cap} and B_{con} on the basis of his experience with those designers and justifiers, e.g., that the company that made the VCR always produces good products, implying that the designers that developed the use plan and the VCR are knowledgeable; that the shop where the VCR was bought is a reliable one; or that the fellow user who explained the working of the VCR has hands-on experience with it. If the inexperienced user has such justified beliefs B_{cap} and B_{con}, his function ascription to the VCR satisfies our passive-user definition.

Now suppose the user has acquired more experience with an artefact, say, because she has successfully carried out a use plan for the artefact, or because the communication of the use plan also involved user training. Then, the user may start to justify the beliefs required by function ascriptions on other grounds than testimony. When an experienced user of a VCR, and a trained operator of a drilling rig, ascribe the usual functions to these artefacts, they have the requisite beliefs B_{cap} and B_{con} and can justify these beliefs on the basis of their experience and training, respectively. In our terminology, such agents are no longer passive users who rely exclusively on testimony, but justifiers of the use plan. The account A by which such users justify their beliefs B_{cap} and B_{con} may consist of only experience, as is typically the case when users ascribe functions to house-hold appliances such as VCRs. Yet it may also consist of expert knowledge, as when users of more professional equipment, such as drilling rigs, ascribe functions to these artefacts.

In real life, it may not be possible, let alone necessary, to distinguish these two ways of ascribing functions to artefacts. Indeed, there is no sharp distinction between inexperienced users who rely on testimony given by designers and justifiers, semi-experienced users who base their function ascriptions on both testimony and on their independent justification for beliefs B_{cap} and B_{con}, and experienced users who may communicate the use plans of artefacts to others and give testimony that they believe B_{cap} and B_{con} on the basis of their experience. Still, there is some reason to distinguish function ascriptions of experienced users from those of professional, expert designers. For instance, when an experienced user is forced to spell out how the artefact works, she may still appeal to the testimony of product designers of the artefact, perhaps through justifiers such as shopkeepers and professional instructors. The distinction between purely experiential evidence for the effectiveness of a use plan and a detailed physicochemical explanation may prompt this retreat.

The desiderata

In table 1.1 on page 5, we formulated four desiderata for a theory of artefacts. We are now in the position to show that our use-plan account of using and designing, together with our ICE-theory of function ascriptions, satisfies these desiderata. For the proper-accidental, support and innovation desiderata, our arguments are fairly brief; for the malfunctioning desideratum, the full argument requires more space.

On the basis of our use-plan approach, we can already make a distinction between proper and improper use (see section 2.7). Proper use is the execution of a use plan that is accepted within a certain community; improper use is the execution of a use plan that is socially disapproved. More specifically, designers who are acknowledged as professionals in a community typically develop socially acceptable use plans. There may be other mechanisms of social acceptance of plans; after all, use plans that are the result of amateur, common-sense designing are often socially accepted as well. Still, because there is a distinction between passive users and professional designers, and a mechanism for the social acceptance of some plans over others, a distinction may be made between proper and improper use plans. This distinction may be combined with the ICE-theory to define *proper function ascriptions* as function ascriptions, by designers, justifiers and passive users, relative to proper use plans, and to define *improper function ascriptions* as function ascriptions relative to improper use plans. On this definition, proper function ascriptions are enduring, because the acceptance of the underlying use plans is socially stable; conversely, improper function ascriptions are transient because the underlying use plans lack stability and social entrenchment. In this way, our philosophy of artefacts accommodates the *proper-accidental desideratum.* It does so in a slightly different way than some existing function theories: we make a distinction between proper and improper *function ascriptions*, on the basis of plans and roles played by agents, and not between the ascription of proper and improper *functions.*

Our description of artefacts also meets the *innovation desideratum.* The historical perspective required to ascribe ICE-functions may be limited to the design process; it need not extend to earlier generations of artefacts. An artefact can therefore straightaway be ascribed the capacity for which designers selected it, even if the artefact is a completely novel one (the case of the first nuclear plant) or if it is selected for a capacity for which it was not selected before (the case of Aspirin selected for preventing blood clots).

Furthermore, the *support desideratum* is met. This is in part achieved by construction, since we require that the functions ascribed to artefacts are physicochemical capacities. Nevertheless, the C-conditions in the two definitions of function ascriptions ensure that the ascribing agent justifiably believes that the artefact has these physicochemical capacities. The relevant evidence may be experience that the artefact has the capacity, testimony by other agents, or scientific or technological knowledge; in all cases, the evidence supports the function ascription by supporting the beliefs that the artefact has the corresponding physicochemical capacity and that this capacity explains, in part, the effectiveness of a use plan. Still, a scientific basis for a function ascription, testimony and even personal experience leave open the possibility that, contrary to reasonable expectations, the artefact lacks the capacity.

This last observation shows how the ICE-theory can satisfy the *malfunctioning desideratum*: by the two definitions, functions may be ascribed to artefacts when the artefact in fact lacks the ascribed physicochemical capacity. In those cases, the artefact may be said to malfunction. Yet there is a catch, which suggests that we satisfy the malfunctioning desideratum only in a limited sense. On our theory, agents cannot ascribe a function to an artefact while they believe that the artefact in question lacks the corresponding physicochemical capacity. By the I-conditions of both of our definitions, the ascribing agent should have the belief B_{cap} that the artefact has the capacity corresponding to the function ascribed. Therefore, an agent can only ascribe an ICE-function to an artefact that lacks the corresponding capacity, as long as he reasonably believes that it has this capacity. So, an artefact that worked up to a certain time may be ascribed its usual function, even though it actually broke down after that point — as long as the ascribing agent is unaware of its present state. Similarly, an artefact (say, the prototype of some new product) that is justifiably thought to have a capacity on the basis of some account, but is not yet tested for having this capacity, may be ascribed this capacity as its function, although it does not in fact have it. But once an agent believes that the artefact lacks the capacity on the basis of unsuccessful use or tests, it seems that this capacity cannot be ascribed as an ICE-function anymore.

Such 'known-to-be-broken artefacts' apparently provide counterexamples to the ICE-theory; our theory cannot ascribe the right proper functions to such malfunctioning artefacts. In the next chapter we argue that the ICE-theory satisfies the *malfunctioning desideratum* in a broader sense, i.e., that it can account for more function ascriptions to malfunctioning artefacts than the 'not-yet-known-to-be-broken' cases accounted for so far.

4.4 Functional roles

Our theory of function ascriptions is, so far, built around three roles that agents may play with respect to artefacts: those of designer, justifier and user. This *dramatis personae* is not exhaustive. In particular, agents may play other roles in which they describe artefacts in functional terms. The first role is that of *observing* artefacts. Agents may consider the use of artefacts while not being part of the communication chain of the relevant use plans. Archaeologists are typical examples of observing agents. A second role is that of *structurally analysing* artefacts. Agents may analyse the physical make-up of artefacts and, especially, their components without considering the use plans of these components. Investigators of accidents may take this role. In the final section of this chapter, we consider the functional descriptions given by these agents. We argue that functional descriptions of artefacts given by *observers* fit our theory of function ascriptions. And we indicate how structural analysts may, despite first impressions to the contrary, justifiably ascribe functions to artefacts in accordance with our theory. The discussion of this second case also shows, however, that some functional descriptions given by structural analysts are not function ascriptions as defined in the ICE-theory. This motivates the introduction of a second type of functional descriptions of artefacts, which we call *ascriptions of functional roles.*

Function ascriptions by observers

Agents may consider the use of artefacts without being part of the communication chain of use plans for these artefacts. They therefore cannot play the role of designers, justifiers or (passive) users. Examples that may immediately come to mind are archaeologists, historians and anthropologists. But there are other, perhaps less well-known examples. Engineers, for instance, at times engage in what is called 'reverse-engineering', meaning that they analyse an artefact designed by other engineers, typically those employed by a rival company. The goal of this analysis is to determine how the 'rival engineers' managed to construct the artefact in terms of its components. In reverse engineering, the engineer is part of the communication chain of the artefact itself — she knows the use plan for the cellular phone designed by the competition and the functions it is ascribed relative to this plan — but not of the use plans for the components of the artefact of which she tries to determine how they are applied in the artefact.

In all these cases the agents, whom we call *observers*, can ascribe functions to the artefact and its components. One may have the impression that such function ascriptions cannot fit our definitions, because they are not made relative to use plans; after all, by the very definition of this agent role, the agent does not know the use plan for the artefact or component.

This impression is wrong: function ascriptions by observers do fit our proposal, albeit in a slightly roundabout way. When observers ascribe functions to an artefact, they intend to ascribe the functions that were also ascribed by the original designers and users of the artefact. On our use-plan analysis, this means that observers have to recover the use plan for the artefact, since this plan was developed by the designers and executed by the agents who use(d) the artefact. Hence, to ascribe functions, observers have to overcome their not being part of the communication chain of the use plans of the artefacts they consider.

There are many ways to do this. Some artefacts carry intentional and unintentional marks of how they are to be used. Designers may have added visual cues such as handles and switches to artefacts in order to communicate (parts of) the use plans for the artefacts. Wear marks on artefacts also provide information about their use plans: observers may study these marks to learn how the artefacts were used. The finding place and the vicinity of other artefacts, whose functions are already established, also provide valuable information. Archaeologists and historians can consult texts or other sources from the period in which an observed artefact was designed or used, to try to find information about its use and thus reconstruct its use plan; anthropologists may observe how an artefact is used by its users and for what purpose. Engineers engaged in reverse engineering may acquire information about the use of components by the engineers that developed the analysed artefact, say by knowing that they follow specific design principles, or always use specific components in specific ways. The sources for this information are, say, shared design methods, information about the other company ranging from the (maintenance) manual they produce to the patents they hold.[14] Finally, observers can use structural similarities between the artefact they consider and artefacts with familiar use plans. A finding that looks like a sword or like a cup is readily identified as an artefact that has a use plan similar to those of the swords and cups we know. A copper wire with a plastic covering is readily identified by someone engaged in reverse-engineering a cellular phone as a component employed for conducting an electrical current. These attempts by observers to become part of the communication chain of the artefact may be successful or not, and may lead to detailed or coarse-grained results. But on our reconstruction, observers at least at first hypothesise a use plan for an artefact before they ascribe a function to it, and then gather evidence for this hypothesis. Their function ascriptions are then made relative to this hypothetical plan and satisfy our two definitions for function ascriptions.

It may be countered that some observers are ultimately uninterested in becoming parts of the communication chain for the use plan of the observed artefact. In reverse engineering, for instance, engineers may settle for a less ambitious goal than reconstructing how the rival engineers reasoned about their products. Gaining insight into the working of the competitor products and some ideas about how to improve on them may be sufficient for commercial reasons. We do not deny the existence or meaningfulness of this practice, but take the agents involved in it as structurally analysing artefacts, a role to which we now turn.

Functional descriptions by structural analysts

Agents may describe the components of artefacts in functional terms, on the basis of analyses of their physicochemical structure alone; in particular, they may do so without considering their use plans. There are various ways in which, in particular, engineers can play this role of *structural analysts.* As experts on the products that they and their colleagues designed, engineers may describe the components of artefacts in explicitly functional terms, e.g., as transportation pipes, engines or switches, and they may do so on the basis of an analysis of the physicochemical structure of these items. As reverse engineers, they may analyse a cellular phone made by the competition and describe a heavy battery located at the bottom of the phone as a lowerer of the phone's centre of gravity. And to give one more example, as investigators of accidents with artefacts, engineers may identify a short-circuited electrical system as a detonator of the explosion.

[14]All these sources of information are less effective than design documents or than information given by the competitor engineer himself or herself. Still, our point is that observers have ways of overcoming their epistemic isolation without being made part of the communication chains for use plans by justifiers or designers. These need not be the optimal ways of finding out the use plan.

At least some of these 'physicochemical functional descriptions' of artefacts fit our proposal: these can be taken as function ascriptions in which references to use plans, hypothetical or not, have been concealed. This concealment mechanism is detailed in section 6.2. Disregarding the value of this promissory note, we admit that not all physicochemical functional descriptions of artefacts fit our two definitions for function ascriptions. In some cases, the hypothesis that the artefact has the designed and communicated use plan that is needed for ICE-style function ascriptions, is simply too far-fetched. Such physicochemical functional descriptions by analysts are best regarded as 'plan-less'. As such, they fall outside the scope of the use-plan approach and of the theory of function ascriptions based upon this approach.

To see this, step by step, consider the three examples given above. Let us start with the expert engineer who describes an artefact, say, a pipe, in functional terms, say, as a device for transporting fluids, but who does so on the basis of its physicochemical structure and its location in an industrial installation. Although no reference is made to a use plan for the pipe, this description arguably satisfies the definition for function ascriptions by designers and justifiers. The short version of the argument goes as follows. The pipe itself has a use plan, which differs from the use plan for the installation as a whole. It may, for instance, be: 'connect the pipe to a pump in order to obtain an artefact with the capacity to displace a fluid from one reservoir to another'. Relative to this plan, the engineer may justifiably ascribe to the pipe the function to transport the fluid, in accordance with the designer/justifier definition for function ascriptions. In this definition, explicit reference is made to the use plan for the pipe; therefore, the functional description given by this engineer is reconstructed as being — despite first appearances — based on more than a structural analysis of the pipe and the installation alone. It can, however, be shown that the reference to a use plan may be omitted; the use plan of the pipe still plays its crucial role in the function ascription, but reference to it and to other non-physicochemical concepts involved are put between brackets. This explains how the description given by the engineer is phrased in terms of physicochemical concepts alone, although his function ascription is compatible with the ICE-theory.

This argument hinges on a general result for function ascriptions to components that we derive in section 6.2. This result is that the definition for function ascriptions by designers and justifiers can be simplified when the designer/justifier is an expert engineer e, who ascribes to a component c a function to ϕ relative to a component use plan 'compose c, c', c'', $\ldots$ in configuration k in order to obtain an artefact with the physicochemical capacity to Φ'. The simplified definition has the following form:

Function ascriptions to components by engineers
(expert analysts who bracket references to use plans):

An engineer e justifiably ascribes the physicochemical capacity to ϕ as a function to the component c relative to the composition of c, c', c'', ... in configuration k of an artefact x with the physicochemical capacity to Φ, and relative to scientific and technological knowledge, iff:

I. e has the belief B_{cap} that c has the capacity to ϕ;
e has the belief 'B_{con}' that x has the capacity to Φ due to, in part, c's capacity to ϕ; and

C. e can justify B_{cap} and 'B_{con}' on the basis of scientific and technological knowledge.

This result shows that in this specific case all direct references to use plans can be bracketed, leaving a definition phrased in physicochemical terms alone.[15]

Let us switch to our second example of structural analysis, namely functional descriptions in reverse engineering. Reverse engineers may describe, for example, the battery of the cellular phone as a lowerer of the centre of gravity. Such descriptions can again be taken, in line with the ICE-theory, as function ascriptions in which references to the use plan for the battery have been bracketed. The difference with the first example is that the engineer makes an (educated) guess about the use plan. She has to hypothesise that the designers of the phone developed and/or executed a use plan for the battery to realise their goal of, say, stabilising the phone in the hand of its user, and that this use plan includes locating the battery at the bottom of the phone. Relative to this hypothetical use plan, the reverse engineer may ascribe to the battery the function of lowering the centre of gravity, and then, by the result mentioned above, bracket this plan and phrase this function ascription in physicochemical terms alone.

This manoeuvre becomes difficult to carry out, or even a symptom of paranoia, when one considers the last example. There, a short-circuited electrical system is described as a detonator of an explosion. One can, of course, think up a use plan for this electrical system by which someone aims to realise his goal of creating an explosion, and then, relative to this hypothetical plan, ascribe the function of detonation to the system. But, unlike in the example of the battery, this hypothesis of deliberate plan designing is generally untenable. If all accidents would be analysed along these lines, many systems would be ascribed functional roles that are different from those for which they were designed and used. The cables of the Tacoma Narrows bridge would be resonators and thus contribute to the collapse of the bridge, although they were designed for supporting it; the ruptured tyre that caused Concorde to crash just after take-off from Paris' Charles de Gaulle airport would be a fuel-tank opener, although it was neither designed nor used for that capacity. In structural analysis, there

[15]The contribution belief mentioned in the second I-condition of this 'engineering' definition is different from the contribution belief mentioned in the designer-justifier definition: because references to use plans are bracketed in the definition of function ascriptions by engineers, the relevant contribution belief is not that a use plan p leads to its goals due to a capacity of the component c, but that an artefact x has a capacity Φ due to the capacity of the component c. We account for this difference by parenthesising this contribution belief as 'B_{con}'.

may not be use plans that correspond to the functional descriptions of the artefacts. Such physicochemical functional descriptions fall outside the scope of our use-plan approach and thus of our definitions for function ascriptions.

To account for these deviant descriptions, we take them as a separate type of functional descriptions within the ICE-function theory and call them *functional roles*:

Ascriptions of functional roles by analysts:

An analyst a justifiably describes a component c as functioning physicochemically as a ψ-er relative to an item x with the physicochemical capacity to Ψ and composed of c, c', c'', ... in configuration k, and relative to scientific and technological knowledge, iff:

I. a has the belief B_{cap} that c has the capacity to ψ;
a has the belief 'B_{con}' that x has the capacity to Ψ due to, in part, c's capacity to ψ; and

C. a can justify B_{cap} and 'B_{con}' on the basis of scientific and technological knowledge.

Since these functional roles are not ascribed relative to use plans of artefacts, they differ conceptually from function ascriptions. The conditions that should be met by ascriptions of functional roles are, however, similar to those that should be met by function ascriptions to components. In that sense, ascriptions of functional roles by analysts can be seen as generalisations of engineering function ascriptions to components by engineers, obtained when one does not bracket the use plans, but denies their existence.

This definition of functional roles bears a marked resemblance to Cummins' (1975, p. 762) definition of causal-role functions (see section 3.3). The main difference is that our definition stresses that functional roles are ascribed by agents, on the basis of justified beliefs instead of a correct account.

If this definition of functional roles would just be added to the two previous definitions, the ICE-theory would become a hybrid. To avoid this, we do not advance the definitions of function ascriptions and of functional roles on an equal footing. We take function ascriptions as central to technology, so that the corresponding definitions should accommodate the bulk of functional descriptions by designers, justifiers and users, and should meet the requirements and desiderata that apply to a philosophy of artefacts. Ascriptions of functional roles are taken as a weaker and broader type of functional descriptions in technology, included to cover a specific use of functional descriptions, but which need not meet all requirements or desiderata.[16]

For ease of reference we list the three ICE-definitions of functional descriptions in table 4.2.

[16]Ascriptions of functional roles meet the support and innovation desiderata. The malfunctioning desideratum is again met in the limited sense that an agent has to believe that a component has the capacity corresponding to the ascribed functional role; hence, once an agent believes that a component does not have a specific capacity, this capacity can no longer be ascribed as a functional role. Finally, ascriptions of functional roles can all be taken as accidental and be said to meet the proper-accidental desideratum when complemented by proper function ascriptions.

Function ascriptions by designers or justifiers:

A designer d or justifier j justifiably ascribes the physicochemical capacity to ϕ as a function to an artefact x relative to a use plan p for x, and relative to account A, iff:

I. d/j has the belief B_{cap} that x has the capacity to ϕ;
d/j has the belief B_{con} that p leads to its goals due to, in part, x's capacity to ϕ; and

C. d/j can justify B_{cap} and B_{con} on the basis of A.

Function ascriptions by passive users:

A passive user u justifiably ascribes the physicochemical capacity to ϕ as a function to an artefact x relative to a use plan p for x, and relative to testimony T, iff:

I. u has the belief B_{cap} that x has the capacity to ϕ;
u has the belief B_{con} that p leads to its goals due to, in part, x's capacity to ϕ;
u believes that a designer d or justifier j of p has B_{cap} and B_{con};

C. u can justify B_{cap} and B_{con} on the basis of T;
u can justify on the basis of T that d/j has B_{cap} and B_{con}; and

E. u received T that d/j has B_{cap} and B_{con}.

Ascriptions of functional roles by analysts:

An analyst a justifiably describes a component c as functioning physicochemically as a ψ-er relative to an item x with the physicochemical capacity to Ψ and composed of c, c', c'', ... in configuration k, and relative to scientific and technological knowledge, iff:

I. a has the belief B_{cap} that c has the capacity to ψ;
a has the belief 'B_{con}' that x has the capacity to Ψ due to, in part, c's capacity to ψ; and

C. a can justify B_{cap} and 'B_{con}' on the basis of scientific and technological knowledge.

Table 4.2: The three ICE-definitions for functional descriptions

Chapter 5

Malfunctioning

In section 4.3, we showed that the ICE-theory satisfies the desiderata for a theory of artefacts. Yet we acknowledged that it meets the malfunctioning desideratum in a formal, limited sense. In this chapter, we go beyond what is required by our own conceptual-engineering method. We examine to what extent the ICE-theory and the use-plan approach account for the phenomena *from which we derived the desiderata. This investigation centres on malfunctioning, but the results extend to other phenomena, in particular that of use versatility.*

Going from desiderata to phenomena serves a dual purpose. First, it increases the phenomenological appeal of our theory. We show in the course of this chapter that the ICE-theory can account for many different aspects of malfunctioning. We do not give a full description of malfunctioning. Yet our investigations make clear how intricate such a description must be, and which ingredients — social, epistemic, practical and action-theoretical — constitute it. Thus, second, we show that much of the phenomenological work is, in fact, done by the theory of using and designing and not by the theory of functions.

We start the chapter by presenting some central varieties of artefact malfunctioning and thus determining the phenomenological work that faces us. Then, in section 5.2, we argue first that the ICE-theory accounts more broadly for malfunctioning if one differentiates between the belief that an artefact exercises *a capacity and the belief that the artefact* has *that capacity. Second, we broaden our analysis by reflecting in section 5.3 on the normative content of malfunctioning claims, and that of claims about artefacts in general. There, we reveal two types of normative content, one related to practical reasons, and found in all function ascriptions, and one related to the privileging of use plans and the role of professional designers, which is found in proper-use claims and some malfunctioning claims. The second type in particular leads us beyond a theory of artefact functions to its social and action-theoretical background.*

W. Houkes, P.E. Vermaas, *Technical Functions*, Philosophy of Engineering and Technology 1, DOI 10.1007/978-90-481-3900-2_5,

5.1 The phenomenon of artefact malfunctioning

In the previous chapter, we presented our ICE-function theory, and we assessed whether it meets the four desiderata we set for a theory of artefacts. The results were positive for the proper-accidental, support and innovation desiderata, but mixed for malfunctioning. Formally, the ICE-theory meets the malfunctioning desideratum: we sketched cases in which various agents — designers, justifiers, passive users — may justifiably ascribe functions to non-working artefacts. What these cases have in common is the proviso that the agent involved should believe, while ascribing the function to ϕ, that the artefact concerned has the capacity to ϕ. If the agent does not have that belief, the I-conditions of the two central ICE-definitions remain unfulfilled, and function ascriptions to the artefact are not supported by the ICE-theory. This proviso holds for some cases of malfunctioning, such as the prototype of a newly engineered type of aeroplane that fails to lift-off at its first test; or a television set that worked properly yesterday, but fails to show broadcasts when it is turned on today. The proviso seems not to hold for other cases that, arguably, involve malfunctioning. Suppose an aeroplane fails to fly after one year of service and is towed back to its manufacturer for inspection. Intuitively, the plane is malfunctioning since its users and designers will ascribe to it the function of transporting people through the air. In the ICE-theory, this function ascription seems unjustified, since the users and designers know that the plane will not fly. Similarly, a broken-down television set is usually still ascribed the function of showing broadcasts, also by agents who know that it cannot perform this function anymore. Yet, in the ICE-theory, this knowledge prohibits agents to continue to ascribe this function, meaning that the owner of the set cannot, strictly speaking, call the shop to complain that the set is malfunctioning.

Thus, technically, the ICE-theory meets all desiderata, but it is intuitively lacking: it does not account for some cases in which one might think that malfunctioning claims are in order. Intuitively, the malfunctioning desideratum is only met in a limited sense. If one takes these limitations seriously, and adopts our assessment that existing theories of technical functions do not satisfy all four desiderata (chapter 3), one might conclude that our desiderata are mutually incompatible, as suggested by Beth Preston (2003).

In this chapter, we argue that the ICE-theory can describe malfunctioning more broadly than argued for in section 4.3, thus taking away worries about limitations or shortcomings of the ICE-theory and about the incompatibility of our four desiderata. In this section, we first present some central varieties of artefact malfunctioning, in order to determine more precisely the phenomenological work facing us. Then, we evaluate the extent to which the ICE-theory can deal with these varieties, singling out the remaining problematic cases. We turn to these in the remainder of the chapter, developing two arguments to improve the phenomenological merits of our approach. The first argument is developed in section 5.2, and directly widens the number of malfunctioning varieties accounted for by the ICE-theory. The argument is based on a distinction between the belief that an artefact *exercises* a capacity and the belief that the artefact *has* that capacity, and ultimately appeals to notions of repair and maintenance of artefacts. The second argument, developed in section 5.3 focusses on the normativity of malfunctioning claims, a feature that one may still find lacking after the work of section 5.2. This argument reveals a complicated network of social, social-epistemic and practical recommendations and requirements, against the background of which claims about malfunctioning artefacts are made.

These arguments go beyond what is required by our own conceptual-engineering method of showing that the ICE-theory meets the four desiderata. Instead, they examine to what extent this theory with the underlying use-plan approach is able to account for the *phenomena* from which we derived the desiderata — not only the phenomenon of malfunctioning, but also that of use versatility, as it was described in chapter 1. Furthermore, the arguments show that much of the phenomenological work is, in fact, not done by the ICE-function theory, but by the use-plan approach that underlies it.

Some varieties of artefact malfunctioning

Before determining to what extent our ICE-function theory successfully accounts for malfunctioning, we need to establish what should be accounted for. Some central varieties of artefact malfunctioning can be distinguished along at least three different lines.[1]

First, degrees of malfunctioning may be distinguished, ranging from slightly sub-optimal performance, through minor and major defects, to outright failure to perform the function. To give a range of simple examples: a light bulb may not emit light of the expected intensity, or it may consume more energy than advertised; it may flicker occasionally or constantly; and the bulb may fail to emit light altogether. Various notions are used in everyday language to describe these cases. We gather them all under the heading of 'malfunctioning'. Insofar as they concern the functioning of an artefact (instead of, say, its safety), we hold our theory of artefacts accountable for all cases along this scale.

Second, malfunctioning statements may concern either artefact tokens or entire types. Token malfunctioning is prototypical, but some malfunctioning claims may concern types. One may, for instance, describe all light bulbs of a type as malfunctioning, because they all consume more energy than they should, or because none of them burns for the advertised number of hours. In the following analysis, we concentrate on token malfunctioning, and most examples — both in this chapter and in everyday life — are of this variety. Yet a theory of artefacts that remains silent on type malfunctioning would be phenomenologically incomplete.

Third, a distinction may be made concerning the situation in which a malfunctioning claim is made. These situations may be distinguished in terms of the role played by the agent making the claim, and in terms of the temporal relation between the malfunctioning event and the malfunctioning claim. We

[1] The first two ways of distinguishing malfunctioning claims are also presented by Franssen (2006).

call the first type of situation 'evaluation-in-action'. In these situations, malfunctioning claims are made by someone practically involved with the artefact, i.e., either a designer or a user (passive user or justifier) of the artefact. We already encountered one subtype of evaluation-in-action in section 4.3, namely evaluation-in-design. A prototype may, for instance, be found to perform poorly during a test, even if the designers had reason to assume that it would work as planned. The other subtype, evaluation-in-use, may be more common. It features in one of the most prominent descriptions of artefact malfunctioning in philosophy, namely Heidegger's analysis of damaged equipment.[2] When using a hammer, a carpenter may find that its head is loose; when turning the ignition of a car, a driver may find that the car fails to start. The second type of situation is called 'external evaluation' and concerns malfunctioning claims made by agents who are not practically involved with the artefact, but are observing such involvements (i.e., the agents play the observer role introduced in section 4.4). Anthropologists who observe traditional methods for curing diseases may, for instance, have sufficient reason to believe in the ineffectiveness of these methods for achieving their alleged purposes, and may formulate their doubts as malfunctioning claims. The third and final type of situation may be called 'post-hoc evaluation'. In such situations, a malfunctioning claim is made after the actual event. A driver may, for instance, after failing to start her car for some minutes, opening the bonnet, etc., decide to call her car dealer, and describe her car as malfunctioning.

Successes

Our previous assessment of the ICE-theory can be mapped onto the spectrum of varieties listed above, taking the third, situational distinction as a guideline.

As said at the start of the chapter, the ICE-theory meets the malfunctioning desideratum in those cases in which agents may ascribe to non-working artefacts the relevant capacities to ϕ as functions. This means that the agent involved should, when ascribing the function to ϕ, believe that the artefact concerned has the capacity to ϕ. Therefore, the ICE-theory accounts straightforwardly the evaluation-in-action varieties of malfunctioning. This success extends to both artefact types and artefact tokens. To start with the latter, the agent may support his capacity belief about the artefact token with experience of previous successful use of tokens of the type (including the presently used token). Someone may justifiably ascribe to a toaster of a specific brand the capacity to produce toast as a function, based on experience with toasters of that brand. This kind of evidence is not available for type malfunctioning, except in the rare case when all artefacts of a given type suddenly break down. However, agents may have justifiable, but false capacity beliefs about an artefact type, based on scientific theories or on testimony. As an example, consider the electric bug zappers described by Preston (1998b, pp. 245–246). All artefacts of this type are ascribed the function to eliminate mosquitoes, and this function ascription is justifiable, at least on the basis of the testimony given by the designers and

[2] Heidegger (1962, §16).

manufacturers of these artefacts. It turns out that no electric bug zappers actually have the capacity to eliminate mosquitoes — as studies into the effects of using the artefacts have shown — and there is some evidence that bug zappers actually attract mosquitoes, without harming them. But as long as these facts remain unknown to the agents ascribing the functions, their function ascriptions to malfunctioning bug zappers (as a type) are justifiable.

This example also shows how our ICE-theory deals with external evaluations. Suppose that the first person who has done research into the effects and capacities of bug zappers has recently reached the conclusion that the artefacts are ineffective, but has not yet communicated this conclusion to others. If this agent comes across someone using a bug zapper, he may externally evaluate the performance of this artefact. The knowledgeable agent then acts as an observer, albeit of an uncommon variety. The observers who were described in section 4.4 are not part of the communication chain of a use plan because no information about use plans is passed to them by designers and/or justifiers. External evaluators place themselves outside of a communication chain: they no longer accept the effectiveness of a use plan and the capacity belief for the artefact. Still, an external evaluator may observe that someone else ascribes a function to an artefact on the basis of false, but — for the ascribing agents — justified capacity and contribution beliefs. As long as the bug-zapper researcher has not informed others about the ineffectiveness of the artefacts, it makes little sense to claim that these others may not ascribe the function of eliminating mosquitoes to the zappers. In the terminology introduced by Bernard Williams (1981), the ascribing agents have an external, but not an internal reason to reject their capacity beliefs and function ascriptions. If one does not accept this distinction between internal and external reasons, external evaluations cannot be expressed as malfunctioning statements: they are better expressed as claims that no function ought to be ascribed, neither by the observer nor by the agent observed. If the distinction between internal and external reasons is acceptable, the ICE-function theory accounts for external evaluations.

This brings us to post-hoc evaluations. In these cases, the agent who makes the malfunctioning claim seems to be aware that his previous capacity belief is false. Hence, we cannot appeal to the distinction between internal and external reasons to save the function ascription. Rather, it seems that the agent making the claim ought not to ascribe the function. This may be acceptable for type malfunctioning: as soon as the ineffectiveness of bug zappers becomes common knowledge, no one has a reason to ascribe the function of eliminating mosquitoes to these artefacts. For some cases of token malfunctioning, the same conclusion applies. A car that was designed for twenty years of service, but breaks down after fifty years, ought perhaps not to be ascribed the function of transporting people anymore. For other cases of token malfunctioning, however, this conclusion is counterintuitive. If the car would break down after just two years, it seems outrageous to claim that its user ought not to ascribe the function of transporting people to the car, so that it is strictly speaking not malfunctioning.

Summing up, the ICE-function theory can account for a large variety of malfunctioning claims, perhaps larger than suggested by the evaluation in section 4.3. Still, it leaves something to be desired, since there are uncovered varieties of artefact malfunctioning, in particular post-hoc evaluations. Our task in the remainder of this chapter is to show whether, and how, the ICE-theory can deal with these remaining cases.

5.2 Having capacities versus exercising them

A first line of argument to show how the ICE-theory may deal with post-hoc evaluations of malfunctioning artefacts is to tease apart the belief that an artefact *has* the capacity to ϕ (the belief B_{cap} that is necessary for function ascriptions) and the belief that the artefact *exercises* this capacity to ϕ. Post-hoc evaluations of malfunctioning seem problematic because, if someone knows that an artefact is malfunctioning, say, because she observed that it does not exercise this capacity to ϕ, it seems that she must reject B_{cap} and therefore the function ascription. On closer inspection, however, this situation leaves room for saving the ICE-theory. In the line of reasoning just given, it is assumed that the belief that an artefact does not *exercise* a capacity to ϕ entails the belief that the artefact does not *have* the capacity to ϕ. If this inference can be challenged, post-hoc malfunctioning may be accounted for by the ICE-theory: an agent ascribes a function to ϕ to an artefact since she believes that the artefact has the capacity to ϕ even though she believes that the artefact cannot exercise this capacity.[3]

In the domain of technical artefacts, the inference can be blocked. There are circumstances in which one may believe that an artefact does not exercise the capacity to ϕ, while (still) believing that the artefact has the capacity to ϕ. Take, for instance, a car that does not drive when properly operated by someone with the requisite skills. In this situation, the car may not have lost the capacity to be driven; whether this is the case depends on additional circumstances. If, for instance, the car was set on fire and is heavily damaged, it seems reasonable to believe that it indeed does not have the capacity to be driven. But if the car has merely run out of petrol or the starting motor is broken, it arguably still has the capacity to be driven: this may be demonstrated by driving the car after its tank has been refilled or after the starter has been replaced. Hence, there are circumstances in which one may believe that a car does not exercise a capacity while believing that it still has that capacity — the two beliefs are not mutually inconsistent.

The circumstances in which artefacts have capacities but do not exercise them can be analysed in general terms. For this, we briefly return to the conditional analysis of capacities that we adopted in section 4.2. According to that analysis, the antecedents in the conditionals concerned are in technology often captured incompletely. One focusses pragmatically on the antecedents that are relevant for use and includes the remaining 'default' antecedents in an unspecified set of normal conditions. More formally, a capacity to ϕ corresponding to

[3] This argument develops the line of reasoning given in (Vermaas 2009b, §6).

an artefact's function is often analysed as the conditional $N \wedge A_d \rightarrow (A \rightarrow C)$, where C are the consequences, A the (non-default) antecedents, A_d the default antecedents and N other normal conditions. For the car, C corresponds to transporting people, A to moving the car's controls properly, and A_d includes the tank being filled with petrol and having a starting motor in good technical condition. An artefact can be taken as exercising the capacity to ϕ iff $N \wedge A_d \rightarrow (A \rightarrow C)$ and A_d hold: in that case the conditional $A \rightarrow C$ is true for the artefact. For the car: if the tank is filled and the starting motor works (and all other normal conditions are true as well), then the car will transport people when its controls are moved properly. It is tempting to take an artefact as having the capacity to ϕ iff $N \wedge A_d \rightarrow (A \rightarrow C)$ holds, but this would be too liberal: it opens the door to arguing that all artefacts have this capacity. The car with the empty tank is an artefact for which the conditional holds, as demonstrated when the tank is refilled such that A_d holds. But the conditional also holds for the heavily damaged car, as would be demonstrated if it would be completely restored in its original state (e.g., by waving a magic wand and saying 'Reparo'[4]), such that A_d holds. Even worse, the conditional may hold for any object whatsoever, as would be demonstrated by bringing it into a state in which A_d holds (e.g., by transforming it into a car[5]). So, in order to come up with a reasonable definition for an artefact having a capacity to ϕ, we need a criterion to restrict the application of the conditional. Technology provides such a criterion, which brings us back to the action-theoretical background of artefacts.

In the characterisation of designing given in section 2.4 we said that designing may produce information about how to maintain and repair artefacts. This information may be communicated to prospective users — think of information in manuals, including lists of how to resolve problems with artefacts — and to technical professionals who assist users in maintaining and repairing artefacts. By this information, transformations of the states of the artefacts are defined that are deemed technologically possible and economically reasonable. Cars with empty tanks can be filled up with petrol by the user, and cars with broken starting motors can be repaired by professional maintenance people. By contrast, cars that are heavily damaged by fire are deemed 'total loss', meaning that there are technological or economical reasons against believing that they may be restored to their original state. Using this action-theoretical background, we can propose the following definition: an artefact has the capacity to ϕ iff $N \wedge A_d \rightarrow (A \rightarrow C)$ and A_d hold *or* $N \wedge A_d \rightarrow (A \rightarrow C)$ holds and the state of the artefact can be transformed by technologically and economically acceptable maintenance or repair into a state in which A_d holds.

By teasing apart the beliefs that an artefact has a capacity and that it can exercise it, we can distinguish three cases:

[4] http://en.wikipedia.org/wiki/Spells_in_Harry_Potter

[5] This transformation would involve magic that even Harry Potter is (as yet) incapable of using.

1. an agent justifiably believes that an artefact has the capacity to ϕ and justifiably believes that it exercises this capacity;

2. an agent justifiably believes that an artefact has the capacity to ϕ and justifiably believes that it does not exercise this capacity;

3. an agent justifiably believes that an artefact does not have the capacity to ϕ.

(These three cases exhaust the possibilities: it follows from the definitions that, if an agent justifiably believes that an artefact does not have the capacity to ϕ, the agent justifiably believes that the artefact does not exercise the capacity to ϕ.)

The second case corresponds to some post-hoc evaluations of malfunctioning artefacts. This shows that the ICE-theory extends successfully to such situations: there, an agent may ascribe the capacity to ϕ as a function to the artefact, because she justifiably believes that an artefact has the capacity to ϕ (i.e., the agent has the requisite capacity belief), although she justifiably believes that it does not exercise this capacity (e.g., because she witnessed an earlier event of malfunctioning).

Three comments are in order.

First, it may seem that introducing the having-exercising distinction overshoots the mark, because it allows too many cases of malfunctioning. Not every case in which an artefact has, but does not exercise a capacity is intuitively a case where malfunctioning claims are appropriate. Again, consider the car. If the starting motor has broken down, the car can be taken as malfunctioning: it can be restored in driving condition, but the average user cannot perform this repair. By contrast, a car that has run out of petrol is not typically described as malfunctioning. Restoring this car in driving condition involves refilling its tank, which may be described as one of the most elementary aspects of car maintenance, left to users. More strongly, a car parked in a garage does not exercise its capacity to transport people, but may be fully functional; 'restoring' this car in driving condition only involves starting it. This action is part of using a car, and not intuitively described as maintenance of even the most elementary type. This shows that the analysis above can be developed by distinguishing between use, maintenance and repair; and between maintenance and repair operations that are communicated to users, and maintenance and repair operations that are not. Post-hoc evaluation can then be characterised through three justified beliefs of the function-ascribing agent: that an artefact has the capacity to ϕ corresponding to its function; that it does not exercise this capacity; and that she cannot, through actions included in a use plan or communicated maintenance and repair operations, transform the state of the artefact into a state in which it again exercises the capacity to ϕ.

Second, malfunctioning becomes a context-dependent phenomenon, since it is related to judgements about the technological and economic feasibility of transformations of the states of artefacts. The technological possibilities for maintenance and repair may, however, change over time, with the state of technology in general, and they may depend on local resources. On our analysis, malfunctioning claims would be equally dynamic. A burned-out car may become easily repairable in the future, and thus judged to be malfunctioning rather than total-loss. A car with a broken starting motor may not be repairable in a region in which no spare parts are available; there it should be described as having lost its function rather than as malfunctioning. Moreover, a car damaged by fire may be considered repairable from an economic point of view if it is considered to be of high financial, cultural or emotional value. And if the prices of spare parts or of labour increase enormously, changing a starting motor may become economically unfeasible.[6] We think that this context-dependence is acceptable: it is not part of the phenomenology that malfunctioning claims are made *sub specie aeternitatis*.

Third, our analysis has focussed on artefacts for which information is available on maintenance and repair. Such information is not available for all artefacts — think of pencils, tennis balls and tea bags. It would be odd to rule out post-hoc evaluations of such artefacts. Our analysis might accommodate them because repair of simple artefacts is typically common knowledge (think of re-sharpening a broken pencil), and/or economically unattractive (think of refilling and resealing a ruptured tea bag). In either case, there is no need to communicate information on repair, but the having-exercising distinction may be applied.

Our analysis in this section has shown that the ICE-theory can handle post-hoc evaluations of malfunctioning artefacts. Therefore, it meets an even wider variety of malfunctioning claims than shown in section 4.3.

Malfunctioning in other function theories: Cummins again

A final note concerns other function theories than ours. For despite the connection with the use-plan approach, our first line of reasoning also provides a means for other function theories to account for malfunctioning. As an example, we briefly consider Cummins' (1975) causal-role theory. As noted in section 3.3, Cummins' theory is considered notorious for failing to account for malfunctioning claims. On Cummins' definition, an item 'x functions as a ϕ in s (or: the function of x in s is to ϕ)' only if that item 'x is capable of ϕ-ing in s' (see page 57). And observation that x is not actually ϕ-ing in s, is generally taken as evidence that the last condition does not hold. With the distinction between x having a capacity and x exercising it, one may now challenge this last conclusion. Observation that the item x is not actually ϕ-ing in s can be taken as evidence that x is not exercising the capacity to ϕ in s in the circumstances at hand. This, however, does not rule out that in other circumstances 'x is capable of ϕ-ing in s', when this last phrase is interpreted as that x has the capacity to ϕ in s. On this interpretation, Cummins' theory does account for malfunctioning. And if one wants to reject this interpretation as a proper

[6] In the Netherlands there are, for instance, little attachable electrical lights available for making cyclists visible at night. These lights have batteries that can be replaced when the lights fail. Yet, when the lights fail, it is cheaper to replace them with new lights than to buy new batteries, leading to the conclusion that a light that fails is immediately beyond maintenance or repair.

interpretation of Cummins' theory, the underlying reasoning still puts us in the position to formulate a function theory that can account for malfunctioning and that is similar to Cummins'. We give it here, without analysing the pros and cons of such a modified causal-role theory:

> Modified causal-role theory of functions:
>
> x functions as a ϕ in s (or: the function of x in s is to ϕ) relative to an analytical account A of s's capacity to Φ just in case x has the capacity to ϕ in s and A appropriately and adequately accounts for s's capacity to Φ by, in part, appealing to the capacity of x to ϕ in s.

5.3 Artefact normativity

With the having-exercising distinction, the ICE-theory accounts for all types of situations in which malfunctioning claims may be made: evaluations-in-action, external evaluations and post-hoc evaluations. Yet it still does not account for each and every malfunctioning claim. There are situations in which artefacts are so damaged that repairing them is technologically or economically unfeasible. In that case, the 'restorability' condition, which allows people to retain their capacity belief and to ascribe the original function to the artefact, is not fulfilled. Thus, a burned-out television set may no longer be ascribed the function to show broadcasts, and a crashed aeroplane no longer has the function to transport people through the air; and because they may not be ascribed these functions, post-hoc evaluation claims regarding these artefacts are ruled out. We accept this small loss in saved intuitions. There is, after all, a point at which a television set stops having its original function, and becomes a former television set. Being beyond repair may be taken as this point.

This is not the only remaining issue for the ICE-theory. On our analysis in the previous section, the capacity belief must also be rejected in case repairing the artefact is economically unfeasible. Therefore, 'total-loss' cars are in the same situation as crashed aeroplanes: their users should stop ascribing to them the function to transport people. This seems somewhat harsh, if only because we have little trouble describing a total-loss car as a car. On a similar note, electronic bug zappers are still called bug zappers, although it is at least impractical to restore their capacity to eliminate mosquitoes — a capacity that they did not have in the first place.

We are prepared to bite this bullet. Any analysis is bound to violate some intuitions on borderline cases. Post-hoc evaluations about artefacts that cannot be restored in working condition, such as total-loss cars and bug zappers, are our borderline cases. Still, we may examine why a malfunctioning claim on these borderline cases seems appropriate, and whether this claim can perhaps be reformulated in such a way that it still is applicable.

This investigation is also in order for another reason. One might think that our entire account of malfunctioning with the ICE-theory is on the wrong track, because it does not show, or shows insufficiently, why malfunctioning is a normative phenomenon. Intuitively, when someone claims that her television set is malfunctioning, she effectively claims something like 'My television set does not show broadcasts. Still, it ought to. Therefore, it is a bad television set.' On our analysis, these statements express the belief that the television set has the capacity of showing broadcasting, but cannot presently exercise this capacity — which seems purely descriptive. The 'ought' statement about the television set can then at most express the belief that the set could previously exercise the capacity — which is, again, descriptive — or, perhaps, express an appeal to repair and maintenance. This invokes standards of technological and economic feasibility, but these might only reflect the current state of technology or availability of resources, apparently making malfunctioning claims as non-normative as the claim that we cannot yet produce superconductors that function at room temperature.

It is not obvious that this lack-of-normativity objection is grounded. For one thing, not every 'ought' statement is normative. Some, such as 'Dutch autumns ought be rainy', merely express rational expectation. Others, such as 'There ought to be snow on Christmas', express little more than hope: in the Netherlands, there is only a small chance of snow on Christmas. Other 'ought' statements, however, clearly express normativity. 'You ought to help others if you can' involves a recommendation, and perhaps a mild rebuke. 'You ought to take care of your children' involves an obligation.[7] Even supposing that malfunctioning claims can be unpacked as, in part, 'ought' statements, it is unclear whether they involve 'ought' statements of a normative variety. The same goes for the evaluative second part of the statement 'My television set does not show broadcasts. Still, it ought to. Therefore, it is a bad television set.' Some authors do not accept evaluative statements regarding artefacts as genuinely normative, since they do not concern intentional actions.[8] Nevertheless, artefact malfunctioning is often taken as a normative phenomenon, just as proper-function claims are occasionally read as involving normativity.[9] However, these interpretations are often disputed.[10] Thus, there is no general

[7]Vaesen (2006) presents a more detailed account of the role of various types of 'ought' statements in technology.

[8]Many authors, following Castañeda (1970), make a distinction between deontic 'ought to do' statements, which are only applicable to intentional actions, and evaluative 'ought to be' statements, which might also be applied to artefacts. Less explicitly, many endorse the view that only deontic statements are genuinely normative. An example of an implicit endorsement of this view is: 'We might say that the circuit is wired up if it is in a state such that, when the key is turned, there ought to be a spark. But the 'ought' here speaks not of duties and values, but just about what you would expect. If there is no spark, then something has gone wrong, but this too means simply loss of expected or intended function. The ignition creates the current; that is what it is for.' (Blackburn 1998, p. 56)

[9]E.g., 'The main *normative* distinction is between proper functioning, on the one hand, and malfunctioning (or dysfunctioning), on the other.' (Neander 1995, p. 111); 'The main difference between system function and proper function is therefore that the latter is normative whereas the former is not. Moreover, the normativity that defines proper function is not reducible to statistical regularity.' (Preston 1998b, p. 224)

[10]Millikan, for instance, admits that malfunctioning claims and proper-function claims may involve normativity, but of a very weak variety: 'Normative terms are not always evaluative,

agreement on the relation between artefact malfunctioning and normativity, and the relation is seldom analysed in the literature.

Here, we might revert to our conceptual-engineering approach. Our malfunctioning desideratum does not mention normativity, and is therefore satisfied by the ICE-theory, perhaps in combination with its background. Since the relations between malfunctioning and normativity and between the malfunctioning phenomenon and theories of technical functions are unclear, there is no reason to impose requirements that are, in this respect, stricter than our desideratum. The doubts of some authors regarding the normativity of malfunctioning claims or even all artefact-related 'ought' statements provide an additional argument against going beyond our initial requirements. This would mean that the assessment of the theory is complete, and that the ICE-theory is completely successful in accounting for most types of malfunctioning claims.

This way of rejecting some of the phenomenological burdens resting on a theory of artefacts may strike some as unacceptable, like our previous acceptance of the remaining limitations of the ICE-theory. To satisfy those readers, we now show how our theory of artefacts, i.e., the combination of the use-plan approach and the ICE-function theory, accounts for the normativity of statements about artefacts. This demonstration is, strictly speaking, super-erogate; those who are convinced of the merits of our theory on the basis of the previous assessments may skip the remainder of this section.

To find out which statements about artefacts might be normative, and in which ways, we first need a general characterisation of normativity. We choose one of the few explicit characterisations available, given by Jonathan Dancy (2006). On Dancy's characterisation, a normative fact, i.e., the content of a normative statement, is a second-order fact, namely that a particular fact or set of facts about the world is relevant to the beliefs, desires or actions of a specific person or of any person.[11]

On the basis of this characterisation and our action-theoretical analysis of using and designing, various statements about artefacts, in particular function ascriptions, and statements about their proper use and malfunctioning, may be shown to be normative.[12]

Function ascriptions

Firstly, on Dancy's characterisation, the effectiveness belief B_{eff} that executing a use plan leads to its goal is normative. Suppose one believes that executing the camping-ground use-plan for making tea (section 2.2) is effective. Then,

but can indicate any kind of measure from which actual departures are possible.' (1999, p. 192)

[11] Dancy (2006, p. 136). Dancy does not characterise normative statements in non-normative terms; his criterion is not reductive.

[12] Dancy's characterisation of normativity has earlier been used to clarify artefact normativity by Franssen (2006). The remainder of this section owes some debts to this paper. In addition, the subsection on proper-use claims and malfunctioning claims is partly based on Houkes (2006).

one has a reason to execute this plan, in case one desires a cup of tea; therefore, on Dancy's characterisation, B_{eff} has normative content. As argued in section 4.2, a function ascription analyses B_{eff} in terms of the belief B_{cap} that the artefact has a physicochemical capacity to ϕ and the belief B_{con} that this capacity contributes to the effectiveness of the use plan. Thus, every function is ascribed relative to an effectiveness belief. This means that function ascriptions are normative as well. Take, for instance, ascribing to a chair the function to stand on. This refers to a second-order fact, namely that a certain capacity of the chair is of practical relevance; in particular, when executing a use plan, the chair contributes to the realisation of the goal of the plan. In other words, function ascriptions give a person, or any person, reasons for actions, namely to manipulate the chair in a certain way. If I want to realise a goal g and I know that an artefact x can be used to realise the goal g, it is practically rational for me to try and realise g by means of x. Thus, function ascriptions carry an element of prescription or recommendation regarding actions.

That function ascriptions are normative makes clear how malfunctioning statements might be normative. In situations in which the user becomes aware of malfunctioning while executing the use plan, she had a practical reason to use the artefact, but it is defeated upon execution of the plan. In this sense, post-hoc evaluations — and external evaluations as well — express the fact that someone does *not* have a reason to use the artefact to realise a goal g — or, perhaps, that someone has a reason not to use the artefact to realise a goal g.[13]

It may be possible to take the latter as a normative fact in Dancy's sense; and there is little doubt that it is evaluative. Yet this does not seem to be all there is to the normativity of malfunctioning statements. Suppose that a car fails to start after two years of service. If its driver then makes a post-hoc evaluative statement like 'My car does not work', she does not only claim that she has a reason not to use the car for transportation (i.e., that one has reason not to use it for this purpose), but also that she *ought to* have had a reason to use it. This is not true for all post-hoc evaluations. If a swivelling chair turns out not to be usable to stand on in order to change a light bulb, the statement 'This chair is unfit to stand on' only expresses the absence of a practical reason — if there is an additional 'ought' statement involved, it amounts to no more than an expression of a previous expectation.

The relevant difference between the malfunctioning car and the malfunctioning swivelling chair seems to be that the car is properly used for driving, and that the swivelling chair is used improperly, or accidentally, for standing on. Therefore, we take a closer look at the normativity of proper-use claims, before examining whether malfunctioning statements have more normative content than expressing a practical reason.

Proper use

Proper use, as discussed in section 2.7, involves the social acceptance and privileging of a particular use plan over others. Testimony is not just a basis for

[13]Franssen (2006, p. 47).

the effectiveness, capacity and contribution beliefs that are involved in function ascriptions, but also provides a basis for this privileging. In teaching each other certain allegedly effective use plans, we simultaneously privilege these plans over other plans, either by remaining silent about other plans, or by explicitly discouraging others from executing or even deliberating about alternative plans.[14] Accepting testimony reflects the social entrenchment or possible institutionalisation of use plans, and the process of forming a 'community of users', such as a team or organisation.

Not all such social entrenchment concerns proper use. Opening hotel doors by means of a credit card, or escaping from a prison cell by means of bedsheets are socially entrenched plans, and one's capacity and contribution beliefs regarding the credit cards and bedsheets are largely or exclusively based on testimony, not on personal experience. Yet few people would claim that bedsheets are used properly to escape from a prison cell, although one might justifiably ascribe this escape function to them. Hence, we do accept testimony for these claims, but that does not show that they are claims about proper use; moreover, we accept the authority of the spokesperson or the one showing this way of using the artefact, but only as a testifier for a usefulness claim.

The relation between testimony and proper use can be restored by the typology of designing introduced in section 2.4 and invoked in section 2.7 when we analysed the notion of 'proper use'. 'Designing' is not an impartial label attached to certain intentional actions. The role of designing is professionally played by some agents and not by others, who are typically only involved in using artefacts. This 'right to design' comes with privileges, most notably that of determining the proper use of an artefact. Conformation is enforced through warranties, peer pressure and other mechanisms. The rationale behind this social mechanism can be understood in terms of practical rationality: if some members of a group specialise in designing use plans and artefacts, and communicate the results of their efforts to other members, the latter are saved the effort to construct use plans and artefacts of their own; and if the professional designers do a good job, everyone profits. This means that the division of labour between designers and users, and the standards for good design (section 2.7), create a social role for testimony. Professional designers can only fulfil their social role if they provide testimony for the capacity and contribution beliefs in function ascriptions, and for the effectiveness of use plans that form the background of these function ascriptions. If they do not, their efforts may be evaluated as bad design. For their part, consumers play their expected social role if they accept this testimony, i.e., if they trust the designers, and act as passive users on our function theory. Proper-use claims reflect this trust or, negatively, the social sanctions on amateur designing.[15]

[14]This peer pressure extends to other characteristics than effectiveness. While taking driving lessons, you are taught that exceeding speed limits is unacceptable, but there is little doubt that it is effective — albeit not always efficient in terms of fuel consumption. To take another example: using someone else's toothbrush to brush one's teeth is probably as effective and efficient as using one's own, but many people shudder at the mere thought.

[15]This social-epistemic story about testimony, professional design and proper use only shows

This means that proper-use claims, supported by testimony, express normative facts that are different from, but related to, the normative facts expressed by usefulness claims and function ascriptions. Ascribing to a car the proper function to drive in, expresses not just that a specific or any person has a practical reason to drive the car. Instead, it expresses the fact that any person using the car ought to use it for *driving*, not for another purpose. This 'ought' can be interpreted in different ways. One way is purely social: any member of a society ought to conform to its social rules, apparently including its division of labour between professional designers and passive users. Another, more complicated interpretation relates proper-use claims to practical reasons: one ought to use an artefact properly because one has a reason to believe that proper use is effective. This does not exclude reasons to believe that other ways of using the artefact might be equally successful. Furthermore, it does not mean that one has a conclusive reason to believe that proper use will be effective. As pointed out in section 2.7, proper use might be irrational, for instance when the artefact malfunctions. In particular, in post-hoc evaluations of artefact performance, one might combine a proper-use claim, that one ought to use x to realise g, with a claim that one has reason not to use x to realise g. This brings us back to the original purpose of this section, to clarify the relation between normativity and artefact malfunctioning.

Malfunctioning

We just found that claims that an artefact x is used properly to g, and hence proper function ascriptions, express conditional obligations, namely that any member of a society ought to use x to realise g, not to realise some other goal h. Yet proper-use claims do more than that: they also express the fact that at least some persons (namely those who know a use plan and who possess the requisite skills) ought to be able to use x to realise g. This 'ought' is not a matter of rational expectation. To see why, suppose that one person uses a swivelling chair to stand on, and another person uses it to sit on, and both fail to achieve their goals. Both users may make malfunctioning claims, but only the second may claim to have used the chair properly. In the first case, a post-hoc evaluation expresses the fact that an alternative, self-designed use plan was ineffective (i.e., that amateur designing was irrational) and the fact that the user has a

a minimal, 'default' structure of the relation between designers and users. One may imagine any number of (realistic) scenarios in which consumers override the testimony of professional designers and their proper-use claims. Groups within a community may not just execute other use plans, providing evidence for alternative rational use, but they may elevate alternative use plans to the status of properness. In our analysis, this means that the social status of professional designers is no longer acknowledged: the alternative use plan is still, by definition, the product of designing, but not of *professional* designing. Another intriguing scenario is that in which a professional designer intentionally leaves room for further development of use plans (or even the artefact) by users. Our analysis provides a framework for discussing these issues, but the exact ways in which these conflicting claims to properness are resolved and in which non-professional designing is socially hindered or stimulated remain to be clarified. This clarification is not our purpose here, and the basic scenario suffices to clarify the normativity of proper-use and malfunctioning claims. More detailed research into the social aspects of function ascriptions is presented in, e.g., Scheele (2005, 2006).

reason not to use the chair for this purpose again. There, the normative story ends. This is not so for the second case. There, a rational expectation was also defeated and an effectiveness belief, capacity belief and function ascription have to be revised, but the expectation, belief and ascription were combined with a proper-use claim. All three may have originally been based on testimony, and all three are based, through the proper-use claim, on the division of labour between professional designers and passive users. In this case, 'The chair ought to have been useful for sitting on' minimally expresses that a social expectation has not been met, or that someone has not lived up to social rules or standards. We expect professional designers to construct rational use plans and useful artefacts, not because they have successfully done so in the past, but because this is their social role.

This brings out more of the normativity of malfunctioning claims. A post-hoc evaluation that is combined with a proper-use claim involves an admission that one's previous beliefs about the artefact were false, as well as a claim to the *right* to have those beliefs. In a post-hoc evaluation, you admit that your practical reasons have been defeated, but you also claim that they *should not* have been. Thus, post-hoc evaluations presuppose that you ought to have reasons for goal-directed manipulation of an artefact. This right is based on the social division of labour between professional designers and consumers. On the basis of this division of labour, and the resulting network of social responsibilities, consumers may, in principle, complain about ineffective artefact use, if they believe that they have faithfully executed the designed and communicated use plan.[16] In practice, this right is enforced through a system of liability, repairs, refunds and other redresses.

This normative content of malfunctioning claims is not related to the function ascription as such, but rather to the social and action-theoretical background of the function ascription. Therefore, this particular normative story is a part of the phenomenology of artefact using and designing that need not (and perhaps ought not to) be accounted for by a theory of artefact functions. Thus, our second way of showing how the ICE-function theory deals with malfunctioning claims fades into its background.

[16]In actual practice, this right to use artefacts and to complain about ineffective use is constrained: one has little cause for complaints if a one-euro watch breaks down after one day, or if a light bulb blows out after five years. However, these limitations seem as much a part of the social background of artefact use as the right to use itself.

Chapter 6

Engineering, science and biology

The ICE-theory accounts for two types of functional descriptions of artefacts: function ascriptions relative to use plans and ascriptions of functional roles based on plan-less physicochemical analyses. We put forward the first type as central to technology. In this chapter we continue our discussion of plan-less functional descriptions, and review a number of engineering functional descriptions that seem to challenge the central role of plan-relative function ascriptions. The discussion stays initially within the domain of technology, but provides us with the means to consider also how the ICE-theory fares in dealing with functional descriptions outside this domain.

Engineers, as we noted in section 4.4, may ascribe functions to components of artefacts by analyses of the physicochemical structure of artefacts only. These functional descriptions can nevertheless be reconstructed as function ascriptions relative to use plans, as we prove in section 6.2.

Physics and chemistry seem also to contain functional descriptions: systems are taken as measurement or preparation devices and substances as, for instance, conductors and solvents. These functional descriptions can easily be accommodated in the ICE-theory, as is shown in section 6.3.

The ICE-theory can, however, not be taken as adequately capturing functional descriptions in biology. This negative result leaves us two routes for contributing to the analysis of biological functions. In section 6.4 we criticise attempts to analyse biological functions as 'as-if' technical functions: the ICE-theory shows that such analyses imply the implausible acceptance of a series of teleological concepts within the domain of biology. The second route is a more daring one. We show in section 6.5 that the ICE-theory can be generalised to an overarching uniform function theory in which agents ascribe, relative to goal-directed patterns, functions to items that are part of those patterns. This generalised ICE-theory applies to biology but has the consequence that biological functions are also ascribed by agents and become teleological.

W. Houkes, P.E. Vermaas, *Technical Functions*, Philosophy of Engineering and Technology 1, DOI 10.1007/978-90-481-3900-2_6,

6.1 Plan-less function ascriptions

The concept of use plans is central to our action-theoretical approach towards artefacts. We analyse artefact using and designing in terms of use plans, and we analyse functional descriptions of artefacts primarily as ascriptions of functions relative to those plans. We also acknowledge a second, plan-less type of functional descriptions — ascriptions of functional roles based on physicochemical analyses of artefacts — but we do not put forward this second type as central to functional descriptions in technology. Function ascriptions relative to use plans are, in the ICE-theory, supposed to account for the bulk of functional descriptions of artefacts.

Engineering functional descriptions

The concept of use plans is our invention, and currently not in use in descriptions of technology. We mentioned in section 2.4 a few engineering methodologists who adopt action-theoretical bases for their accounts of designing,[1] and we may attempt to analyse these accounts in terms of our use plans. In general, however, design methodologists tend to suppress references to actions by users; methodologists rather talk about the goals of users and on the creation of new functional objects, and characterise designing accordingly.[2] Consequently, our claim that functional descriptions of artefacts are primarily function ascriptions relative to use plans is bound to be reconstructive: in real life, agents are clearly capable of ascribing functions without considering use plans or user actions. This reconstruction becomes problematic as soon as existing function ascriptions cannot be plausibly analysed as related to use plans. At the end of chapter 4 we already gave some examples of such plan-less function ascriptions by observers and structural analysts of artefacts and here we continue their analysis.

A common picture of engineering designing confirms the idea that engineers regularly ascribe functions without considering use plans. In this picture, designing consists of two more or less distinct parts. In a first intentional part engineers consider the users and their goals in order to find out what functions the artefacts to be designed should have for the users to realise the goals. And in a second structural part, engineers focus on the artefacts for determining what their physicochemical structure should be like in order for these artefacts to have the required functions. This division suggests that engineers can ascribe functions to the artefacts without considering the ways in which they are used nor the goals involved in these uses; in the second part of the design process engineers analyse the physicochemical structure of artefacts, and those analyses seem sufficient for ascribing functions to these artefacts.[3]

These cases of functional descriptions on the basis of physicochemical descriptions of artefacts alone, apparently undermine our view that the bulk of

[1] Hubka and Eder (1988); Roozenburg and Eekels (1995, §4.3); Eekels and Poelman (1998, ch. 4); Brown and Blessing (2005).

[2] E.g., Gero (1990), Stone and Wood (2000) and Pahl, Beitz, Feldhusen, and Grote (2007).

[3] See (Vermaas 2009a) for an analysis of different ways in which design reasoning is simplified in models of designing.

functional descriptions of items in the technical domain are related to use plans. It seems that, if we want to retain our ICE-theory, we should acknowledge that plan-less ascriptions of functional roles are the rule in technology, and not a secondary functional notion.

In this chapter, we address this challenge. We retain our function theory and our view that plan-relative function ascriptions comprise the bulk of functional descriptions of artefacts. In section 4.4, we already argued that function ascriptions by observers fit our definitions of plan-relative function ascriptions. There, we also sketched an argument to the following effect: functional descriptions of artefacts and their components, given by engineers in their roles as structural analysts, can in some, but not all, cases be taken as ICE-function ascriptions. This argument hinges on the possibility for engineers to conceal — or 'bracket', as we call it in this chapter — the reference to use plans in their function ascriptions. In the next section, we consider in detail engineering functional descriptions based on analyses of the physicochemical structure of artefacts, and give the proofs promised in section 4.4. As a result of the discussion, the challenge to our theory largely dissolves: in general, engineering functional descriptions of components based on analyses of their physicochemical structure, can be reconstructed as ICE-function ascriptions. However, when engineers give functional descriptions of artefacts as a whole, based on analyses of their physicochemical structure, these descriptions cannot in general be reconstructed as plan-relative ICE-function ascriptions.

Functional descriptions in science and biology

The central role of use plans also makes the ICE-theory vulnerable if it is applied to functional descriptions in the natural sciences. In physics and chemistry, one can find functional descriptions of substances; copper, for instance, may be described as a conductor and ethanol as a solvent of fats. It will be hard to construe such descriptions as plan-relative function ascriptions in the sciences, since use plans seem out of place in physics and chemistry. Biology, which also features functional descriptions, seems an even more hostile environment for use plans.

A response against the charge that the ICE-theory cannot accommodate functional descriptions in the sciences is to deny that these descriptions are purely scientific. It can be claimed that functional descriptions of substances in physics and chemistry are technical descriptions of those substances, and that they can therefore be understood as plan-relative function ascriptions after all. A line of defense against the charge that our function theory cannot accommodate biological functional descriptions is to deny that it should. The ICE-theory is based on our use-plan approach. This approach makes sense in the domain of technical artefacts that are used and designed by agents to achieve goals. The ICE-theory should therefore account for functional descriptions of artefacts, and not of items for which the use-plan approach does not make sense, such as biological items. This last claim implies a conceptual distinction between the functions of technical artefacts and those of biological items, and thus enforces an inter-domain pluralism.[4]

We take these lines of defense, but not without considering how the ICE-theory is related to these scientific and biological functional descriptions. In section 6.3 we focus on plan-less functional descriptions of substances in physics and chemistry, and in section 6.4 we focus on biology and discuss whether the ICE-theory can be a tool for understanding biological functional descriptions.

Components, semimanufactured products and substances

The discussion of plan-less function ascriptions in technology, physics and chemistry is not merely meant for defending the ICE-theory, but also for showing its flexibility and versatility. We argue that the ICE-theory can accommodate functional descriptions of artefact components in engineering, and functional descriptions of substances in science, by taking both as function ascriptions. This may be seen as an advantage over some other function theories.

A characteristic of components and substances is that they typically have various capacities for which they are rationally and properly used. Flat metal rings that fit a bolt have the capacity to make the head of the bolt wider. These rings are therefore applied for distributing the force exercised by the head on the material it fastens, and for letting the bolt fit a hole wider than its head. Moreover, the rings have a thickness, which allows them to be used even if the bolt is slightly too long for properly fastening the material. A chemical substance such as ethanol dissolves fats and is used for this capacity. But it also expands gradually when its temperature rises and is therefore applied in thermometers; it is a mild intoxicant and is therefore added to drinks; and it is flammable and is therefore used as fuel in engines and lamps. Semimanufactured products such as flour and bricks are similarly multifunctional. Flour nourishes and is applied for that in, say, bread. But flour has also the capacity to bind fluids and is used for making sauces; it has a white colour and was therefore once used for powdering hair.

With the ICE-function theory, one can ascribe all the corresponding functions related to the different uses of components, semimanufactured products and substances, without necessarily taking one of them as proper and the others as accidental. Ethanol can be ascribed the function of dissolving fats relative to a surface-cleaning use plan, and it can be ascribed the function of expanding relative to a temperature-measuring use plan. This advantage is shared with the causal-role function theory and with those intentional theories in which the functions of items are related to user intentions. Evolutionist function theories and intentional function theories in which designer- or constructor-intentions determine the functions of items run into problems when applied to components, semimanufactured products and substances. For which capacities are those flat rings reproduced or designed? For which capacities are flour and ethanol reproduced? One possibility is that components and semimanufactured products are designed for all their current rational uses, but this may amount to assigning an unreasonably broad belief base to the designers concerned. Alternatively, one

[4] See also the discussion in section 3.5, especially footnote 41 on page 67.

may assume that components and semimanufactured products are designed for only one or two of their current uses, but this has the consequence of turning the functions related to the remaining uses into accidental ones, regardless of how proper these remaining uses may be. More fundamentally, it seems counterintuitive to describe substances such as ethanol as *designed* products in the first place, since they occur naturally. A similar dilemma arises for evolutionist theories in which functions are related to use. Are components, semimanufactured products and substances reproduced for all the capacities for which they are rationally used, or only for a few of them? In the first case, the ethanol in your last tequila also had the function to expand when its temperature rises. In the second case, this function of ethanol is presumably always accidental.

These problems for the intentional and evolutionist theories need not be insurmountable. One can assume that components, semimanufactured product and substances are designed/reproduced for a few of their current uses, and then give the functions related to the other current uses an 'upgraded' accidental status; one can, for instance, take these 'other' functions as *ongoing exaptations*, as Preston (1998b) does (see section 3.5). The ICE-function theory, however, provides a straightforward way of ascribing all current functions to components, semimanufactured products and substances, as is demonstrated by the different cases that we consider when discussing plan-less function ascriptions.

6.2 Engineering

In this section, we consider engineers who give functional descriptions of artefacts on the basis of an analysis of their physicochemical structure alone. We first discuss such 'physicochemical' functional descriptions of *components*. We argue that, in this case, the relevant definition of function ascriptions can in general be simplified and re-interpreted as one that refers to physicochemical concepts alone. This shows that functional descriptions of components, although based on analyses of their physicochemical structure, cannot just be taken as ascriptions of functional roles, but can also be reconstructed as function ascriptions. This twofold understanding is not always possible; we give a criterion under which physicochemical functional descriptions to components can be taken as plan-relative function ascriptions. Secondly, we consider engineering functional descriptions of *artefacts as a whole* (including semimanufactured products). For this case, we show that the relevant definition of function ascriptions cannot in general be simplified as one that refers to physicochemical concepts alone. Hence, functional descriptions of artefacts as a whole that are based on analyses of their physicochemical structure alone, cannot, in general, be taken as plan-relative function ascriptions.[5]

Engineering function ascriptions to components

Artefacts typically consist of components and, typically, these components are described in functional terms. Metal rings have the function to distribute the

[5] Vermaas (2006).

forces exercised by bolts, switches have the function to interrupt electrical currents, pipes to transport substances in engines, and so on. Now consider an engineer who gives a functional description of a component, based on an analysis of the physicochemical structure of the component and of the artefact as a whole. If this functional description is to be understood as an ICE-function ascription, it should be possible to derive, from the definitions that comprise this theory, a definition for function ascriptions to components that refers to physicochemical concepts alone. Here, we give such a derivation, which provides a way for understanding physicochemical functional descriptions of components as ICE-function ascriptions.

Let us start by noting that the use-plan approach to artefact using and designing also applies to components. Components have use plans that are developed by designers and executed by users. The difference with ordinary artefacts is that 'component users' typically also play the role of designers, namely of artefacts composed of components: components are items that are used by designers to compose artefacts with. The use-plan approach thus already provides an action-theoretical background for considering components: a component c has a use plan q which is developed by a component designer d, and communicated to other designers d', who play the role of users of c when they design an artefact x by using the component c according to its component plan q. This handful of variables is introduced to make the necessary distinctions between an artefact x and its components c, and between their respective use plans p and q. In the example of the pipe in an engine, for instance, q is aimed at transporting a fluid, and p at producing propulsion. The component designer d and the artefact designer d' may be the same agents but are typically not. Engineers, when they design, often use existing components, designed by other engineers and made available within the profession through manuals and technical handbooks — thus constituting the communication chain which distributes the use plans of components.

This action-theoretical reconstruction provides the background against which an engineer e who considers component c of an artefact x may be taken as ascribing to this component a function to ϕ. More specifically, this reconstruction provides a use plan q relative to which this function can be ascribed, and a communication chain by which the engineer can appeal to testimony for ascribing to c the relevant capacity to ϕ. As a consequence, this function ascription refers to all kinds of non-physicochemical concepts, such as use plans, communication and testimony.

To save the appearance of a physicochemical functional description, we have to suppress these references to the action-theoretical background. This can be done in three steps. First, we assume that engineers take the action-theoretical background as a default, to which no explicit references need to be made when ascribing functions to components. That is, when engineers ascribe such functions, they focus on those conditions that are part of the definitions of function ascriptions, and put between brackets the conditions — given in table 4.1 on page 84 — that characterise their own roles in the communication chain. This bracketing suggests that engineers ascribe functions to components as justifiers, or possibly as designers; the definition for function ascriptions by passive users contains direct references to other agents in the communication chain of the use plans of the components, which is hard to square with bracketing this chain. This consequence makes sense, for engineers are generally seen as agents who can explain how artefacts work on the basis of their own knowledge, instead of relying on fellow engineers for these explanations. So, let us assume that an engineer e ascribes as a designer or justifier the function to ϕ to the component and let the account A by which e justifies his beliefs consist of scientific and technological knowledge.

As a second step, we choose a formulation of the use plan q for c that minimises references to non-physicochemical concepts. This formulation is: 'compose c, c', c'', … in configuration k in order to obtain an artefact x with the physicochemical capacity to Φ'. This is a use plan for a number of components, including c, and its potential executors are designers who want to obtain an artefact (which may be a component itself) with the capacity to Φ. By choosing this formulation, we do not narrow the scope of our theory: the use plan for every component can have this form. Substitution of this reformulated use plan in the definition for function ascriptions by designers or justifiers gives the following result:

> An engineer e justifiably ascribes the physicochemical capacity to ϕ as a function to the component c, relative to the plan q to compose c, c', c'', … in configuration k in order to obtain artefact x with the physicochemical capacity to Φ, and relative to scientific and technological knowledge, iff:
>
> I. e has the belief B_{cap} that c has the capacity to ϕ;
> e has the belief B_{con} that q leads to its goal of an artefact with the capacity to Φ due to, in part, c's capacity to ϕ; and
>
> C. e can justify B_{cap} and B_{con} on the basis of scientific and technological knowledge.

This definition still refers to the component use plan q, to its goal (obtaining an artefact x with the capacity to Φ), and to the actions that comprise q (compose c, c', …). In the third and final step, we remove or at least suppress these references. This can be done by interpreting the use plan q, not as a set of instructions formulated by the component designer d to be carried out by engineers e who want to have an artefact with the capacity to Φ, but as a technological claim about how capacities of components 'add up' to capacities of configurations of these components. On this reinterpretation, the component use plan q is transformed into a structural rule: 'components c, c', c'', … in configuration k constitute an artefact with the physicochemical capacity to Φ'. This reinterpretation of the component use plan puts between brackets its action-theoretical background: the component plan no longer comprises actions, aimed at a goal, which is developed and communicated among designers. Instead, the component plan expresses technological knowledge about the physicochemical structure of artefacts and their components, knowledge that is part of the account A invoked when engineers design artefacts. If this reinterpretation is substituted in the above definition, the component use plan is replaced with the mentioned rule and the contribution belief becomes a belief 'B_{con}' about this rule. One then obtains the following ICE-theory definition for engineering function ascriptions to components:

> An engineer e justifiably ascribes the physicochemical capacity to ϕ as a function to the component c relative to the composition of c, c', c'', ... in configuration k of an artefact x with the physicochemical capacity to Φ, and relative to scientific and technological knowledge, iff:
>
> I. e has the belief B_{cap} that c has the capacity to ϕ;
> e has the belief 'B_{con}' that x has the capacity to Φ due to, in part, c's capacity to ϕ; and
>
> C. e can justify B_{cap} and 'B_{con}' on the basis of scientific and technological knowledge.

(We gave this definition in section 4.4, anticipating the current analysis of function ascriptions by engineers.)

By this three-step process, we bracket all references to non-physicochemical concepts in ICE-function ascriptions to components. This shows that a physicochemical functional description of a component c by an engineer need not always be an ascription of a functional role. It may also be interpreted as a function ascription relative to a bracketed use plan.

A physicochemical functional description is only interpretable in both ways, i.e., as an ascription of the functional role of ϕ-er and as an ascription of the function to ϕ, if this bracketing process makes sense. That is, the assumption of an action-theoretical background — consisting of designers d of the component c who developed its use plan q relative to which the capacity to ϕ can be ascribed as a function to c — must be plausible; if it is not, the assumption that this background is bracketed is untenable.

The example of the ascription of the functional role of detonator to an electrical system during an investigation of an accidental explosion in an industrial installation, provides a case in point: it may be an integral part of a functional description of a component that there are no designers who have developed a corresponding use plan. This bars us from interpreting the ascription of a functional role as a bracketed function ascription. This type of genuinely plan-less functional descriptions becomes relevant in section 6.4, when we discuss the application of the ICE-function theory to biology.

Engineering function ascriptions to artefacts as a whole

Let us turn to another case of physicochemical functional descriptions. Consider an engineer who gives a functional description of an artefact as a whole, based on an analysis of the artefact's physicochemical structure. Can such a description be taken as an ICE-function ascription in which references to use plans are bracketed? Can we, again, show that our definitions for function ascriptions to artefacts can be made to refer to physicochemical concepts only? For some types of artefacts, the answer is positive. We have, however, not been able to find a general positive answer: unlike for components, it seems impossible to find, for all artefacts, reformulations of their use plans in which references to non-physicochemical concepts are sufficiently minimised.

Let e again be the engineer, x the artefact, p its use plan and d the designer of x and p. The engineer may be the designer d or another agent in the communication chain. Assume that the engineer brackets the action-theoretical background and, as a designer or justifier, ascribes to x the function to ϕ, relative to scientific and technological knowledge (account A). The definition for function ascriptions by designer or justifiers still refers to the use plan p, which comprises actions that users have to carry out, and which is also characterised by the goals that users realise by executing p. These action-theoretical notions can be bracketed by reformulating p as a technological rule instead of as an set of instructions issued by the designer d. Yet this interpretation does not fully suppress the nature of the plan. Even as a rule, p still refers to actions that an agent has to take with x in order to achieve specific goals. For some artefacts, it may be possible to rephrase these actions and goals in physicochemical terms. Consider, for instance, artefacts that have use plans with a goal state that is straightforwardly characterised as a physicochemical state of affairs S and with actions that barely need their executors' attention. Examples are air fresheners such as sachets of lavender or more modern perfumed gadgets. These artefacts should simply be bought and retained somewhere in the house, and then distribute aromatic substances throughout the house without the need of further actions on the user's part. The use plans for these artefacts may be: 'buy and retain x for obtaining goal S'. Such plans are easily re-interpreted as technological rules of the type 'x leads to state of affairs S'. Substituting this rule into our definition, engineering function ascriptions to these artefacts get the following form, which mentions only physicochemical terms:

> An engineer e justifiably ascribes the physicochemical capacity to ϕ as a function to an item x relative to the rule that the presence of x leads to S, and relative to scientific and technological knowledge, iff:
>
> I. e has the belief B_{cap} that x has the capacity to ϕ;
>
> e has the belief 'B_{con}' that the presence of x leads to S due to, in part, x's capacity to ϕ; and
>
> C. e can justify B_{cap} and 'B_{con}' on the basis of scientific and technological knowledge.

As soon as one considers other cases, action-theoretical elements creep back in. Take, for instance, artefacts with plans that are usually not spelled out in the communication between agents. The reason may be that the plan is sufficiently well-known to make further communication about it superfluous. Examples are brooms, books, bridges and vacuum cleaners. Alternatively, there may be too many use plans for the artefacts, which makes communicating them more or less hopeless. In such cases, it is typically left to the users to determine which plan to carry out. Examples of this type are semimanufactured products such as flour, bricks, planks and sour cream. In both types of cases, the use plans for the artefacts may be colloquially described as 'use x for goal g', where the goal g may be stated with some precision — when x has a standard use, e.g., 'read this book to learn about artefact functions'; or the use plans may be kept vague — when x has multiple uses, e.g., 'use sour cream to improve all your cold dishes'. In any case, the actions that comprise this plan for achieving g are omitted. With a bit of arm-twisting, one might re-interpret this plan as a technological rule 'x leads to state of affairs S'; success would mean that the definition of engineering function ascriptions has again the form given above. Still, it cannot be maintained that this definition is phrased in physicochemical terms alone. The contribution belief required by the I-condition, for instance, is still a belief about those uncommunicated actions: it refers to the actions users have to take in order to achieve the state of affairs S. Although one might say that the I-condition only implicitly refers to these actions, they still have to be considered explicitly if an engineer e is to justify this contribution belief, as required by the C-condition.

Hence, functional descriptions of whole artefacts on the basis of an analysis of their physicochemical structure alone, can typically not be reconstructed as plan-relative function ascriptions. Hence, if such functional descriptions exist in engineering, they cannot be accommodated by the ICE-theory, except in some cases.

The conclusion that can be drawn from the discussion in this section, is that physicochemical functional descriptions of components need typically not be taken as plan-less functional descriptions. These functional descriptions allow a reconstruction as plan-relative function ascriptions. Only those few cases that do not satisfy our criterion should be taken as genuine ascriptions of functional roles. That there are physicochemical functional descriptions of components therefore does not undermine our claim that plan-relative, ICE-function ascriptions account for the bulk of functional descriptions of artefacts. Physicochemical functional descriptions of whole artefacts cannot, however, in general be reconstructed as function ascriptions relative to use plans. If such physicochemical functional descriptions exist, they undermine our claim and even our ICE-theory: they can typically be taken as neither plan-relative function ascriptions nor as ascriptions of functional roles, since the definition for those latter ascriptions applies to components only.

6.3 Physics and chemistry

Functional descriptions also play a role in physics and chemistry. In these sciences, preparation and measurement devices can be described in functional terms, and substances are sometimes characterised by functional names such as 'conductor' or 'solvent'. The ICE-theory cannot be straightforwardly applied to these functional descriptions in the sciences, since 'use plan' is not a notion that is part of physics or chemistry. One can, however, argue that these functional descriptions should be taken as technical ones, not as descriptions that belong to physics and chemistry proper. The distinction between science and technology is a flexible one, as revealed in studies on the interplay between scientific and technological research[6] and on the role of instrumentation in science.[7] This flexibility allows us to claim that whenever devices are described scientifically, they are described as having specific physicochemical capacities — a preparation device can interact with systems so that these systems acquire a specific well-defined state; a measurement device can establish correlations between the states of systems and the device — and whenever they are described technically, these capacities can also be ascribed to the devices as functions.[8] Similarly, one may argue that a scientific description of physical and chemical substances consists merely of a description of the physicochemical properties and dispositions of these substances and that, when considered from a technical point of view, some of these properties or dispositions can be ascribed as their functions. Describing copper as a conductor, for instance, implies from a physical point of view only that specific electrons of copper atoms can move freely from one atom to another, and implies from a technical point of view that copper can be ascribed the function of conducting an electrical current.

If one accepts this, functional descriptions in physics and chemistry do not pose a problem to the ICE-theory. If preparation devices, measurement devices and substances are described functionally, they are considered from a technological point of view, and our action-theoretical analysis applies to these devices and substances. We can thus describe them as having use plans — preparation devices have use plans aimed at producing well-defined samples, and ethanol has use plans aimed at, for instance, removing fats from surfaces and at indicating the temperature in thermometers — and accommodate in the ICE-theory the functional descriptions of these devices and substances as ascriptions of functions relative to these use plans.

Plan-less designing of substances

One phenomenon in science challenges not so much the ICE-theory as the underlying use-plan approach. This phenomenon is the synthesis of new substances, in physics and especially chemistry, with the aim of having substances with specific physicochemical capacities. In some cases, this synthesis is motivated technolog-

[6]See, e.g., Faulker (1994).

[7]The interest in experimentation and instrumentation is rapidly increasing among philosophers of science, after groundbreaking works such as Hacking (1983), Galison (1986) and Gooding, Pinch, and Schaffer (1989). Some recent contributions to this growing field are collected in Radder (2003).

[8]The position that preparation and measurement devices are not themselves part of science is less plausible. Measurement devices, for instance, play a fundamental role in quantum mechanics, since the states this theory assigns to physical systems primarily have a meaning in terms of measurements performed on those systems, instead of in terms of a description of the physical properties of those systems. Hence, devices that can be ascribed functions are sometimes essential elements of scientific theories. Another example may be the Carnot-cycles that are central to thermodynamics.

ically: producing materials that are superconductive at room temperature, for instance, or synthesising new chemical substances with desired pharmaceutical uses. But in other cases, the motives are more purely scientific: the synthesis of elements with an atomic number higher than 100, is, for instance, a purely scientifically motivated enterprise, and so is producing very large molecules and crystals.

In science, this synthesis of new substances is sometimes described as designing. This defines a type of designing that, especially when it is done without any technological motivation, does not fit our use-plan reconstruction of designing. There are, of course, plans or procedures for synthesising these new substances, but these are not use plans for the substances themselves, but rather ones for all kinds of other substances and instruments. Hence, although it can be argued that there are no functional descriptions in physics and chemistry that are at odds with the ICE-theory, there is a type of designing in these sciences that does not fit the use-plan approach. In that sense our approach is indeed confined to technical artefacts.

6.4 Biology

Biology is a scientific discipline that prominently features functional descriptions. Not coincidentally, it is the main focus in most philosophical function theories. The ICE-theory has been developed specifically for the technical domain. To some extent, we intend to move away from the biological bias in research on functions and concentrate on issues particular to technical artefacts. Still, the current focus on biology in the literature inevitably raises the question how the ICE-theory fares in accommodating functional descriptions in biology.

An additional reason for taking up this question is that function ascriptions to artefacts are relevant to the literature on biological functional descriptions. In this literature, function ascriptions to artefacts are often considered to be philosophically unproblematic and therefore often act as a standard against which the 'checkered' understanding of function ascriptions in biology is measured.[9] Furthermore, some authors actively import technical functions into the biological domain to make sense of biological function ascriptions; a well-known example is the *intentional stance* as defended by Daniel C. Dennett (1971, 1990, 1995), in which functional descriptions in biology are made sense of by interpreting organisms as if they were artefacts.

By now, we hope to have shown that function ascriptions to technical artefacts are not at all easy to understand: the standard that technology provides for biology may be a bit checkered itself. Although our primary purpose is to analyse technical functions, we accept the challenge of checking whether the ICE-function theory is useful for understanding biological functional descriptions. More specifically, we examine whether the ICE-function theory accommodates biological function ascriptions and, if not, whether it can be generalised

[9] Ariew and Perlman (2002, p. 1).

to accommodate those functions. In doing so, we also review the idea that biological function ascriptions can be understood by taking them *as if* they were function ascriptions to technical artefacts.[10]

Application of the ICE-function theory to biology

A cursory assessment of the different definitions that make up the ICE-theory shows that, if this theory is applicable to biology at all, it is by the definitions for ascriptions of functional roles (see page 100 for all definitions). These definitions are phrased in concepts that make sense within biology, which cannot be said for the definitions of function ascriptions by designers, justifiers and passive users. On closer inspection, the definition for function ascriptions to components turn out not apply to the biological domain either.

The definitions for function ascriptions to artefacts by designers, justifiers and users are not literally applicable to biological items because they refer to use plans, a notion that seems alien to biology. More generally, these definitions presuppose an action-theoretical background of artefact using and designing. This background is clearly not available in the biological domain. According to current neo-Darwinian orthodoxy, there are neither intentional designers of biological items nor use plans for these items that were developed to contribute to goal realisation by other agents; even adherents of creationism should find it somewhat steep to accept such a background.

This point merits spelling out. The use-plan approach does not merely imply that artefacts are items that are intentionally selected or shaped by agents, but provides a far richer account. On it, all artefacts are embedded in use plans as means for realising the goals associated with those plans; designers develop those plans and communicate them to other agents, thus creating communication chains between agents who distribute the use plans. Hence, merely taking biological items as designed or created by a deity is not sufficient to reproduce the action-theoretical background of our definitions for function ascriptions. One also needs to take them to be designed as means to human ends and to accept that the designer(s) informed humans about these uses.[11]

These features are shared by the definition of function ascriptions to components. The action-theoretical background described is bracketed in this definition, but it is still presupposed. Function ascriptions to components are distinguished from ascriptions of functional roles, because the former presuppose that there exists a use plan for the component. The agent ascribing the function need not know this use plan, but must suppose that it exists and that

[10]Some functional descriptions of biological items are associated with human interferences with biological organisms, ranging from breeding to genetic engineering. Such descriptions can be counted as belonging to both the biological and technical domain. In order to consider pure biological cases, we ignore these hybrids.

[11]McLaughlin (2001, ch. 7) rejects the analogy between function ascriptions to artefacts and biological items by arguing that artefacts and organisms are associated with different goals and that designing differs from natural selection. Our position, that function ascriptions to artefacts presuppose an action-theoretical background that is absent in biology, may be taken as a third argument.

the agents who are part of the communication chain set up by the designer of the component know it. One may ignore these points and apply the definition for function ascriptions to components to biological items anyway. In that case, one treats this definition as a *truncated* function ascription to components; the action-theoretical background that is presupposed by it is amputated and becomes a ghost limb to the ascription. In the ICE-theory, such truncated function ascriptions in fact count as ascriptions of functional roles.

The upshot is that biological items can be ascribed only functional roles in the ICE-theory. If the desiderata for function ascriptions to biological items include at least analogues of the proper-accidental and malfunctioning desiderata,[12] the ICE-theory is inadequate to biology. For it does not distinguish between proper and accidental functions, but indiscriminately casts functional descriptions in biology as ascriptions of functional roles. Moreover, in the ICE-theory, one cannot ascribe the capacity to detect light to the eyes of a blind person as soon as one has the belief that these eyes do not have this capacity.[13]

This does not rule out that functional descriptions in biology can be understood as descriptions that arise by taking biological items *as if* they were artefacts. But applying the ICE-theory to biology does show that the presuppositions of this position are much more detailed and possibly much less attractive than presented by some authors defending this position. In the literature, the presupposition is often taken to be that the biological item is designed by a rational agent in response to problems posed by its biological context. Designing is then taken in its simple, object-oriented sense of intentionally determining the physicochemical structure of the object concerned, not in the action-theoretical sense that we propose. For Lewens (2004), who analyses what he calls the 'artefact model' of evolution as 'the approach to the organic world that treats it as though it were designed',[14] the antecedent seems indeed design in this simple sense. Matthen (1997), in his description of the 'product analogy' to understand biological functional descriptions, takes one step towards our action-theoretical sense of designing, by requiring that users are also identified. These users are not agents, but the organisms that benefit from the item to which the function is ascribed. The user of a liver, for instance, is the body that uses it to metabolise fats.[15]

Suppose that one takes a biological item as if it were designed, in the sense of having its physicochemical properties intentionally determined by an agent. This is not yet sufficient for applying the ICE-theory to the item. It is sufficient for applying an intentional function theory, say, the one by Neander (1991b). But such an application will immediately confront as-if accounts for biological functional descriptions with the proliferation problem; biological items may then be ascribed wildly unsupported functions, or functions related to 'hidden motives' of the as-if designer: the liver may, say, be ascribed the function to display nature's 'ingenuity', or ichneumon wasps the Schopenhauerian function

[12] See, e.g., Lewens (2004, pp. 88–89).
[13] The argument given in section 5.2 may help resolve this point.
[14] Lewens (2004, p. 39).
[15] Matthen (1997, p. 31).

of showing nature's inherent cruelty.

Applying the ICE-function theory to biological items requires more detailed presuppositions. Biological items should be embedded in the action-theoretical background discussed above. One has to suppose that the liver has been selected by agents as part of the development of a use plan for the liver. This plan was meant for other designers and communicated to those other designers. The goal of this plan may have been the composition of organisms as a whole. If those other designers decided to carry out this plan by designing organisms, say mice, they again did so as part of the development of a use plan for the mice, and communicated this mice use plan to other agents, who then count as users of the mice.

On first sight, this more detailed antecedent need not seem problematic for those who take biological functional descriptions as as-if artefact-function ascriptions. After all, they did not intend to paint a correct picture in the first place, so adding a few outlandish brush strokes seems to do little harm. But, in the end, these extra assumptions may decrease plausibility, because biologists who describe biological items in functional terms may become more reluctant to take those items as if they are artefacts. The users of a biological item can, for instance, not be identified as the organisms containing that item, as Matthen wants it: the use plan of a liver is, in the ICE-theory, communicated to other designers, not to the body that contains the liver. Lewens' view, that the artefact model of evolution 'only becomes practically applicable and psychologically attractive to inquirers' for items that are the result of processes that create 'systems with traits that have the kind of functional complexity reminiscent of designed objects',[16] also becomes difficult to maintain. It seems plausible on the original presupposition: when biological items appear similar to designed objects (in Lewens' sense) to biologists, then these biologists may be prone to take them as if they are designed objects (in Lewens' sense) and take the original presupposition as plausible. But resemblances between biological items and designed artefacts (in Lewens' sense) are no longer sufficient to accept the more detailed presuppositions; instead, biologists need to take biological items as similar to objects that have use plans and that are designed to be used. The latter as-if assumptions appear far less plausible than the former, more modest ones.

6.5 A biological and generalised ICE-theory

Even if the ICE-theory does not apply to biology, one may still try to construct an 'ICE-like' function theory for biology and/or a generalised theory of which both the ICE-theory for artefacts and its biological counterpart are instances. Let us call this biological counterpart of the ICE-theory the b-ICE-theory, and the generalised theory the g-ICE-theory. It is not difficult to construct such theories: just take any function theory adequate to the biological domain as the b-ICE-theory and define the g-ICE-theory as the theory that yields ICE-function ascriptions for the technical domain and b-ICE-function ascriptions for

[16] Lewens (2004, pp. 119–120).

the biological domain. Since this manoeuvre seems a cheap trick, which does not result in an 'ICE-like' theory for biological function ascriptions, we do not carry it out here. Instead, we formulate one example of an 'ICE-like' b-ICE-theory. By this, we prove that there exist function theories for biology that are similar to the ICE-theory. Furthermore, we show that, if one indeed takes function ascriptions to artefacts as a standard for function ascriptions in biology, one may obtain results that differ considerably from currently existing proposals.

One way of transposing a theory from one domain to another is by keeping the structure of the theory intact and translating those of the theory's key concepts that are particular to the first domain into analogue concepts that apply to the second domain. This procedure is not without ambiguities nor does it guarantee success. The etiological theories by Millikan (1984, 1993) and by Neander (1991a, 1991b), for instance, are geared to the biological domain but also made applicable to artefacts by translating, among others, the biological concept of reproduction into a technical analogue. This has not only led to two different translations in the writings of Millikan and Neander — one-shot selection by designers and long-term selection processes by, e.g., users — but also to function theories for technical artefacts that fail to satisfy our desiderata (see the discussion in sections 3.2, 3.4 and 3.5).

For arriving at an ICE-like theory for biological function ascriptions, translating the key concepts of artefact and use plan to biological concepts is not sufficient: the detailed action-theoretical background of the ICE-theory also needs to be transposed. This means that counterparts should be found for concepts such as designing, using and communication. A first attempt at this task could be to take use plans as goal-directed patterns of behaviour of organisms, to take designing as the process of natural selection of those patterns,[17] communication as the passing on of the results of that selection through genetic information, and using as the manifestation of that genetic information in new organisms. This translation leads to an evolutionist function theory, which would place such a b-ICE-theory in good company within the biological domain.

One problem remains. In the ICE-theory designing and using do not just refer to processes, but also define the roles that agents may take when ascribing functions to artefacts. Hence, the biological counterparts of these key concepts should be processes and also define the roles taken by agents when they ascribe functions to biological items. Furthermore, these two roles must be different: designers start communication chains that support function ascriptions, whereas users prolong these chains.

[17] This may bring to mind Krohs' (2009) proposal to generalise the concept of design to one that also has as an instance a non-intentional notion of design applicable to biological items. In the light of our use-plan analysis, this proposal once again takes designing in its limited, object-oriented sense of intentionally determining the physicochemical structure of the object. For artefacts, Krohs characterises design in terms of blueprints, which determine physical characteristics, but not ways of use. This forces us to come up with an alternative transposition to biology. The function theory Krohs derives from his general notion of design was discussed in section 3.5.

A discoverer d or justifier j justifiably ascribes the physicochemical capacity to ϕ as a function to an item x relative to a behavioural pattern p for x, and relative to account A, iff:

I. d/j has the belief B_{cap} that x has the capacity to ϕ;
d/j has the belief B_{con} that p leads to its goals due to, in part, x's capacity to ϕ; and

C. d/j can justify B_{cap} and B_{con} on the basis of A.

A layperson l justifiably ascribes the physicochemical capacity to ϕ as a function to an item x relative to a behavioural pattern p for x, and relative to testimony T, iff:

I. l has the belief B_{cap} that x has the capacity to ϕ;
l has the belief B_{con} that p leads to its goals due to, in part, x's capacity to ϕ;
l believes that a discoverer d or justifier j of p has B_{cap} and B_{con};

C. l can justify B_{cap} and B_{con} on the basis of T;
l can justify on the basis of T that d/j has B_{cap} and B_{con}; and

E. l received T that d/j has B_{cap} and B_{con}.

Table 6.1: The b-ICE-definitions for function ascriptions to biological items

A biological ICE-function theory

The following translation has all required features. Let designing be the process of *discovering* or *first describing* a particular set of behaviours of an organism as a goal-directed pattern p. The biologist who is engaged in this process then takes the role of discoverer d. As a discoverer, she communicates the discovered pattern to other biologists who then learn about it and can 'use' it for better understanding the organism and the behaviour it engages in. This process of learning is the biological counterpart of the process of using and defines the agent role of layperson, in analogy to the expert-layperson distinction in social epistemology.

These translations transpose the full action-theoretical background of technical functions into an 'epistemic' background relative to which biologists may ascribe functions to biological items. Similar to function ascriptions in technology, biologists may ascribe a physicochemical capacity to ϕ as a function to a biological item relative to the pattern p if they have a justifiable effectiveness belief B_{eff} about the pattern based, in part, on a justifiable belief B_{cap} that the item has the capacity to ϕ, and a justifiable belief B_{con} that this capacity to ϕ contributes to the effectiveness of the pattern.

The function theory obtained in this way again consists of two definitions for function ascriptions. They are listed in table 6.1;[18] justifiers j are biologists who did not discover the behavioural pattern p relative to which biological func-

[18]If this b-ICE-function theory is spelled out more precisely, one should also transpose the definitions that characterise the different agent roles in technology (see table 4.1 on page 84) to biology. These transposed definitions, which are presupposed by the definitions of b-ICE-function ascriptions given in table 6.1, then fix the agent roles of discoverer, justifier and layperson.

tions are ascribed, and are laypersons in that sense, but who can nevertheless justify the different beliefs underlying the function ascriptions. One can develop this b-ICE-theory in analogy with the ICE-theory by showing that for function ascriptions to organs, one can define a physicochemical function ascription in which references to goals, behaviour and the epistemic background described above are bracketed; and one can enrich it with the ascription of pattern-less functional roles.

It is not our aim to fully develop this theory here, or to assess it with respect to what may be expected from function theories for biological items. We limit the discussion to a few general remarks.

First of all, the b-ICE theory may seem odd compared to existing theories of biological functions. It is not a theory about biological items *having* functions independent of the beliefs of biologists, but one about justifiable *function ascriptions* to biological items. One may therefore take the theory for a merely (social) constructivist one, in which function ascriptions are determined by whatever biologists may make each another believe. However, both these observations are too quick. Firstly, there may be a function theory about biological items *having* functions that counts as an ontological counterpart to the 'epistemic' b-ICE-theory. Such ontological counterpart theories are given in (Vermaas 2009b) for the ICE-theory for technical functions and generalisations thereof. However, in those counterpart theories, items — artefacts and biological items — still have their functions relative to beliefs of agents. Secondly, the b-ICE-theory is not merely about beliefs of agents; as the ICE-theory itself, the b-ICE-theory requires that function ascriptions by discoverers and justifiers are supported by an account A. This typically consists of biological and other scientific knowledge. Therefore, the b-ICE-theory, as its technical counterpart, satisfies the support desideratum.

Secondly, the b-ICE-theory allows function ascriptions to organisms as a whole, whereas other function theories in biology typically do not. Whenever one can find a goal-directed pattern of behaviour of an organism in which a capacity of the organism as a whole plays a contributing role, then that capacity can be ascribed as a function. For example, a constrictor snake may be ascribed the function of strangling prey, relative to its behavioural pattern of predating.

The b-ICE-theory also has advantages. First, it satisfies the malfunctioning desideratum for function ascriptions in biology, just like the ICE-theory does. Second, it can be made to satisfy the proper-accidental desideratum for biological items by distinguishing various behavioural patterns relative to which functions are ascribed. Proper functions — say the capacity to pump in the case of hearts — are ascribed relative to the behavioural patterns of one type, and accidental functions — the capacity of those hearts to make a thumping noise — relative to patterns of another type. One way to identify the relevant types might be through the contribution of behaviour patterns to the survival and/or fitness of the organism involved, but this is just a guess.

Third, in the b-ICE theory, function ascriptions do not depend on one particular biological theory. Function ascriptions by biologists who employ(ed) different theories than current neo-Darwinian ones, for instance, can also be accommodated by the b-ICE-theory.

Fourth, the theory is neutral with respect to the issue of teleology within biology and shows how goals enter function ascriptions. The b-ICE-theory is compatible with the position endorsed by, e.g., Searle, that '[t]here is no intrinsic purpose whatever to the origin and survival of biological species'.[19] And it is compatible with the position that nature via, e.g., the process of evolution, reveals a certain directedness of organisms. In both cases, the discoverers incorporate or identify, respectively, the goals in function ascriptions by isolating a certain set of behaviours of the organism as a goal-directed behavioural pattern.

We leave a further assessment of this b-ICE-theory to our colleagues in the philosophy of biology. But if this theory or a similar one is acceptable for biology, then our little exercise shows that the ICE-theory can be employed to defend the position that biological function ascriptions should be understood *as if* they are function ascriptions to artefacts. The presupposition of this position need not be that biological items are designed in our action-theoretical sense, but merely that the epistemic background of function ascriptions to biological items is similar to the action-theoretical background of function ascriptions to artefacts. By endorsing this alternative presupposition, one keeps the intentional element of biological function ascriptions that Dennett argues for, but avoids the consequence that biological items are designed by agents to be used by other agents.

A generalised ICE-function theory

A more general remark is that the b-ICE-function theory may, in turn, be transposed to disciplines other than biology. It may be argued that if, in the biological domain, functions can be ascribed relative to an epistemic background that consists of scientists who communicate their discoveries to one another, then such function ascriptions are also possible in other sciences. For behavioural and social sciences, such as sociology and psychology, this may be good news; for physics and chemistry, it is most certainly not. So, if one accepts this argument, one needs some constraints to prevent that discovered patterns in physics and chemistry lead to function ascriptions. An obvious candidate for such a constraint is that one should not take these physical and chemical patterns as goal-directed.

The question whether the ICE-theory itself applies to the biological domain, and whether it can be generalised to a theory that applies to this domain, is connected to the question whether a uniform concept of function is used across the different domains that involve functional descriptions. If there is, one expects function theories to apply in all those domains. We are agnostic about this issue. In fact, it seems possible to argue for both a negative and a positive answer on the basis of our results. On the one hand, we take the definition of function ascriptions relative to use plans, which is part of the ICE-theory, as tenable in the technical domain and untenable in the biological domain. In this sense, we accept inter-domain pluralism: ICE-function ascriptions relative

[19] Searle (1995, p. 16).

to use plans are particular to the technical domain and, in this respect, there is no uniform concept of function. On the other hand, we could argue that function ascriptions relative to use plans can be transposed to b-ICE-theory function ascriptions relative to goal-oriented patterns of behaviour. It then seems possible to define an overarching g-ICE-function theory, which has the ICE-theory and the b-ICE-theory as special instances. Function ascriptions as defined by this overarching theory are then uniform, if one at least acknowledges that the concept of function ascriptions relative to behavioural patterns make sense in the biological domain.

Chapter 7

The nature of artefacts

In this book we have presented our philosophy of artefacts. We started from an action-theoretical perspective, clarifying the intuitive connection between artefacts as useful objects and the goals, beliefs and actions of agents. After analysing artefact using and designing, we focussed on the characteristic of artefacts that has drawn most philosophical attention: their functions. We criticised existing function theories on the basis of our desiderata and we presented our alternative. This ICE-function theory analysed how functions are justifiably ascribed to artefacts, and connected function ascriptions to the goals and actions for which objects are employed, and to the physicochemical structure of these objects. Then, we argued that our account meets the desiderata and finally explored how our ICE-theory applies to other domains than the technical.

Instead of focussing immediately on artefacts, we have constructed accounts of acting with artefacts and a theory of function ascriptions. In this chapter, we take stock and present some consequences for a metaphysics of artefacts.

We start by arguing, in section 7.1, that technical functions are conceptual drawbridges: functional descriptions allow users and engineers to connect and *disconnect intentional and structural descriptions of artefacts. This analysis of the conceptual or epistemic role of functional description of artefacts challenges the metaphysical position that functions can be taken as the essences of artefacts, as we show in section 7.2; our account suggests rather that if artefacts have essences, it is that they are objects embedded in use plans. We explore this alternative position in section 7.3 and find that it has several shortcomings. Finally, we argue that it may be supplemented with another common intuition about artefacts, namely that they are man-made objects. In section 7.4, we sketch how these two perspectives on artefacts may be combined profitably.*

W. Houkes, P.E. Vermaas, *Technical Functions*, Philosophy of Engineering and Technology 1, DOI 10.1007/978-90-481-3900-2_7,

7.1 Functions as conceptual drawbridges

In this book, we have constructed a theory of artefact functions in two steps. Firstly, in chapter 2, we gave an action-theoretical analysis of using and designing, casting use plans in the leading role. Secondly, in chapter 4, we formulated conditions under which agents may justifiably ascribe functions to artefacts. These conditions make explicit the connection between functions, goals and actions by embedding function ascriptions in use plans.

These results raise questions regarding the importance of functions in the characterisation of artefacts. As we stated at the outset, artefacts are frequently described in functional terms, or even named after their function. In our own examples, we have used toasters, televisions and painkillers. Some philosophers have taken these functional characterisations so seriously that they regard functions as the essences of artefacts — meaning that, for artefacts, the correct answer to the question 'What is that thing?' necessarily involves a functional description.

We examine this function essentialism in the next section. First, as a preliminary, we turn our attention to the conceptual role of functions. For, no matter how seriously one takes technical functions, it cannot be denied that artefacts are not *always* characterised in functional terms. Moreover, although agents use functional characterisations such as 'toothbrush', users in particular seldom use the notion of function to describe their dealings with artefacts.

These considerations from ordinary language cohere with some of our results. Function ascriptions involve beliefs about the effectiveness of a use plan, and about capacities of items that contribute to this effectiveness. Thus, functions connect two descriptions of artefacts that have been called their 'dual nature'.[1] On the one hand, artefacts can be described structurally, in terms of their physicochemical capacities; on the other hand, they can be described teleologically in terms of goals and actions. Only in combination, the structural and teleological (or physical and intentional) descriptions provide a complete characterisation of artefacts.[2] On our analysis, function ascriptions also feature elements of both descriptions and show one relation between the physical and intentional characterisations of artefacts. Thus, the notion of 'function' is like a bridge connecting the intentional, use-plan description of artefacts and a description of their physicochemical capacities.[3]

Perhaps surprisingly, this bridging role of functions can be used to explain why artefacts are often *not* characterised in functional terms. Moreover, our analysis shows how artefacts may be described exclusively in either teleological or structural terms — without undermining the 'dual-nature' thesis that both are needed for a complete characterisation. We now describe the process in which users and designers selectively (de-)emphasise intentional and structural descriptions of artefacts. Continuing the earlier metaphor would liken functions to drawbridges of a curious type, which can be hauled up from either side. To

[1] Kroes and Meijers (2000, 2002, 2006).

[2] Kroes and Meijers (2006, p. 2).

[3] Vermaas and Houkes (2006a).

avoid overstretching the imagination,[4] we switch metaphors: we describe the emphasising process in terms of 'cloaking' and 'highlighting' elements of the full characterisation.

We first focus on the ways in which users may describe artefacts. By our general description of using, users need not describe functions to these artefacts at all. They may just describe artefacts as means to realise their goals: a towel is for drying your hair, a toaster is for making toast. If users *do* ascribe a function to an artefact, in line with our ICE-theory, this involves references to both the use plans and the physics of the artefact. The latter reference may, however, be coarse-grained and cursory. In ascribing the function to ϕ to an artefact, a user merely needs to believe that the artefact involved has the physical capacity to ϕ. In justifying this, the user may rely on designer testimony, and thus avoid extensive structural descriptions of artefacts. Hence, although function ascriptions by users refer to physical capacities of artefacts and to use plans for those artefacts, the conditions for these function ascriptions make clear how a physical description of the artefacts is de-emphasised in favour of a description in terms of goals and actions alone. More specifically, one physicochemical capacity is *highlighted*, while the rest of the physical structure is *cloaked*, often literally in the phenomenon of black-boxing.

Let us turn to engineers. We showed in section 6.2 how our theory of function ascriptions accommodates the self-understanding of engineers, in which engineering is primarily regarded as a physicochemical enterprise. We presented two ways in which references to use plans can be minimised.

The first way considers function ascriptions to components of artefacts. Like artefacts, components have use plans, developed by component designers and executed by users, who are typically themselves designers of artefacts. Unlike ordinary use plans, use plans for components are easily bracketed. They can be taken as claims about how capacities to ϕ of components 'add up' structurally to capacities to Φ of configurations of these components. Engineers can justify their beliefs about structural 'addition' on the basis of their scientific and technological knowledge. Hence, the action-theoretical background for function ascriptions to components is exactly that: a background. Thus, if designers ascribe to a component a capacity to ϕ, relative to a capacity to Φ of a configuration of components, they *highlight* a capacity of the component that contributes to the function of the artefact and *cloak* the use plans of both the component and the artefact — reversing the cloaking-highlighting mechanism that is typical for users.

The second way concerns function ascriptions to artefacts as a whole. Here, our results are less univocal. For artefacts with well-established or indeterminate use plans, engineers may again bracket references to the actions and the goals making up the use plans of the artefacts. Then, they ascribe functions to these artefacts on the basis of their physicochemical make-up. For other artefacts, this result does not hold. When ascribing functions to ϕ to artefacts,

[4]The product of our conceptual engineering is highly innovative: a two-sided drawbridge outfitted with spotlights and cloaking devices!

engineers need in general to refer also to the user actions and goals in order to justify their beliefs that the use plans are effective. This leads to a more complicated cloaking-highlighting mechanism. If designers ascribe a function to ϕ to an artefact, they — just like users — highlight the capacity to ϕ of the artefact that is relevant to its use plan. They cannot, however, always cloak the action-theoretical conditions for function ascriptions. In justifying the beliefs that underlie the function ascription by means of scientific and technological knowledge, designers have to consider the goal and the actions that are part of the use plan. Returning to the methodologists' picture of engineering design as divided into an intentional part and a structural part (see section 6.1): in designing, engineers may focus on the structural part in which they determine what the physicochemical structure of an artefact should be like if it has to have a specific function. Yet in order to determine that function, engineers need to consider the users and their goals in the intentional part.

Results in cognitive science partly confirm this analysis, in particular the highlighting mechanism for users.[5] Experiments have been done to find out when people categorise artefacts in terms of their function, and when they refrain from doing so. The leading question is whether functions are so-called core properties, i.e., properties that are necessary and sufficient for determining whether or not an item belongs in a category.[6] It turns out that functions are used as the basis of artefact categories, but only in circumstances in which people could reasonably infer that the physical structure of the artefact would support effective use. Barton and Kamatsu (1989), for instance, asked subjects to categorise items on the basis of rudimentary physical information. A sample question was whether items made of different materials, which both reflect images, are both mirrors. Barton and Kamatsu conclude from the subjects' positive response to this question that functions are core properties. On the ICE-theory, these function ascriptions would indeed be justifiable because the subjects have effectiveness and (rudimentary) physical-support beliefs, both based on testimony. When, however, effective use is not obviously supported by the physical structure, for instance if an artefact has a highly non-standard composition, people typically refrain from using functional categories and describe the artefact in terms of either its physical structure or very indeterminate goals. Malt and Johnson (1992), for instance, offered subjects descriptions of items with different combinations of function ascriptions and information about the physical structure, such as combinations of 'provides warmth to the upper body when worn over a shirt' with the physical description of either a sweater or a rubber garment. Subjects typically described the first item as a sweater and the latter item as 'a rubber thing' or 'some rubber garment'.

Even two-year olds are reported to describe artefacts in functional terms, if

[5] The presentation of these results is largely based on Romano (2009).

[6] The existence of core properties is a matter of debate between proponents of various theories about the nature of categories and concepts; see, for instance, Smith and Medin (1981) for an overview of exemplar, prototype and theory-theories. Still, that the empirical results which are significant for our analysis were gained by proponents of a controversial theory does not make the results any less significant.

the relation between use and physical structure is sufficiently clear. Kemler-Nelson, Russell, Duke, and Jones (2000) showed a simple object to children and described it in functional terms. They did not demonstrate the use of the object, but allowed children to manipulate it. Children then turned out to describe other objects in similar functional terms, provided that they could comprehend how the function was supported by the physical structure. In earlier experiments, such as those reported by Gentner (1978) and Landau, Smith, and Jones (1998), children were not allowed to manipulate the object and no effort was made to communicate or facilitate a physical-support belief. In these circumstances, only children from the age of five upwards, who have much more elaborate beliefs about the physical behaviour of objects than two-year olds, were found to ascribe functions. These results suggest that even very young children ground function ascriptions on beliefs about the relation between effective use and the physical structure supporting it, and that they refrain from using functional terms when lacking information about (the relation of) use and structure.

7.2 Against function essentialism

In the previous section we have discussed the ways in which artefacts are described in contexts of using and designing. The picture that emerges is that of a network of structural, teleological and intentional notions, all used to describe artefacts. Functional notions are useful go-betweens in this conceptual network. However, we argued that, in this role, functions are dispensable: the cloaking and highlighting of parts of the network are as important for understanding how users and designers describe artefacts as function ascriptions.

Given these different ways in which notions from this network can be employed, limiting a philosophy of artefacts to a theory of function ascriptions and functional descriptions would be short-sighted. But we can draw a stronger conclusion on the basis of the previous chapters, namely that functions *cannot* be understood as the (real or nominal) essences of artefacts. Hence, no matter how far one is willing to go in reconstructing ordinary language, it is *impossible* to construct a framework in which function is the most fundamental notion for describing artefacts.

This conclusion may not seem particularly interesting from the perspective of conceptual analysis. If we shift attention to the metaphysics of artefacts, however, our call for de-emphasising functions is in discord with a small chorus of philosophers who might be called ‘function essentialists’. In this section, we present an argument against such function essentialism — an argument that involves little more than the translation of some earlier conclusions in metaphysical terms — and we give one example of the problems that a function essentialist metaphysics, viz. Lynne Rudder Baker’s (2000, 2004) constitution view, encounters with regard to the phenomenology of artefact use.[7]

[7]Houkes and Vermaas (2004) and Houkes and Meijers (2006).

In metaphysics, artefacts have not suffered from complete neglect. Both Aristotle and Heidegger can be interpreted as providing ontological views on artefacts, and more recently several well-known contributors to general metaphysics, such as Van Inwagen (1990), Wiggins (2001) and Baker (2004), have commented on the ontological respectability of artefacts. Moreover, some hotly debated topics in metaphysics use artefacts as prime examples, if not focussing explicitly on them. Both Theseus' ship and Goliath (and/or Lumpl) are artefacts as well as metaphysical conundrums.[8]

Here, we simply assume, following Baker (2004) and in line with the general thrust of this book, that artefacts are ontologically respectable. We will, moreover, not describe and criticise the views alluded to above. Instead, we argue against a supposition that many ontological accounts of artefacts have in common, despite their other, marked differences. This supposition is that functions may be regarded as the essence of artefacts.[9]

Our claim is that accommodating the phenomenology of artefact using and designing in a theory of artefacts seriously compromises a commitment to function essentialism. This phenomenological argument may be developed by briefly summing up some earlier parts of this book — sections 2.7 and 5.3 — that dealt with classifications of use as proper or rational.

Throughout this book, we have argued that these salient features of everyday artefact use can be analysed by directly considering using and designing, and describing such activities in terms of the use plans executed and constructed in them. This action-oriented approach accounts for various evaluations of using by appealing to both standards for use plans and the social role of certain designers. In this way, we made a conceptual distinction between rational and proper use, and also showed how these concepts can be related.

We will not repeat the many respects in which this use-plan account is adequate to the phenomenology here. Our present point is that, in this respect, it outperforms accounts that take technical functions as the most fundamental notion. Suppose, for instance, that such an account only distinguishes proper functions and accidental features; many function theories do, as we have seen in chapter 3. On this categorisation, the colour of the casing of a vacuum cleaner, its making noise when switched on, and its casting a shadow in daylight are all accidental features of this artefact, although some might be related to its proper

[8]Various recent contributions to what appears to be a very lively debate in general metaphysics are collected in Rea (1997).

[9]Sample expressions of function essentialism are: 'Artefacts are collected up [...] under functional descriptions that are precisely indifferent to specific constitution and particular mode of interaction with the environment. A clock is any time-keeping device, a pen is any rigid ink-applying writing implement, and so on.' (Wiggins 2001, p. 87); 'What makes something a clock is its function of telling time, no matter what it is made of' (Baker 2004, p. 100), '[...] what makes two artifacts members of the same kind is that they perform the same function' (Kornblith 1980, p. 112). These references show one of the differences that function essentialism cuts across, namely between those who think that artefacts have only nominal essences (Wiggins, Schwartz (1978)), and those who believe that artefacts have real essences or 'underlying natures' (Putnam (1975, p. 162), Kornblith, Baker). A position in which the proper functions of some technical artefacts are part of their essences is defended by Elder (2004, 2007).

function to clean floors of dust. This may, to some extent, capture the notion of proper use, if the account is supplemented with an appropriate mechanism for privileging one supposed capacity. It runs into troubles, however, with regard to alternative use. For there is an intuitive distinction between features that are used, albeit improperly, and features that are not used for practical purposes at all.

Hence, the function essentialist needs at least a ternary typology, which distinguishes capacities from functions and, on top of that, distinguishes some functions as the essential, proper ones. This may be a natural move to make, but it does not lead to phenomenological adequacy. For it seems that, in such an account, all evaluative work has to be done by the notion of proper function. It seems that at least the claim to properness can be captured in functional terms, characterising proper use as use that is 'in accordance with' the proper function of an artefact. But no matter what the merits of this metaphorical characterisation may be, it leaves unanalysed the notion of rational use, which we took as the primary evaluative notion in our phenomenology. Our account has available the standards for use plans and the interaction between using and designing, but these resources obviously address actions rather than functions.

The function essentialist may claim that the relevant standards and the analysis of professional designing, redesigning and other activities can be translated into 'function talk'. But at that point a pattern emerges. On the one hand, it is certainly possible to describe the phenomenology by consistently emphasising functions at the price of actions, thus saving a previous commitment to function essentialism. On the other hand, it seems hopeless to establish this view through such a systematic reconstruction. Rather, the repeated translation of action-theoretical vocabulary and analyses into function talk strongly suggests that the former is more fundamental than the latter. The only way to take away the suspicion that an adequate theory of functions is parasitic on an analysis of artefact using and designing would be to establish the commitment to function essentialism on grounds that are either different from the phenomenology of using and designing that we considered in previous chapters or independent from phenomenology altogether. We showed in chapter 3 that existing theories of function are based on such phenomena as malfunctioning and the distinction between proper and improper use. Hence, the phenomenological basis of these theories is a part — and only a small part — of that of our analysis of using and designing, which makes the former option highly unlikely. This leaves the latter option, to establish function essentialism without considering phenomenology. Given our own phenomenological starting-point, this route seems unattractive: it would divorce an analysis of artefacts from the involvements that intentional agents have with them, which gave rise to the existence and descriptions of these entities in the first place.[10]

[10]One possible source of intuitions to which the function essentialist may appeal seems independent from the phenomenology of artefact use and designing, namely intuitions concerning the persistence of artefacts in various circumstances. Indeed, authors such as Baker and Wiggins often use such considerations, arguing for instance that statues obviously cease to exist when smashed to pieces; see, for instance, Baker (2000, pp. 37–38) for intuitions re-

The function essentialist's continual need to accommodate the phenomenology by translating action-theoretical notions and accounts is not surprising, given our own theory of function ascriptions. This theory takes functions, both proper and possible, not as simple properties of an artefact but as capacities that an agent ascribes to an artefact relative to a (purportedly rational) use plan for that artefact and relative to a body of evidence. If this is indeed the only adequate theory of technical functions on offer, it follows immediately that 'function' is not the most fundamental notion for describing artefacts. Since our theory describes function ascriptions in action-theoretical, epistemological and structural terms, it implies that it may be possible to translate every action-theoretical account into functional terms, but that it is necessary, to translate functional terms to action-theoretical ones. Thus, our analysis of function ascriptions undermines function essentialism, dispelling a myth[11] and thereby, ironically, lowering its own status. Apparently, technical functions look better when observed casually and from a distance.

Let us give one example of the problems that a function-essentialist metaphysics of artefacts runs into on a close encounter with the phenomenology of artefact using. This example concerns the constitution view, proposed by Lynne Rudder Baker (2000) as a general metaphysical theory and applied, in a recent paper, (Baker 2004), and book, (Baker 2007, ch. 3), to artefacts. Although it is just one possible way of developing a constitution view of artefacts, we take it as our reference point here. We first summarise Baker's metaphysics of artefacts and then derive a problem from its explicit commitment to function essentialism.

According to Baker, constitution is the glue of the material world. The concept is meant to describe virtually all metaphysical relations between middle-sized or everyday objects and material aggregates. Constitution is conceived as a relation between things of fundamentally different kinds; it is not a watered-down identity relation. Reality must be fragmented before it can be glued together. Baker's way of doing this is by determining an object's *primary kind*: 'Every object has its primary kind essentially, and entities of different primary kinds have different persistence conditions. Constitution is a relation between

garding the persistence of a dismantled manor house, a carburetor and Carracci's *Landscape with the Flight into Egypt*. We believe that intuitions regarding artefact persistence may ultimately be based on intuitions regarding the possibilities of repairing or re-assembling them. A watch ceases to exist when smashed to pieces, but continues to exist when taken apart by a careful and competent watchmaker. Hence, even these intuitions spring from our practical involvement with artefacts rather than from some deeper, more pristine source.

[11] Function essentialism is not an exclusively philosophical myth. In cognitive science, it is upheld by Bloom (1996, 1998), who analyses technical functions primarily in terms of author intentions rather than of actual or potential usefulness. In support of this view, Matan and Carey (2001) report an experiment in which subjects were asked to categorise artefacts on the basis of information about their intended use and about alternative ways of using the artefacts; typically, subjects preferred categorisations based on intended rather than alternative use. However, rather than supporting function essentialism, this shows that subjects categorise artefacts on the basis of information about the *use* of artefacts: even if they use functional terms, judgements about proper and alternative use are ultimately at stake and functions are ascribed on this basis. This result therefore underwrites our phenomenological argument against function essentialism rather than confirming the latter.

things of different primary kinds.'[12] Hence, the key to determining an object's primary kind, i.e., its fundamental ontological type, is to study its persistence conditions.[13] Besides this notion of primary kind, the constitution view employs that of *circumstances*: a thing of a given primary kind does not constitute another thing of another primary kind *per se*, but only in certain circumstances. In this way, water molecules in a particular arrangement constitute a river and pigments on a canvas, presented at an exhibition, constitute a painting.[14]

Baker has applied this general view to artefacts. She characterises the material basis of an artefact as an aggregate, meaning that its identity conditions are exhausted by the identity conditions for its parts and that its primary kind is a hybrid or, more accurately, a list of the various primary kinds of its parts. The material basis of a hammer, for instance, is an aggregate of wood and steel that would be different if either part was damaged or replaced and that cannot be described more compactly than in terms of wood and steel. The artefact itself, however, has a different and more unitary primary kind, namely its (proper) function, 'determined by the intentions of its designer and/or producer'.[15] Summarising Baker's own formulation, an object x is an artefact if and only if: (Art.1) it has one or more designers d — presumably, product designers in our typology — whose intentions in part determine its proper function and on whose intentions the existence of x depends; (Art.2) x is constituted by an appropriate aggregate, selected or arranged by d in order to realise x's proper function.

This analysis of the nature of the artefact x is supplemented with two conditions on the aggregate a that constitutes x and with a characterisation of the circumstances in which a constitutes x. An aggregate a is appropriate if and only if: (Agg.1) a contains enough items of suitable structure to enable the proper function of x to be performed and (Agg.2) the items in a are available for assembly in a way suitable for enabling the proper function of x to be performed.[16] The circumstances in which a constitutes x cannot be specified

[12] Baker (2004, p. 100).

[13] These need not be formulated in structural or geometrical terms: Baker allows and even recommends the identification of objects in terms of their causal capacities and functions. Constitution is meant to relate things with different causal powers; if it would not, those who think that entities are told apart by their causal capacities would conclude that constitution is an identity relation, which Baker emphatically denies. The causal powers ascribed to constituted objects should, by the way, be conceived rather broadly: Baker, for instance, repeatedly uses the capacity of a flag to make a veteran cry as an example.

[14] Another ingredient of the constitution view that is less important for our purposes, is a distinction between two ways of having properties, namely 'essentially' or 'derivatively'. Intuitively, some properties are possessed by both the higher-order object and its material basis: if, for instance, the mechanism in an ordinary clock produces a loud ticking noise, the clock does. On the constitution view, the mechanism has this property essentially; it could not cease to make the noise without ceasing to be the mechanism that it is. The clock has this property only derivatively; repairing it might stop it ticking without ending its existence altogether. Again, an appeal to intuitions regarding the persistence conditions of objects is made to settle this issue.

[15] Baker (2004, p. 102).

[16] The appeal to 'suitability' means that this constitution view of artefacts has problems dealing with malfunctioning statements. In cases that are not analysed as caused by 'human

precisely and exhaustively, but comprise, among other things, that a is in the presence of knowledgeable persons who deliberately and successfully set out to create x by manipulating a.

The central definition of the constitution view of artefacts then reads: a constitutes x at time t if and only if a and x are spatially coincident at t in x-favourable circumstances. To the conditions on aggregate a, artefact x and x-favourable circumstances listed above, Baker adds two modal conditions that underline the importance of the circumstances and that introduce an asymmetry between a and x: any different aggregate in x-favourable circumstances necessarily constitutes an artefact of the same functional type as x (say, a larger hammer); and a may not constitute a hammer, namely in hammer-unfavourable circumstances (say, when found on a hitherto unexplored asteroid).

This function-essentialist application of the constitution view leads to a problem in cases where, intuitively speaking, artefacts acquire secondary or superseding proper functions. Take, for instance, again Aspirin. The active ingredient in Aspirin tablets, acetylsalicylic acid, has, as we often mentioned, recently attracted the attention of pharmacologists because of its beneficial properties other than alleviating pain and reducing fever. Many of these supposed effects, such as the prevention of various forms of cancer, are still under review, but it has been established that Aspirin works to prevent the formation of blood clots. Indeed, the drug is now being prescribed to cardiovascular patients, who regularly need to take tablets that contain a lower dosage of acetylsalicylic acid. Suppose that someone who takes Aspirin as a prescription drug visits a friend, and suddenly remembers that he has forgotten to take his medicine today. Instead of going back home, he asks his friend whether she has any Aspirin. She takes a package of tablets, which she has bought a short time before to use as painkillers but has not used so far, from her medicine cabinet and gives her guest a tablet. Since the package contains Aspirin with a high dosage, it would seem that the tablets in it constitute painkillers, i.e., artefacts with the proper function of alleviating some ache. But since the guest takes it as a blood thinner, *in lieu* of his prescription drug, it also seems that this particular tablet is a blood thinner. So what happened to the painkiller that was in the unopened package? Did it cease to exist during the process of being used as a blood thinner? If so, at what point? Or did the tablet acquire a new proper function at some time

failure', artefacts appear to malfunction because the aggregate that constitutes them is *not* suitable for performing the artefact's function. Planes crash and pens run out of ink. Simply claiming that a suitable aggregate constitutes the artefact is unattractive, but the only alternative is to say that the unsuitable aggregate did not constitute an artefact at all, which runs in the face of claims by engineers and potential users. It may be plausible to modify the constitution view, taking into account the normative content of malfunctioning claims as presented in section 5.3. One could let the conditions that determine whether an aggregate is appropriate or not refer to the beliefs of intentional agents, either designers or users and not to *de facto* properties of the aggregate in question. The original constitution view presented in Baker (2000) offers this possibility, whereas her later application to artefacts, discussed here, does not. The original view, however, has the disadvantage that the relation between the higher-order object and its material basis depends largely, or even exclusively, on the circumstances: there are no conditions of appropriateness on the aggregate, which seems implausibly liberal.

during this process, constituting both a painkiller and a blood thinner? Do all tablets in the container constitute two artefacts?

One might respond that the dosage of acetylsalicylic acid makes it the case that the aggregate is suitable as both a painkiller and a blood thinner, but that its height makes it more appropriate for the former role than for the latter; therefore, the pharmacologists designing this drug have chosen this particular aggregate instead of one with a lower dosage, which they deemed more appropriate as a blood thinner. The symmetry between the two functions is easily restored, however, if we suppose that, contrary to fact, painkiller-Aspirin and bloodthinner-Aspirin would contain the exact same dosage of acetylsalicylic acid, or if we imagine that a new prescription drug for preventing colon cancer contains the exact same dosage of acetylsalicylic acid as original painkiller-Aspirin. If our example is re-enacted in this possible world, claiming that it amounts to non-standard or even improper use of painkiller-Aspirin is counterintuitive.

Once it is admitted, in any of these scenarios, that all Aspirin tablets have two proper functions, Baker's constitution view of artefacts entails that the tablets constitute two functional objects: one aggregate with a specific dosage of acetylsalicylic acid spatially coincides with, but is not identical to, an object with the primary kind of killing pain and an object with the primary kind of preventing blood clots (or, alternatively, colon cancer). We think that there is something distinctly unattractive in this type of 'ontological stacking', in which two distinct functional objects are spatially coincident; for one thing, the relation between the two artefacts remains unclear. Nor is there, in this particular case, anything sacred about the number three; multifunctional material objects may constitute many, perhaps indefinitely many distinct artefacts, all of which would have a different functional essence.[17]

Baker has admitted this much. In a commentary on an earlier, slightly different version of this argument,[18] she responded to the puzzle that multifunctional objects create for her constitution view by suggesting that artefacts have disjunctive essences (Baker 2006). This would mean that, in the example above, all Aspirin tablets constitute one functional object, which has the primary kind of either killing pain or preventing blood clots.

Although this would give artefacts a unique ontological status, it damages their metaphysical reputation. If an essential feature is, following Aristotle and almost all of Western metaphysics, including Baker herself, the answer to the question 'What is it most fundamentally?', a disjunctive essence is an oxymoron. The absolutism inherent to this way of doing metaphysics leaves no room for doubt. Artefacts would therefore have no essences: functions do

[17] Stratification does not resolve this issue. Suppose one claims that, in the Aspirin example, the blood thinner handed to the ailing guest constitutes a painkiller and is itself constituted by a tablet containing acetylsalicylic acid. However neat, this ontological organisation leads to counterintuitive results. The relation of constitution is, according to Baker (2000, p. 45), not transitive. In the current example, the tablet would then constitute a blood thinner, and the blood thinner a painkiller, but the tablet would not constitute a painkiller.

[18] Houkes and Meijers (2006).

not qualify for reasons of multiple-personality-disorder, and the commitment to function essentialism means that no other candidates are on offer. Hence, function essentialism may lead naturally to the 'artefact nihilism' that is often associated with Van Inwagen.[19]

7.3 Plan relativism

In the previous section we criticised philosophies of artefacts that elevate functions to essential status, arguing that they make artefacts a metaphysically dodgy class of objects.

This raises the question 'What are artefacts, if they are not primarily functional objects?' In this section, we sketch a metaphysics of artefacts that takes seriously the use-plan approach taken in this book: it defines artefacts as objects that are embedded in use plans, i.e., as objects that are useful, elevating the action-theoretical vocabulary of chapter 2 to essential status. We study some consequences of this definition, before supplementing it with another intuition regarding artefacts in the final section of this book.

So let us attempt to derive a definition of artefacts from the analyses in the previous chapters. There, we analysed notions such as 'use', 'design' and 'function', and we did so on the basis of a primitive notion of 'artefact': use plans, for instance, were distinguished from plans in general because they include the manipulation of at least one material object other than our own body. In doing this, we left it open to which objects our central notions apply. Alternatively, however, one could take 'artefact' to be defined implicitly by our characterisations of using and designing. Making this explicit leads to:

> **A 'useful materials' definition of artefacts:**
>
> An object x is an artefact a if and only if x is manipulated in the course of executing a specific use plan p, which is designed, communicated and evaluated in accordance with the use-plan analysis of using and designing.

Hence, fresh, running water (x) is a cooling agent (a) in the context of generating nuclear electricity (p), and it is a cleaning agent (a') in the context of washing one's hair (p'); a piece of steel and plastic is a screwdriver when building a garden shed, and an opener when opening soda bottles; a complicated configuration of various materials is an airplane when flying across the Atlantic, and a museum piece once it has gone out of service and is on exhibit. On this reading, it is no longer a question whether our theory applies to other items than artefacts: everything to which it applies is an artefact by definition. The theory can thus at most be used to distinguish artefacts, as a category of objects, from other types of objects, namely those that are not embedded in use plans.

[19] Van Inwagen (1990, p. 111) maintains that references to artefacts should be reformulated as references to 'artefact-wise' arrangements of simples.

One way of developing this definition into a 'plan-based metaphysics of artefacts', is to cast it in terms of the constitution view, which was discussed in the previous section. As we argued, Baker's application of this view to artefacts founders on a commitment to function essentialism, but the definition just given suggests an alternative application, although that does require some reshuffling of variables. An aggregate could be taken to constitute an artefact in circumstances that include justified beliefs about the efficacy of a use plan and about the physical support of this plan by capacities of the aggregate. Then, to rephrase the definition above in terms of Baker's constitution view, an aggregate x constitutes an artefact a with respect to a use plan p if and only if: (Art.1) p includes the manipulation of x as one of the steps towards realising the goal state g; (Art.2) the designers d of p selected, arranged or produced x as part of their intention to contribute to the realisation of g, and they communicated p.

This way of setting up a metaphysics of artefacts brings out the plan relativism of our suggestion. Aspirin tablets (x) would then constitute painkillers (a) in the context of the older use plan (p) that serves to alleviate pain (g) by swallowing a specific dosage of Aspirin (manipulating x); the same aggregates (x) constitute blood thinners (a') in the context of another, more recently designed and communicated plan (p') that serves to prevent blood clots (g') by swallowing a different dosage of Aspirin (again, manipulating x, which is now a different action). In this view, a tablet could constitute one artefact, i.e., be a painkiller, until it is manipulated in the course of executing another use plan, i.e., when the tablet is taken from its package by a cardiovascular patient taking his daily dosage: the painkiller then ceases to exist, and a blood thinner comes into existence.

Formulated in these terms, our definition offers a framework for discussing the persistence conditions for artefacts. These conditions are related to the use plans for these artefacts and to the aggregates that are manipulated in these plans; thus, they have one foot in the intentional, and one in the physical realm, so to speak.[20] Where one artefact ends and another begins is of less interest; in fact, our definition suggests that drawing such boundaries may be as context-sensitive as determining the differences between plans. One might, for example, be inclined to say that an artefact a ceases to exist as soon as it becomes impossible (or at least generally irrational) to execute the use plan p with respect to which the aggregate x constituted a. A car that is wrecked beyond repair in a crash is then no longer a car, but a mere aggregate of twisted steel and plastic. Similarly, when the white, fresh-smelling aggregate that comes in tubes is used to fill a small crack in a wall, it ceases to be toothpaste — assuming that no-one in his right mind would then still use it as such — and becomes filling material instead. In both cases, the persistence of the artefacts

[20] Thomasson (1999), placing herself in a tradition that goes back at least to Husserl's *Logical Investigations*, has proposed to take an analysis of the dependence relation as a general metaphysical framework. The considerations in this section and the next develop some of the specifics of the dependence relation for artefacts, in a way that appears to be compatible with Thomasson's general proposal and her analysis of artefacts as depending on both mental states and material objects (see, more recently, her (2007a, 2007b)).

is intimately connected to realising goals by means of material objects. This makes the metaphysics of artefacts a partly practical concern.

This indicates how a plan-centred metaphysics of artefacts may be developed. In the remainder of this section, we explore some of its consequences. All of these consequences might individually be taken as showing the absurdity of this metaphysics; however, we shall postpone judgements as long as possible, i.e., until the end of this section.

A first consequence of the above useful-materials definition of artefacts is that it becomes impossible to make a principled distinction between artefacts and natural objects. At least since Aristotle, the distinction between nature and technology, *physis* and *techne*, natural objects and artificial objects has been a staple of all philosophical thinking about artefacts. Moreover, the distinction seems to express certain deep-rooted, albeit culture-relative values: calling something 'natural' or 'artificial' does not, or not just, express some matter of fact, but may also act to recommend or discredit the item in question. Nowadays, at least in the Netherlands, 'artificial' often carries negative connotations: it counts as a recommendation if foodstuffs only contain natural flavourings, clothing is advertised as composed of all-natural materials, and the prospect of 'artificial intelligence' fills many people with dread. Thirty years ago, these negative connotations and the corresponding nightmare scenarios of dominating robots and self-replicating machines were less outspoken or even absent altogether — as shown by even a sketchy knowledge of science-fiction movies and novels.

A plan-centred metaphysics of artefacts has difficulty accommodating these intuitions. As soon as there is a designed, communicated and justifiable use plan in which an object is manipulated, that object is by definition an artefact. This does not just make it impossible to draw a fixed line between natural and artificial objects, but it actually makes it difficult to draw the line anywhere at all. Manipulating a forest by clearing a path through it, or intentionally using mountains as part of telecommunication systems converts these objects into artefacts. Even Jupiter is an artefact, since it is intentionally manipulated for accelerating an interplanetary probe towards the outer regions of our planetary system. Thus, plan-relativism regarding artefacts leads to a dilemma: either we regard virtually all objects as artefacts, leaving no place for natural objects[21] or we abolish the distinction between natural and artificial objects and try to devise other categories (e.g., 'improvised' or 'permanent'; 'goal-enabling' or 'goal-contributing'; 'self-replicating' or 'produced') to accommodate our intuitions. Since the latter seems possible, although not within the scope of this book, this dilemma does not amount to a fatal objection against a plan-centred view.

[21] This pan-artifactualism is not without precedent. In *The Question concerning Technology* (1977), Heidegger uses water-power plants in the river Rhine and some other examples to show that all objects are 'standing reserves' (*Bestände*) in a world-encompassing, inescapable technological system called 'the enframing' (*das Gestell*). Thus, Heidegger seems to employ the fact that something is used as a criterion for regarding it as part of the technological system — following the same track as our plan-relativist metaphysics of artefacts.

The same goes for a consequence that is relevant to a less intuitive way of doing metaphysics. Most proposals that are currently entertained in analytic metaphysics reconstruct or revise everyday language and the intuitions expressed in it to some extent. This price is paid to achieve a more rigorous or parsimonious categorisation of objects. Proponents of such rigour may frown upon the fact that, on our proposal, artefacts would not have clear and precise identity criteria; they would, in this respect, inherit the vagueness of use plans. When we introduced the notion of 'use plan' in chapter 2, we listed a number of ways in which use plans might be told apart, e.g., by having different goal states, different orderings of the same actions, and different objects manipulated.[22] We admitted that these criteria were far from exhaustive, that they might be related, and that their application was highly context-sensitive. Artefacts, when defined in terms of use plans in which they are manipulated, would be just as indeterminate. Take, for instance, the tea bags from section 2.2. Suppose that someone uses a tea bag by moving it up and down in the tea, making a quick cup of tea, and that someone else uses it in a teapot, letting the tea draw for some time before pouring a cup. The actions included in these plans seem sufficiently distinct to tell them apart. Yet it is unclear whether we ought to regard the tea bags manipulated in them as different artefacts. One might base an answer on the strength of one's intuitions about proper use, or beliefs about the quality of the tea produced by executing the plans, but there is no denying that the identity criteria for artefacts become messy and context-sensitive.

Quine's dictum 'No entity without identity' may then be imposed with its usual force: if one seeks to define a category of entities, but one does not succeed in giving precise criteria for claiming when two entities in this category are the same, the definition should be rejected or the sought category is ontologically suspicious.[23] Hence, plan-relative artefacts are discredited by their vagueness and join mental properties, events, and numbers (to name but a few) as victims of one of the most ruthless programs of ontological puritanism in the history of philosophy. Those who sympathise with Quine's physicalism and distrust of abstract objects may argue that this program resembles trimming a very unkempt beard rather than a massacre. Others might conclude that Quine's dictum sacrifices too many bona fide entities to the elusive ideal of metaphysical rigour, and might therefore not object to such vague objects as plan-relative artefacts. We just note their vagueness and go on exploring.

A third consequence of a plan-centred metaphysics of artefacts, the final one studied here, is that the notion of 'artefact' looses any semblance of absoluteness. Throughout this book, we have emphasised the versatility of use plans, finding them in everyday and expert use, in abbreviated and bracketed forms, etc. Moreover, we stressed the fact that every agent can be a designer, although only some designers may be socially accepted as such. From an epistemic or action-theoretical perspective, this versatility of use plans reflects the truism

[22] On the plan-centred metaphysics of artefacts discussed in this section, it is impossible to list 'artefacts used' among the difference or identity criteria of use plans, on pain of circularity.

[23] Quine's criterion is widely adopted in debates about the status of events, actions, mental states and other abstract objects; for an elaborate defense of the criterion, see Lowe (1995).

that objects can be used in multiple ways, i.e., that they may be manipulated for various purposes. Rephrased as a metaphysical thesis about artefacts, however, the same versatility leads to the view that an object is one artefact with respect to one use plan or community of users, and a different artefact with respect to another plan and community. There would be no single correct answer to the question what type of artefact an object is: one person's screwdriver is another's can opener.

This relativism can be mitigated somewhat by appealing to proper use. Only some designers are accepted as such in a social division of labour; thereby, some use plans are privileged and play a role in a network of responsibilities and requirements, as briefly indicated in section 5.3; other use plans merely involve recommendations for use. Categorising an object as specimen of a type of artefact may be, as we indicated earlier, one way of expressing that, among all known effective plans, a specific use plan is proper.

This puts some limitations on the versatile nature of artefacts; it may, for instance, be said that a screwdriver is not a paint-can opener, even if it is used as such. For other items and in other situations, however, relativism still runs rampant. This goes perhaps most clearly for items that are very large or extended in time. The river Rhine, for instance, is in some places properly used as a cooling agent, in the sense that people are not allowed to swim or fish in it; in other places, it is properly used as swimming water, but may not be used for diving or draining industrial waste. In all cases, there are regulations and responsibilities that strongly resemble those in the network created by claims about proper use — suggesting that we might just as well say that the Rhine has been designed as and is properly used as a cooling agent in one place, and has been designed as and is properly used for recreational purposes in the other. Similarly, the proper use of old buildings may have changed several times during their existence, and it may not be clear whether this means that several artefacts exist consecutively or simultaneously: did the Colosseum ever cease to be an arena? Is Cordoba's Mezquita a church, a mosque, just a stunning work of art, or all of the foregoing for various groups of people? In all cases, the problems with the nature of the artefact reflect vagaries of the use plan, which is more indeterminate for rivers and buildings than for 'hand-held' artefacts such as can openers. But items that are equipped with ever more 'functionalities', such as cellular phones and organisers, would be many things simultaneously: they play a role in many determinate use plans, or none clearly, if their use is very open-ended.

This relativism does not amount to a reductio. It does, however, commit one to a thesis of relative identity with respect to artefacts. It is an open question whether such a thesis, on which two objects can be the same F without being the same G (where F and G are both artefact types), is sustainable, or whether it is needed.[24] One might therefore be prepared to bite this bullet.

[24] It would be fair to say that, although the jury has still to reach unanimity, relative identity is believed to be innocent by a minority. The thesis of relative identity is primarily associated with Geach (1967), and was later defended by Noonan (1980) and forcefully attacked by, among others, Wiggins (2001, ch. 1). Garbacz (2002) has recently presented a proposal for

A related consequence of our plan-centred metaphysics of artefacts, however, seems puzzling irrespective of this ongoing debate in general metaphysics. This consequence can be derived from several features of use plans, in particular the identification of plans in terms of practical rationality (discussed in section 2.2). If, for instance, the ordering of actions within a plan can be changed without loss of rationality, the series of actions in reverse order is the same plan as the original. This leads to a curious phenomenon, which may be called *system dissolution*. Consider two artefacts that may be used while making breakfast: a tea bag that is manipulated in the course of executing the 'Sophisticated-Dutch' plan for making tea, and a toaster that is manipulated for making toast, for instance in the way described at the beginning of chapter 2. Both of these plans are designed, communicated (even standardised) and defeasibly rational, in terms of both effectiveness and efficiency. Yet they may be straightforwardly combined into an equally effective and more efficient breakfast-making plan, by exploiting the fact that both use plans include waiting periods:[25] one might make a piece of toast while the tea bag is suspended in the hot water, since the extraction process takes place without the need for further actions. If one wants to have both tea and toast for breakfast, combining the execution of both plans saves time in comparison to executing them consecutively. Hence, the tea-and-toast plan may be differentiated, in terms of rationality, from both the tea plan and the toast plan; and if these two plans cannot be reduced to their component actions,[26] the tea-and-toast plan cannot be reduced to the component plans. And if one wants to have tea, toast *and* the morning news for breakfast, combining the execution of three plans into a 'Continental-Breakfast plan' is more efficient than combining two and then executing another.

This phenomenon of effective plan combination, whereby plans dissolve into larger plans, is intriguing from an action-theoretical perspective, and has, to the best of our knowledge, never been studied,[27] but the practical relevance of plan combination is hard to deny. Such combinations are not merely a matter of personal designing, either, but may be communicated to other users and may even be standardised. The Dutch national railway company, for instance, advertises train travelling by showing people that comfortably read books while speeding to their destination. Doubts about veracity left aside, this shows that two use plans (for travelling and reading) may be profitably combined. Similarly, eating bags of potato chips and drinking beer while watching soccer is a combination of plans that is often communicated, albeit not often in positive terms. From

the logical representation of statements of relative identity, and Carrara (2009) reviews the consequences of the relative identity thesis for artefacts.

[25] Here, we appeal to another feature of use plans, described by Bratman (2000), namely their temporal extension.

[26] This feature might be called the 'normative irreducibility' of plans, and is argued for by, for instance, Pollock (1995, §5.2).

[27] Combinations of plans, goals and artefacts bring into mind the 'totality of involvements' of the early Heidegger, and the 'lifeworld' or 'practice' of the later Wittgenstein. Although these notions are intuitively appealing, they are also very ambiguous and indeterminate. Analysing the phenomena addressed by these notions in an analytically satisfactory way would appear to be a big challenge for a philosophy of action and of artefacts.

an action-theoretical point of view, one may wonder whether efficiency is the only value at stake here (some combinations, such as the potato-and-television one, may express a lifestyle rather than a concern for efficiency), and where the combination of plans ends or becomes too idiosyncratic.

Once we look at this phenomenon from the perspective of our plan-centred metaphysics of artefacts, however, it becomes puzzling rather than intriguing. On this metaphysics, the tea bag, the toaster and the newspaper are *identified* as the various objects manipulated in the course of executing the Sophisticated-Dutch plan for making tea, Anna's plan for making toast, and the simple plan for reading the morning news respectively. All of these plans are designed, communicated, rational, and even proper, and therefore satisfy the conditions introduced in the useful-materials definition of artefacts, as given at page 148. However, the Continental-Breakfast plan is just as much designed, communicated and proper (or at least standardised) as the three individual plans, and we have just argued that executing it is more efficient and therefore more rational than executing the three plans consecutively. Now suppose that we identify the tea bag, toaster and newspaper with respect to the Continental-Breakfast plan. Since this plan is relevantly different from the individual plans, we are forced to conclude that the artefacts manipulated in it are also different from the artefacts manipulated in the individual plans: our plan-centred metaphysics would be a farce if differences at the plan level would not be reflected in differences at the artefact level. Therefore, a Continental-Breakfast-toaster is a different artefact than a toast-making-toaster.

This, however, seems absurd. Intuitively, artefacts do not dissolve when they are included in a system,[28] no matter how important systems of artefacts may be in our everyday lives. This judgement is based on an intuition about the nature of artefacts that seems hard to discard; harder, perhaps, than the intuition about natural and artificial objects, which is commonly accepted as being in need of explication. Moreover, unlike the other consequences discussed above, the system dissolution of artefacts is an immediate consequence of the useful-materials definition of artefacts and several features that were earlier used *in favour* of a planning approach; hence, it seems to indicate the intuitive boundaries of this planning approach, showing that it is advantageous as a theory of (artefact) actions, but at least incomplete as a theory of artefacts.

Still, the plan-centred metaphysics of artefacts avoids the problems of function essentialism. Given its basis in our use-plan account, it does not run into trouble regarding the phenomenology of artefact using and designing, such as the problem of ontological stacking that beset the function-essentialist constitution view of artefacts. Metaphysics may not be an area of philosophy where

[28]Intuitions may not converge for all artefacts, since some artefacts may plausibly be regarded as parts of a system rather than 'stand-alone' items. Both computer software and hardware, for instance, can arguably be regarded as artefacts that complete or complement each other: an operating system is an artefact, but it is useless without a desktop computer, which is as useless without the operating system. Similarly, cans and can openers may be regarded as 'symbiotic' artefacts. Whether a principled distinction can be drawn between such symbionts and components of a single artefact is an open question.

theories are judged by their relative merits; yet, so far, the only alternative to accepting plan relativism is nihilism about the existence and nature of artefacts. Therefore, in the final section of this book, we examine whether we can supplement a plan-centred metaphysics of artefacts in such a way that some of its counterintuitive or otherwise unpalatable consequences may be avoided.

7.4 Useful and man-made materials

In this final section, we briefly consider one way of increasing the intuitive appeal of our plan-centred metaphysics, namely to supplement it with a characterisation of artefacts as man-made objects.[29] One way to introduce the resulting combination, and need thereof, is as follows. In the introductory chapter, we mentioned Randall Dipert's theory of artefacts as one of the sources of inspiration for our own work. In retrospect, however, one might object that our analyses and results elaborate on only one part of Dipert's theory, and that we have more or less ignored the rest, including, ironically, Dipert's notion of artefact. For in his (1993), Dipert distinguishes three notions that, in combination, describe the realm of artificial objects: instrument, tool and artefact.[30] These notions are supposed to define classes of objects that are subsets of each other: instruments are objects used for practical purposes; tools are objects used and changed with regard to their use; and artefacts are objects used, changed with regard to their use, and objects that communicate that they have been so changed. If our theory is compared to Dipert's, it may seem that we mainly discuss instruments rather than artefacts. Our notion of designing, in particular, contains no reference to physical modification; although this 'poietical' aspect is an integral part of product designing as we described it in section 2.5, we explicitly introduced product designing as only one type of designing. Many conclusions drawn in this book depend on the general notion of designing; without it, the distinction between rational and proper use would collapse, and our account of the acquisition of additional proper functions would fail. In our account of designing, and of using as well, the notion of use plan holds centre stage. And although we share an emphasis on communication with Dipert, the content of this communication is the use plan and not the fact that the artefact has, somehow, been physically changed to enable or facilitate use. Thus, the 'useful-materials' definition on page 148 makes explicit that, throughout this book, we focus on instruments rather than artefacts or tools.

This suggests that our metaphysics of artefacts may be profitably supplemented with a view on production. Yet we first show that we cannot not just

[29] Characterising artefacts as man-made objects is a central starting point in Hilpinen's (1992, 1993, 2008) analyses of artefacts, especially in his (2008).

[30] Dipert (1995) formulates a similar distinction, but phrased in terms of properties instead of types of objects. The main reason is that, on Dipert's characterisations, the distinction between instruments, tools and artefacts cannot be regarded as an absolute one. Presumably to avoid a relative identity thesis about objects, Dipert reformulated his characterisations in terms of properties. Since we regard relative identity as a non-fatal consequence, we see no pressing need to adopt Dipert's reformulation.

replace the useful-materials definition with another, and only then go on to combining the two.

The most straightforward definition of artefacts as man-made objects is the following:

A 'man-made materials' definition of artefacts:

An object x is an artefact a if and only if x has been intentionally produced by an agent m.

The reference to intentions is needed to avoid some counterintuitive consequences, such as the classification of waste products such as sweat and soot as artefacts; the other elements seem relatively unproblematic. However, appearances deceive, for there are at least two major problems with this definition.

Firstly, if one characterises artefacts exclusively on the basis of their origin, i.e., their history of production, one cannot account for the rich dynamics of using artefacts. Some artefacts, like Aspirin, acquire new uses that co-exist peacefully with the old ones. In other cases, the new use replaces the old one. This replacement may be gradual — plate armour went from being a protective device to being exclusively ceremonial; or it may be sudden — according to both Roman-Catholic and Islamic canonical law, churches, mosques and objects used in religious ceremonies can be desecrated and thus cease to be useful for their original purpose when certain (unspecified) sacrilegious acts take place in or with them. A man-made metaphysics of artefacts rules out a priori that changes in using artefacts may lead to changes in their nature: the original production of an artefact determines its nature once and for all.[31]

Secondly, it is difficult to develop the man-made metaphysics in such a way that its intuitive difference with the plan-centred metaphysics of artefacts is retained. To illustrate this: if 'useful-materials' artefacts are analysed in a constitution view, an artefact with at least one new and essential causal property comes into existence as soon as an agent deliberately manipulates an object while executing a use plan. It does not matter whether the object existed prior to this manipulation: the artefact, not the aggregate that constitutes it, is created by its inclusion in a use plan. The Aspirin case, for example, was analysed in terms of an aggregate of acetylsalicylic acid, which constitutes a painkiller in one plan and a blood thinner in another; the painkiller may then be said to cease to exist when the blood thinner is created. Since the man-made-materials definition of artefacts does not mention physical modifications of the object, i.e., says nothing about getting one's hands dirty, it is compatible with the plan-centred metaphysics and thus does not present a genuine alternative to it.[32] Developing an appropriately 'hands-on' notion of producing or making is,

[31]A similar criticism, aimed at theories that characterise artefact *functions* exclusively in terms of author's intentions, has been formulated by Preston (2003).

[32]This point about a 'man-made metaphysics of artefacts' can be illustrated with Hilpinen's analysis of artefacts. In his (1992, p. 61) he states the following condition for artefacthood: 'An object o is an artifact made by an agent Ag only if it satisfies some type-description D included in the intention I_A which brings about the existence of o', and later adds the

however, no mean feat either. If one requires an artefact to be produced from raw materials, the stereotypical driftwood raft or any other product of assembly is not an artefact. Alternatively, if producing means that something is physically changed to serve a practical purpose, Jupiter is again turned into an artefact by using it to accelerate interplanetary probes, since this use affects Jupiter (albeit very marginally). Closer to home, every episode of use results in wear and tear, leading to an enormous variety of use-related physical modification of objects. Moreover, since everyone who knows some basic physics ought to be aware of these changes, they are, to some extent, intentional effects. That they are side effects, perhaps even negligible effects, rather than the goal of production or use is, of course, a relevant observation, but spelling it out would require a distinction between intentions that are relevant and those that are irrelevant to the production process. This distinction puts one's feet firmly on the path that leads from a naive intentionalist theory to an analysis in terms of use plans; those interested in retracing this path may want to consult section 3.2 before setting out. We surmise that developing a sustainable notion of production leads to our notion of product designing, which is part of a general theory of designing that also avoids the first problem of a man-made metaphysics. Therefore, we conclude that exclusively characterising artefacts as man-made materials leads to fatal problems or collapses into a plan-centred approach.

Still, characterising artefacts as man-made materials has one marked advantage: it seems to provide a basis for distinguishing artificial objects from natural objects. Information about the use plan, especially about the proper or standard plan, may be needed to characterise an artefact as a token of a type, but we saw earlier that relativising the nature of artefacts to plans runs the risk of eliminating the distinction between the natural and the artificial. Characterising something as an artefact, rather than as a token of an artefact type, seems an absolute matter, which does not vary among agents or change in time. One may increase one's abilities to discriminate between artefacts and natural objects, but it would be counterintuitive to relativise this distinction to one's abilities. Characterising artefacts as man-made materials provides the requisite absolutism. On this characterisation, an object becomes an artefact on its original production: if it was intentionally produced by a human being, it is an artefact for everyone, everywhere and always. This absolutism is not just a matter of intuitions, for it seems that people are very good at identifying man-made materials, and that this identification is in large part independent of time and culture: people can tell Stone Age artefacts from rocks, and are able to identify the products of other cultures as artefacts, even if they have no idea

condition that some properties of the artefact should be counterfactually dependent on the agent's intention. On the plan-centred metaphysics of artefacts, the relevant type-description D could be taken to refer to the goal state to which execution of the use plan is supposed to contribute (examples would be 'whitener', 'container' or 'conductor'), the 'creative' intention I_A could be taken as the intention that is central to designing, i.e., to contribute to other agents realising their goals, and the intention-dependent property would be any new causal power of the constituted artefact (again: there should, on the constitution view, be at least one of these). Then, Hilpinen's conditions would express a plan-centred metaphysics instead of a conception of artefacts as man-made materials different from it.

what they are for.[33] Knowing that something is an artefact is, to some extent, different from knowing what it can be used for.

This suggests a combination of the useful- and man-made-materials definitions. The latter serves, to some extent, to distinguish artefacts from other types of objects in an appropriately absolutist way; however, it cannot be used to distinguish types of artefacts. The former can be used to characterise types of artefacts, albeit in a highly context-sensitive manner; yet it performs poorly in distinguishing artificial and natural materials. Thus, these two views are complementary: the characterisation of artefacts as man-made objects might be used to mark out the domain of application of the theory of function ascriptions, in which artefact types can then be defined implicitly. This retains the relativism at the level of artefact types, but makes possible a more absolute distinction between artefacts and natural objects. In this way, more intuitions might be accommodated by the combination than by either view individually.

> **A dual definition of artefacts:**
>
> An object x is an artefact a of type t if and only if: (1) x has been intentionally produced by an agent m; and (2) x is manipulated in the course of executing a specific use plan p, which is designed, communicated and evaluated in accordance with the use-plan analysis of using and designing.

This proposal contains several concepts that should be analysed in more detail. A major complication in doing this is that the two conditions cannot be neatly separated. It is tempting to regard the first condition as determining the domain of artefacts, contrasting it with natural objects and perhaps other categories, and the second condition as determining types of artefacts within the broader domain of artificial objects. This temptation should be resisted, as the troubles with developing the man-made metaphysics show: to develop the notion of intentional production in such a way that it accommodates our intuitions about artefacts, the relevant (designer or user) intentions ought to be specified, and there is ample reason to suppose that this specification must rely on the notions used in the second condition.

There is also phenomenological evidence against the independence of the two conditions — and, more importantly, against the independence of the domain of artificial objects and the various artefact types within it. By introducing, on the basis of beliefs about effective use, artefact types within the domain of

[33] Once again, some results of experiments in cognitive science bear this out: Gelman and Bloom (2000) report that both children and adults are more likely to categorise an object as some kind of artefact when they learn that it was created intentionally than when they are informed that it is the result of a chance process. Gelman and Bloom take this as supporting the claim that artefacts have functional essences; we would rather say that the information about the production of an item only leads subjects to conclude that it is an artefact and that their categorising it as token of an artefact type may be the result of the experimental setup. In general, people do not need to categorise an item as a token of an artefact type in order to categorise it as an artefact.

artefacts, objects outside of this domain might be suggested as useful equivalents. If someone has forgotten to bring his homemade tent pegs to the camping ground, he may start looking for pieces of wood that have approximately the same shape and rigidity, and modify them slightly to fit their use. Through design and communication of an appropriate use plan, even unmodified natural objects might gradually be regarded as artefacts, to which regulations about proper use apply. Reversing the process, artefacts may gradually be accepted as parts of nature; in the Netherlands, this is the case for some man-made landscapes, created through digging peat, building dikes or even constructing highways and railroads. If we accept that the boundary between natural and artificial objects changes along with changes in (proper) use, the two conditions in the last definition become ever more intertwined.

Still, the basic elements in a characterisation of artefacts have gradually become clear in the course of this discussion: proper and effective use, the physical structure supporting this use, and the history of design and/or production of the item.[34] All these elements but one — the production history — are familiar from earlier parts of this book. Throughout this book, we outlined ways in which these elements are related, both in evaluative reconstructions of using, designing and function ascriptions, and in more descriptive analyses of the way in which artefacts are used and designed, and in which they are (not) ascribed functions.

The arguments in this chapter show, negatively, that these analyses cannot be directly extrapolated to a theory about the nature and categorisation of artefacts: significant conceptual effort is needed to accommodate intuitions about artefacts, even if a phenomenologically adequate theory about using, designing and functions is in place. However, the arguments also show, positively, that most of the elements for a theory of artefacts are at hand. We have analysed designing and proper and effective using, and ways in which these actions are supported by (beliefs about) the physical structure of an artefact. The only element overlooked in our lengthy prospecting of the domain of artefacts is production.

Still, the challenge in constructing a satisfactory theory of artefacts does not lie in giving a list of relevant elements, but in showing how they are related — just as the challenge in constructing a theory of artefact functions does not lie in identifying intentions, causal roles and evolution as the relevant elements, but in relating these in a phenomenologically satisfactory way. Our identification of the relevant elements therefore only provides the starting point for constructing

[34]Recent results in cognitive science show that precisely these four elements play a role in the way in which people categorise artefacts. Chaigneau, Barsalou, and Sloman (2004) report that considerations about use and supporting structure are not always decisive — as suggested by the results that we referred to in section 7.1. In situations where information about use and/or structure is lacking, subjects categorise items on the basis of information about its design and production history. Given our own results, this does not come as a surprise: in section 5.3, we indicated that the testimony of professional designers is needed to make claims about proper use. Chaigneau, Barsalou and Sloman regard their results as undermining Bloom's function essentialism (see section 7.2); we might, in addition, interpret them as confirming our focus on proper use rather than functions.

a theory of artefacts.[35] Going beyond this would require a further inquiry into what is at stake in using the notion of artefact. As indicated in this chapter, it appears to play various roles, such as distinguishing artificial objects from natural ones, expressing standards for proper use, and identifying useful equivalents. Regarding artefacts as man-made objects focuses on some of these roles, regarding them as useful materials focuses on others: there appears to be an inherent duality in the notion of artefact.

The best way to construct a theory of artefacts that integrates the man-made and useful perspectives and that accommodates all relevant intuitions is to start a new project of conceptual engineering — first giving a list of desiderata, based on relevant intuitions about artefacts, then designing a candidate theory and testing it against the criteria. This project does not need to start from scratch: all components are available. Yet constructing a theory of artefacts may not be off-the-shelf engineering: not only do some components need considerable fine-tuning, but there may well be conflicts between intuitions and partial analyses.[36] It is not a given that all desiderata can be satisfied; in engineering practice, conflicts between client wishes regularly need to be resolved by ignoring some of them. We managed to satisfy all desiderata for our theory of technical functions. It remains to be seen whether the same may be done for a theory of artefacts.

[35] In philosophy, a similar starting point is provided by Losonsky (1990), who identifies physical structure and the purpose and manner of use as determining the nature of artefacts. We would add the distinction between proper and effective use and considerations about design and production to this list. In cognitive science, Barsalou, Sloman, and Chaigneau (2004) have proposed the HIPE-theory, which is officially a theory of artefact functions, but is better understood as one of artefacts. On this theory, design *H*istory, *I*ntended use, *P*hysical structure, and the *E*vents resulting from use are the relevant elements in artefact categorisation. Again, both production and a conceptual integration of the various elements are lacking. Thus, it seems that in philosophy and elsewhere, the major challenge in constructing a theory of artefacts has so far been avoided.

[36] Reading Ingold (2000) gives a good impression of the possible differences between analysing making and, as we have done, analysing designing. In section 2.5, we reconstructed product designing as a type of designing; the resulting emphasis on deliberation and planning are markedly different from Ingold's focus on skilled movement and interaction with the environment — i.e., on embodiment and situatedness.

Bibliography

Akrich, M. (1992). The de-scription of technical objects. In W. E. Bijker and J. Law (Eds.), *Shaping Technology / Building Society: Studies in Sociotechnical Change*, pp. 205–224. Cambridge, MA: MIT Press.

Allen, C., M. Bekoff, and G. Lauder (Eds.) (1998). *Nature's Purposes: Analyses of Function and Design in Biology*, Cambridge, MA. Bradford Books.

Ariew, A. and M. Perlman (2002). Introduction. In A. Ariew, R. Cummins, and M. Perlman (Eds.), *Functions: New Essays in the Philosophy of Psychology and Biology*, pp. 1–4. Oxford: Oxford University Press.

Aunger, R. (2002). *The Electric Meme*. New York: Free Press.

Baker, L. R. (2000). *Persons and Bodies*. Cambridge: Cambridge University Press.

Baker, L. R. (2004). The ontology of artefacts. *Philosophical Explorations 7*, 99–111.

Baker, L. R. (2006). On the twofold nature of artefacts. *Studies in History and Philosophy of Science 37*, 132–136.

Baker, L. R. (2007). *The Metaphysics of Everyday Life: An Essay in Practical Realism*. Cambridge: Cambridge University Press.

Barsalou, L. W., S. A. Sloman, and S. E. Chaigneau (2004). The HIPE theory of function. In L. Carlson and E. van der Zee (Eds.), *Functional Features in Language and Space: Insights from Perception, Categorization and Development*, pp. 131–148. Oxford: Oxford University Press.

Barton, M. E. and L. K. Kamatsu (1989). Defining features of natural kinds and artifacts. *Journal of Psycholinguistic Research 18*, 433–447.

Basalla, G. (1988). *The Evolution of Technology*. Cambridge: Cambridge University Press.

Bell, J., N. Snooke, and C. Price (2005). Functional decomposition for interpretation of model based simulation. In M. Hofbaur, B. Rinner, and F. Wotawa (Eds.), *Proceedings of the 19th International Workshop on Qualitative Reasoning (QR-05), Graz, Austria, 18-20 May 2005*, pp. 192–198.

W. Houkes, P.E. Vermaas, *Technical Functions*, Philosophy of Engineering and Technology 1, DOI 10.1007/978-90-481-3900-2,

Bigelow, J. and R. Pargetter (1987). Functions. *Journal of Philosophy 84*, 181–196. Reprinted in Allen et al. (1998, pp. 241–259).

Blackburn, S. (1998). *Ruling Passions*. Oxford: Clarendon.

Bloom, P. (1996). Intention, history, and artifact concepts. *Cognition 60*, 1–29.

Bloom, P. (1998). Theories of artifact categorization. *Cognition 66*, 87–93.

Bratman, M. (1987). *Intentions, Plans, and Practical Reason*. Cambridge, MA: Harvard University Press.

Bratman, M. (1999). *Faces of Intention: Selected Essays on Intention and Agency*. Cambridge: Cambridge University Press.

Bratman, M. (2000). Reflection, planning, and temporally extended agency. *Philosophical Review 109*, 35–61.

Brey, P. (2008). Technological design as an evolutionary process. In P. E. Vermaas, P. Kroes, A. Light, and S. A. Moore (Eds.), *Philosophy and Design: From Engineering to Architecture*, pp. 61–75. Dordrecht: Springer.

Brown, D. C. and L. Blessing (2005). The relationship between function and affordance. In *ASME 2005 IDETC/CIE Conference, September 24-28, 2005, Long Beach, California, USA*. DECT2005-85017.

Buller, D. J. (1998). Etiological theories of function: A geographical survey. *Biology and Philosophy 13*, 505–527.

Carlson, W. B. (2000). Invention and evolution: The case of Edison's sketches of the telephone. In J. Ziman (Ed.), *Technological Innovation as an Evolutionary Process*, pp. 137–158. Cambridge: Cambridge University Press.

Carrara, M. (2009). Relative identity and the number of artefacts. *Techné 13*, 108–122.

Castañeda, H. N. (1970). On the semantics of the ought-to-do. *Synthese 21*, 449–468.

Chaigneau, S. E., L. W. Barsalou, and S. A. Sloman (2004). Assessing the causal structure of function. *Journal of Experimental Psychology: General 133*, 601–625.

Clarkson, J., R. Coleman, S. Keates, and C. Lebbon (Eds.) (2003). *Inclusive Design: Design for the Whole Population*, Berlin. Springer-Verlag.

Clarkson, J. and S. Keates (Eds.) (2003). *Countering Exclusive Design: An Introduction to Inclusive Design*, Berlin. Springer-Verlag.

Collins, H. M. and R. Evans (2002). The third wave of science studies: Studies of expertise and experience. *Social Studies of Science 32*, 235–296.

Constant, E. (2000). Recursive practice and the evolution of technological knowledge. In J. Ziman (Ed.), *Technological Innovation as an Evolutionary Process*, pp. 219–233. Cambridge: Cambridge University Press.

Cummins, R. (1975). Functional analysis. *Journal of Philosophy 72*, 741–765. Reprinted in Allen et al. (1998, pp. 169–196).

Dancy, J. (2006). Ethical non-naturalism. In D. Copp (Ed.), *The Oxford Handbook of Ethical Theory*, pp. 122–145. Oxford: Oxford University Press.

Davies, P. S. (2000). The nature of natural norms: Why selected functions are systemic capacity functions. *Noûs 34*, 85–107.

Davies, P. S. (2001). *Norms of Nature: Naturalism and the Nature of Functions*. Cambridge, MA: MIT Press.

Dawkins, R. (1976). *The Selfish Gene*. Oxford: Oxford University Press.

Dennett, D. C. (1971). Intentional systems. *Journal of Philosophy 68*, 87–106. Reprinted in Dennett (1978, pp. 3–22).

Dennett, D. C. (1978). *Brainstorms: Philosophical Essays on Mind and Psychology*. Cambridge, MA: Bradford Books.

Dennett, D. C. (1990). The interpretation of texts, people and other artifacts. *Philosophy and Phenomenological Research 50(S)*, 177–194.

Dennett, D. C. (1995). *Darwin's Dangerous Idea*. New York: Norton.

Dick, P. K. (1957). *Eye in the Sky*. New York: Ace.

Dipert, R. R. (1993). *Artifacts, Art Work, and Agency*. Philadelphia: Temple University Press.

Dipert, R. R. (1995). Some issues in the theory of artifacts: Defining 'artifact' and related notions. *Monist 78*, 119–135.

Dorst, K. and P. E. Vermaas (2005). John Gero's Function-Behaviour-Structure model of designing: A critical analysis. *Research in Engineering Design 16*, 17–26.

Dreier, J. (2001). Humean doubts about categorical imperatives. In E. Millgram (Ed.), *Varieties of Practical Reasoning*, pp. 27–49. Cambridge, MA: MIT Press.

Eekels, J. and W. A. Poelman (1998). *Industriële Productontwikkeling. Deel 1: Basiskennis*. Utrecht: Lemma.

Elder, C. (2004). *Real Natures and Familiar Objects*. Cambridge, MA: MIT Press.

Elder, C. (2007). On the place of artifacts in ontology. In E. Margolis and S. Laurence (Eds.), *Creations of the Mind: Theories of Artifacts and Their Representation*, pp. 33–51. Oxford: Oxford University Press.

Erden, M. S., H. Komoto, T. J. van Beek, V. D'Amelio, E. Echavarria, and T. Tomiyama (2008). A review of function modeling: Approaches and applications. *Artificial Intelligence for Engineering Design, Analysis and Manufacturing 22*, 147–169.

Faulkner, W. (1994). Conceptualizing knowledge used in innovation: A second look at the science-technology distinction and industrial innovation. *Science, Technology and Human Values 19*, 425–458.

Fleck, J. (2000). Artefact–Activity: The coevolution of artefacts, knowledge and organization in technological innovation. In J. Ziman (Ed.), *Technological Innovation as an Evolutionary Process*, pp. 248–266. Cambridge: Cambridge University Press.

Franssen, M. (2006). The normativity of artefacts. *Studies in History and Philosophy of Science 37*, 42–57.

Galison, P. (1986). *How Experiments End*. Chicago: University of Chicago Press.

Garbacz, P. (2002). The logics of relative identity. *Notre Dame Journal of Formal Logic 43*, 27–50.

Geach, P. T. (1967). Identity. *Review of Metaphysics 21*, 3–12.

Gelman, S. A. and P. Bloom (2000). Young children are sensitive to how an object was created when deciding what to name it. *Cognition 76*, 91–103.

Gentner, D. (1978). What looks like a jiggy but acts like a zimbo? A study of early word meaning using artificial objects. *Papers and Reports on Child Language Development 15*, 1–6.

Gero, J. S. (1990). Design prototypes: A knowledge representation schema for design. *AI Magazine 11*(4), 26–36.

Godfrey-Smith, P. (1994). A modern history theory of functions. *Noûs 28*, 344–362.

Goldman, A. I. (1999). *Knowledge in a Social World*. New York: Oxford University Press.

Goldman, A. I. (2001a). Experts: Which ones should you trust? *Philosophy and Phenomenological Research 63*, 85–110.

Goldman, A. I. (2001b). Social routes to belief and knowledge. *Monist 84*, 346–367.

Goldsmith, S. (1997). *Designing for the Disabled: A New Paradigm*. London: Architectural Press.

Gooding, D., T. Pinch, and S. Schaffer (Eds.) (1989). *The Uses of Experiment*, Cambridge. Cambridge University Press.

Gould, S. J. and E. S. Vrba (1982). Exaptation: A missing term in the science of form. *Paleobiology 8*, 4–15. Reprinted in Allen et al. (1998, pp. 519–540).

Griffiths, P. E. (1993). Functional analysis and proper functions. *British Journal for the Philosophy of Science 44*, 409–422. Reprinted in Allen et al. (1998, pp. 435–452).

Hacking, I. (1983). *Representing and Intervening*. Cambridge: Cambridge University Press.

Haggard, P. (1998). Planning of action sequences. *Acta Psychologica 99*, 201–215.

Heidegger, M. (1962). *Being and Time*. Oxford: Blackwell.

Heidegger, M. (1977). *The Question concerning Technology and Other Essays*. New York: Harper and Row.

Hilpinen, R. (1992). Artifacts and works of art. *Theoria 58*, 58–82.

Hilpinen, R. (1993). Authors and artifacts. *Proceedings of the Aristotelian Society 93*, 155–178.

Hilpinen, R. (2008). Artifact. In E. N. Zalta (Ed.), *The Stanford Encyclopedia of Philosophy (Fall 2008 Edition)*. http://plato.stanford.edu/archives/fall2008/entries/artifact/.

Houkes, W. (2006). Knowledge of artefact functions. *Studies in History and Philosophy of Science 37*, 102–113.

Houkes, W. (2008). Designing is the construction of use plans. In P. E. Vermaas, P. Kroes, A. Light, and S. A. Moore (Eds.), *Philosophy and Design: From Engineering to Architecture*, pp. 37–49. Dordrecht: Springer.

Houkes, W. and A. W. M. Meijers (2006). The ontology of artefacts: The hard problem. *Studies in History and Philosophy of Science 37*, 118–131.

Houkes, W. and P. E. Vermaas (2004). Actions versus functions: A plea for an alternative metaphysics of artifacts. *Monist 87*, 52–71.

Houkes, W. and P. E. Vermaas (2006). Planning behaviour: Technical design as design of user plans. In P.-P. Verbeek and A. Slob (Eds.), *User Behaviour and Technology Development: Shaping Sustainable Relations between Consumers and Technologies*, Volume 20 of *Eco-Efficiency in Industry and Science*, pp. 303–210. Dordrecht: Springer.

Houkes, W., P. E. Vermaas, K. Dorst, and M. J. de Vries (2002). Design and use as plans: An action-theoretical account. *Design Studies 23*, 303–320.

Hubka, V. and W. E. Eder (1988). *Theory of Technical Systems: A Total Concept Theory for Engineering Design*. Berlin: Springer-Verlag.

Hudson, J. A. and R. Fivush (1991). Planning in the preschool years: The emergence of plans from general event knowledge. *Cognitive Development 6*, 393–415.

Hudson, J. A., B. B. Sosa, and L. R. Shapiro (1997). Scripts and plans: The development of preschool children's event knowledge and event planning. In S. L. Friedman and E. K. Scholnick (Eds.), *The Developmental Psychology of Planning: why, how, and when do we plan?*, pp. 77–102. Mahwah, NJ: Lawrence Erlbaum.

Ingold, T. (2000). On weaving a basket. In T. Ingold (Ed.), *The Perception of the Environment*, pp. 339–348. London: Routledge.

Jackson, F. (1998). *From Metaphysics to Ethics: A Defense of Conceptual Analysis*. Oxford: Oxford University Press.

Kemler-Nelson, D. G., R. Russell, N. Duke, and K. Jones (2000). Two-year-olds name artifacts by their function. *Child Development 71*, 1271–1288.

Kitcher, P. S. (1993). Function and design. In P. A. French, T. E. Uehling, and H. K. Wettstein (Eds.), *Midwest Studies in Philosophy*, Volume XVIII, pp. 379–397. Minneapolis: University of Minnesota Press. Reprinted in Allen et al. (1998, pp. 479–503).

Kornblith, H. (1980). Referring to artifacts. *Philosophical Review 89*, 109–114.

Kroes, P. A. (2003). Screwdriver philosophy; Searle's analysis of technical functions. *Techné 6*(3), 22–35. http://scholar.lib.vt.edu/ejournals/SPT/.

Kroes, P. A. and A. W. M. Meijers (2000). Introduction: A discipline in search of its identity. In P. A. Kroes and A. W. M. Meijers (Eds.), *The Empirical Turn in the Philosophy of Technology*, Volume 20 of *Research in Philosophy and Technology*, pp. xvii–xxxv. Amsterdam: JAI/Elsevier Science.

Kroes, P. A. and A. W. M. Meijers (2002). The dual nature of technical artifacts: Presentation of a new research program. *Techné 6*(2), 4–8. http://scholar.lib.vt.edu/ejournals/SPT/ See also the commentaries in the same issue.

Kroes, P. A. and A. W. M. Meijers (2006). The dual nature of technical artefacts. *Studies in History and Philosophy of Science 37*, 1–4.

Krohs, U. (2009). Functions as based on a concept of general design. *Synthese 166*, 69–89.

Lackey, J. (2003). A minimal expression of non-reductionism in the epistemology of testimony. *Noûs 37*, 706–723.

Lackey, J. and E. Sosa (Eds.) (2006). *The Epistemology of Testimony*, Oxford. Oxford University Press.

Landau, B., L. B. Smith, and S. S. Jones (1998). Object shape, object function, and object name. *Journal of Memory and Language 38*, 1–27.

Latour, B. (1992). The sociology of a few mundane artifacts. In W. E. Bijker and J. Law (Eds.), *Shaping Technology / Building Society: Studies in Sociotechnical Change*, pp. 225–258. Cambridge, MA: MIT Press.

Leudar, I. and A. Costall (1996). Situating action IV: Planning as situated action. *Ecological Psychology 8*, 153–170.

Lewens, T. (2004). *Organisms and Artifacts: Design in Nature and Elsewhere.* Cambridge, MA: MIT Press.

Lipton, P. (1998). The epistemology of testimony. *Studies in History and Philosophy of Science 29A*, 1–31.

Losonsky, M. (1990). The nature of artifacts. *Philosophy 65*, 81–88.

Lowe, E. J. (1995). The metaphysics of abstract objects. *Journal of Philosophy 92*, 509–524.

Mahner, M. and M. Bunge (2001). Function and functionalism: A synthetic perspective. *Philosophy of Science 68*, 75–94.

Malt, B. C. and E. C. Johnson (1992). Do artifact concepts have cores? *Journal of Memory and Language 31*, 195–217.

Matan, A. and S. Carey (2001). Development changes within the core of artifact concepts. *Cognition 78*, 1–26.

Matthen, M. (1997). Teleology and the product analogy. *Australasian Journal of Philosophy 75*, 21–37.

McLaughlin, P. (2001). *What Functions Explain.* Cambridge: Cambridge University Press.

Merleau-Ponty, M. (1962). *The Phenomenology of Perception.* London: Routledge.

Millikan, R. G. (1984). *Language, Thought, and Other Biological Categories: New Foundations for Realism.* Cambridge, MA: MIT Press.

Millikan, R. G. (1989). In defense of proper functions. *Philosophy of Science 56*, 288–302. Reprinted in Allen et al. (1998, pp. 295–312).

Millikan, R. G. (1993). *White Queen Psychology and Other Essays for Alice.* Cambridge, MA: MIT Press.

Millikan, R. G. (1999). Wings, spoons, pills, and quills: A pluralist theory of function. *Journal of Philosophy 96*, 191–206.

Mokyr, J. (1996). Evolution and technological change: A new metaphor for economic history? In R. Fox (Ed.), *Technological Change*, pp. 63–83. London: Harwood Publishers.

Mokyr, J. (2000). Evolutionary phenomena in technological change. In J. Ziman (Ed.), *Technological Innovation as an Evolutionary Process*, pp. 52–65. Cambridge: Cambridge University Press.

Mumford, S. (1998). *Dispositions.* Oxford: Oxford University Press.

Neander, K. (1991a). Functions as selected effects: The conceptual analyst's defense. *Philosophy of Science 58*, 168–184.

Neander, K. (1991b). The teleological notion of "function". *Australasian Journal of Philosophy 69*, 454–468. Reprinted in Allen et al. (1998, pp. 313–333).

Neander, K. (1995). Misrepresenting and malfunctioning. *Philosophical Studies 79*, 109–141.

Nelson, K. (Ed.) (1986). *Event Knowledge: Structure and Function in Development*, Hillsdale, NJ. Lawrence Erlbaum.

Noonan, H. W. (1980). *Objects and Identity.* The Hague: Nijhoff.

Norman, D. A. (1990). *The Design of Everyday Things.* New York: Perseus. First published as *The Psychology of Everyday Things.*

Pahl, G., W. Beitz, J. Feldhusen, and K.-H. Grote (2007). *Engineering Design: A Systematic Approach* (third ed.). Berlin: Springer-Verlag.

Perlman, M. (2004). The modern philosophical resurrection of teleology. *Monist 87*, 3–51.

Polanyi, M. (1962). *Personal Knowledge.* London: Routledge.

Pollack, M. E. (1990). Plans as complex mental attitudes. In P. R. Cohen, J. Morgan, and M. E. Pollack (Eds.), *Intentions in Communication*, pp. 77–103. Cambridge, MA: MIT Press.

Pollock, J. L. (1995). *Cognitive Carpentry: A Blueprint for Building a Person.* Cambridge, MA: MIT Press.

Preston, B. (1998a). Cognition and tool use. *Mind & Language 13*, 513–547.

Preston, B. (1998b). Why is a wing like a spoon? a pluralist theory of functions. *Journal of Philosophy 95*, 215–254.

Preston, B. (2000). The functions of things: A philosophical perspective on material culture. In P. M. Graves-Brown (Ed.), *Matter, Materiality and Modern Culture*, pp. 22–49. London: Routledge.

Preston, B. (2003). Of marigold beer: A reply to Vermaas and Houkes. *British Journal for the Philosophy of Science 54*, 601–612.

Preston, B. (2006). The case of the recalcitrant prototype. In A. Costall and O. Dreier (Eds.), *Doing Things with Things: The Design and Use of Everyday Objects*, pp. 15–27. Aldershot: Ashgate.

Putnam, H. (1975). The meaning of 'meaning'. In H. Putnam (Ed.), *Mind, Language and Reality*, pp. 215–271. Cambridge: Cambridge University Press.

Radder, H. (Ed.) (2003). *The Philosophy of Scientific Experimentation*, Pittsburgh. University of Pittsburgh Press.

Rea, M. (Ed.) (1997). *Material Constitution: A Reader*, Lanham, MD. Rowman and Littlefield.

Ridder, J. de (2006). Mechanistic artefact explanation. *Studies in History and Philosophy of Science 37*, 81–96.

Romano, G. (2009). *Thoughtful Things: An Investigation in the Descriptive Epistemology of Artifacts.* Ph. D. thesis, Eindhoven University of Technology.

Roozenburg, N. F. M. and J. Eekels (1995). *Product Design: Fundamentals and Methods.* Chichester: John Wiley & Sons.

Rosenbaum, D. A. (1991). *Human Motor Control.* San Diego: Academic Press.

Rosenman, M. A. and J. S. Gero (1998). Purpose and function in design: From the socio-cultural to the techno-physical. *Design Studies 19*, 161–186.

Ryle, G. (1949). *The Concept of Mind.* London: Hutchinson.

Scheele, M. (2005). *The Social Aspects of Artefact Use*. Ph. D. thesis, Delft University of Technology.

Scheele, M. (2006). Function and use of technical artefacts: The social conditions of function ascription. *Studies in History and Philosophy of Science 37*, 23–36.

Schwartz, S. P. (1978). Putnam on artifacts. *Philosophical Review 87*, 566–574.

Searle, J. R. (1995). *The Construction of Social Reality*. New Haven: Free Press.

Smith, E. E. and D. L. Medin (1981). *Categories and Concepts*. Cambridge, MA: Harvard University Press.

Sorabji, R. (1964). Function. *Philosophical Quarterly 14*, 289–302.

Sperber, D. (2007). Seedles grapes: Nature and culture. In E. Margolis and S. Laurence (Eds.), *Creations of the Mind: Theories of Artifacts and Their Representation*, pp. 124–137. Oxford: Oxford University Press.

Stone, R. B. and K. L. Wood (2000). Development of a Functional Basis for design. *Journal of Mechanical Design 122*, 359–370.

Suchman, L. A. (1987). *Plans and Situated Action*. Cambridge: Cambridge University Press.

Thomasson, A. (1999). *Fiction and Metaphysics*. Cambridge: Cambridge University Press.

Thomasson, A. (2007a). Artifacts and human concepts. In E. Margolis and S. Laurence (Eds.), *Creations of the Mind: Theories of Artifacts and Their Representation*, pp. 52–73. Oxford: Oxford University Press.

Thomasson, A. (2007b). *Ordinary Objects*. Oxford: Oxford University Press.

Vaesen, K. (2006). How norms in technology ought to be interpreted. *Technè 10*(1), 117–133. http://scholar.lib.vt.edu/ejournals/SPT.

Van Inwagen, P. (1990). *Material Beings*. Ithaca, NY: Cornell University Press.

Verhoeven, J. D., A. H. Pendray, and W. E. Dauksch (1998). The key role of impurities in ancient Damascus steel blades. *JOM 50*(9), 58–64.

Vermaas, P. E. (2002). Technological innovation as an unusual and non-biological evolutionary process. *Studies in History and Philosophy of Modern Physics 33*, 735–739.

Vermaas, P. E. (2006). The physical connection: Engineering function ascriptions to technical artefacts and their components. *Studies in History and Philosophy of Science 37*, 62–75.

Vermaas, P. E. (2009a). The flexible meaning of function in engineering. In *eProceedings of the 17th International Conference on Engineering Design, Stanford, California, USA, August 24-27, 2009*, pp. 2.113–2.124. Design Society.

Vermaas, P. E. (2009b). On unification: Taking technical functions as objective (and biological functions as subjective). In U. Krohs and P. Kroes (Eds.), *Functions in Biological and Artificial Worlds: Comparative Philosophical Perspectives*, Vienna Series in Theoretical Biology, pp. 69–87. Cambridge, MA: MIT Press.

Vermaas, P. E. and K. Dorst (2007). On the conceptual framework of John Gero's FBS-model and the prescriptive aims of design methodology. *Design Studies 28*, 133–157.

Vermaas, P. E. and W. Houkes (2003). Ascribing functions to technical artefacts: A challenge to etiological accounts of functions. *British Journal for the Philosophy of Science 54*, 261–289.

Vermaas, P. E. and W. Houkes (2006a). Technical functions: A drawbridge between the intentional and structural natures of technical artefacts. *Studies in History and Philosophy of Science 37*, 5–18.

Vermaas, P. E. and W. Houkes (2006b). Use plans and artefact functions: An intentionalist approach to artefacts and their use. In A. Costall and O. Dreier (Eds.), *Doing Things with Things: The Design and Use of Everyday Objects*, pp. 29–48. Aldershot: Ashgate.

Walsh, D. M. and A. Ariew (1996). A taxonomy of functions. *Canadian Journal of Philosophy 26*, 493–514.

Wiggins, D. (2001). *Sameness and Substance Renewed.* Cambridge: Cambridge University Press.

Williams, B. (1981). Internal and external reasons. In B. Williams (Ed.), *Moral Luck*, pp. 101–113. Cambridge: Cambridge University Press.

Wimsatt, W. C. (1972). Teleology and the logical structure of function statements. *Studies in History and Philosophy of Science 3*, 1–80.

Wright, L. (1973). Functions. *Philosophical Review 82*, 139–168.

Ziman, J. (Ed.) (2000). *Technological Innovation as an Evolutionary Process*, Cambridge. Cambridge University Press.

Index