

Techné:

Research in Philosophy and Technology

Special Issue

Artefacts in
Analytic Metaphysics

Wybo Houkes and Pieter Vermaas,
Special Issue Editors

Techné: Research in Philosophy and Technology

Editor-in-Chief
Editors
of Twente

Joseph C. Pitt, Virginia Tech
Peter-Paul Verbeek, University

Pieter Vermaas, Delft University

of Technology
Book Review Editor
Managing Editor

Tom Staley, Virginia Tech
Ashley Shew, Virginia Tech

Special Issue Editors Wybo Houkes, Eindhoven Univ. of Technology
Pieter Vermaas, Delft University of Technology

Special Issue: Artefacts in Analytic Metaphysics

CONTENTS

WYBO HOUKES and PIETER E. VERMAAS, Artefacts in Analytic Metaphysics: Introduction	74
LYNNE RUDDER BAKER, The Metaphysics of Malfunction	82
MARZIA SOAVI, Antirealism and Artefact Kinds	93
MASSIMILIANO CARRARA, Relative Identity and the Number of Artifacts	108
WYBO HOUKES and PIETER E. VERMAAS, Produced to Use: Combining Two Key Intuitions on the Nature of Artefacts	123
PAWEL GARBACZ, What is an Artefact Design?	137
ULRICH KROHS, Structure and Coherence of Two-Model- Descriptions of Technical Artefacts	150
BOOK REVIEWS	162

Artefacts in Analytic Metaphysics: Introduction

Wybo Houkes
Department of Philosophy and Ethics of Technology,
Eindhoven University of Technology
and
Pieter E. Vermaas
Department of Philosophy,
Delft University of Technology

Keywords: artefacts; analytic philosophy; metaphysics

Artefacts increasingly become the subject of philosophical attention. In our field of philosophy of technology, they obviously already held centre stage, most notably in, for instance, the work of Don Ihde (1990), of Peter-Paul Verbeek (2005) and in the Delft *Dual Nature of Technical Artifacts* research program (Kroes and Meijers 2002, 2006). But outside of our field artefacts have also become a topic of analysis, as is witnessed in a series of recent publications. Research on artefacts is arguably suitable for cross-disciplinary research, since artefacts play a role in technology but also in, say, biology, psychology, cognitive science and architecture. Yet, some of that recent work seems to be conducted in relative isolation of the analysis of artefacts in the philosophy of technology, a situation which calls for establishing exchange and interaction between our field and the other fields involved. This development can also be witnessed in recent publications, and this special issue is another contribution to this exchange and interaction. Lewens (2004), for instance, wrote a monograph on the artefact model in the philosophy of biology, and this will be followed up with an edited volume (Krohs and Kroes 2009) in which analyses of functions of both biological items and artefacts are contrasted and integrated. Comparably cross-disciplinary volumes have been published on artefacts in the philosophy of psychology and technology (Costall and Dreier 2006) and in the philosophy of engineering and architecture (Vermaas et al. 2008). The importance of scientific instruments and experimentation for epistemology and the philosophy of science has been scrutinised in, e.g., Radder (2003) and Baird (2004).

Another subdiscipline of philosophy in which work on artefacts has appeared is analytic metaphysics. In this subdiscipline, traditional metaphysical inquiries into the nature, constitution and categorisation of reality are made by using the methods of analytic philosophy, such as formalisation and conceptual analysis. Philosophers have discussed the nature and categorisation of artefacts (Elder 2004; Baker 2007; Thomasson 2007b) and a volume has been published in which artefacts are approached from the perspective of metaphysics and cognitive science (Margolis and Laurence 2007). In this work, philosophy of technology seems not to play a detectable role. With the collection of papers we present in this special issue, we aim to strengthen artefacts as a topic for philosophical research. In particular, we want to start a cross-disciplinary exchange and interaction between philosophy of technology and analytic metaphysics. In this introduction, we set the stage for this exchange. We first present the way in which artefacts have typically been studied in analytic philosophy. Then, we sketch some

promising, very recent developments regarding the philosophy of artefacts. Finally, we give an overview of the papers in this issue.

State of the artefacts

To describe the background against which many papers in this issue have been written, we briefly take stock of the traditional situation of artefacts in analytic philosophy, in particular metaphysics. For this purpose, we distinguish two perspectives that have shaped most existing work on artefacts.

On the one hand, artefacts may be considered in a “detached” way. This does not mean that they are analysed as if they were completely independent from human interests. Rather, artefacts are compared with objects that are independent from human interests or it is examined whether artefacts are sufficiently independent to qualify as objects or as members of a natural kind. Many of the resulting issues belong to ontology or metaphysics, such as questions concerning the persistence conditions or the (relative) identity of artefacts.

On the other hand, artefacts may be regarded as means to human ends or as playing more intricate roles in human existence. We continually use, adapt, or even design artefacts for all kinds of purposes, and most of our knowledge about artefacts stems from and is applicable for practical purposes. Conversely, artefacts shape our everyday life and concerns, not only by enabling actions that are otherwise impossible, but also by influencing our choices, lifestyles and worldviews. These involvements with artefacts are of central importance in all these accounts, which may therefore be also labelled as “involved”. From a (sub-)disciplinary perspective, “involved” analyses of artefacts may raise epistemological concerns such as the justification of function ascriptions to artefacts. More broadly conceived, they also encompass action-theoretical analyses of artefact use and design and other attempts to arrive at what might be called a phenomenology of everyday life.

Arguably, the detached and involved perspectives on artefacts are intimately related. In analytic philosophy, however, they have been carefully distinguished. This distinction is part and parcel of the traditional focus of the few studies that pay attention to artefacts. Those studies share three prominent features:

1. Metaphysical dominance. Artefacts feature in analytic metaphysics, but hardly anywhere else – explaining why this special issue focuses on metaphysics rather than epistemology or action theory. An agenda-setting example of the metaphysical dominance is Van Inwagen’s (1990, ch. 13) “Denial Thesis” concerning artefacts and other composite, non-living material objects. This thesis concerns the existence of artefacts as material objects apart from their constitutive atoms – a concern that is immediately recognisable as ontological. By association, the “detached” metaphysical perspective is shared by the various responses to Van Inwagen’s argument: both the existence question and the concepts used to answer it are the same, although the answer is different. Even the intuitions of many metaphysicians appear to have been shaped by the Denial Thesis.¹ For example, Crawford Elder observes, without apparent irony, that many of his readers may find it hard to believe that a desk exists in addition to the pieces of wood out of which a carpenter fashions it (2004, pp. 131-132).

Furthermore, the dominance of metaphysical studies partly explains why the involved perspective on artefacts is typically ignored. The concerns of metaphysics appear to require a detached perspective: our involvement with objects is supposed to be irrelevant to their “real” nature; if it

is not, this reflects negatively on their metaphysical status. The metaphysical realism embraced by many analytic philosophers after the slow demise of logical empiricism is based on the assumption that “real” objects exist, have properties and can be classified independently of our experience and knowledge. Inverted, this assumption says that objects that do not show this independence are not real. Thus, if artefacts cannot be studied from a detached perspective, they are of no concern to metaphysics.

2. *Non-specificity.* Work on artefacts in analytic metaphysics is seldom specific. Efforts to analyse artefacts are typically a small part of much more encompassing philosophical projects, from David Wiggins’s (2001) plea for absolute identity to Lynne Rudder Baker’s (2000) constitution view. Furthermore, these efforts focus on one amorphous super-category of artefacts represented by a few paradigm cases of chairs, ships, clocks, statues and screwdrivers. It seems that only those philosophers who aim at a very complete and/or a very general understanding of the world care, at some point in their projects, to examine artefacts.

Non-specificity is not the same as inaccuracy. Perhaps artefacts are analysed correctly, as a first approximation or even ultimately, in the context of one or more larger metaphysical projects. However, one of the points raised in this issue is that more attention for specific philosophical details regarding artefacts – such as analyses of their use and design – would not only increase our understanding of artefacts, but could also contribute significantly to more encompassing projects in analytic metaphysics. Moreover, there are more artefacts than chairs, ships, clocks, statues and screwdrivers, and considering their differences may lead to valuable distinctions.

3. *Function focus.* Many philosophers who have studied artefacts characterise them as primarily *functional* objects. This “function focus” takes two different forms. One continues the ontological line of inquiry by defending the claim that functions are the essences of artefacts (e.g., Kornblith 1980; Wiggins 2001) – usually, but not necessarily,² combined with the claim that this essence is nominal rather than real. Similarly, some authors who discuss artefacts in the context of more general metaphysical issues, appeal to functions when determining the persistence conditions of artefacts (Baker 2000; 2004). Another form of function focus is found, outside of metaphysics narrowly defined, in general analyses of the notion of function (e.g., Cummins 1975; Millikan 1984; Neander 1991; Preston 1998). Such analyses are usually motivated by the problems regarding apparently teleological language in biology, but some claim to analyse functional discourse in any domain whatsoever, including that of artefacts. Typically, the application to artefacts of such general function theories is taken to be relatively unproblematic, and little effort is made to defend these applications, let alone to adapt them to any specific features of artefact functions. This again illustrates the non-specificity of existing analyses of artefacts.

But there are exceptions. Beth Preston (1998) develops her general, pluralistic function theory partly on the basis of a detailed consideration of artefact use and design. In the philosophy of technology especially Peter Kroes and Anthonie Meijers (2002; 2006) advanced a research program that more principally countered non-specificity by taking an empirical turn (2000) and analysing technical artefacts within technology. In this program, called *The Dual Nature of Technical Artifacts*, artefacts were explicitly taken as “(i) designed physical structures, which realize (ii) functions, which refer to human intentionally”, thus also taking artefacts as functional objects and relating them explicitly to designing. This research program aimed at understanding artefacts as “hybrid” objects that can only be described adequately in a way that somehow combines the physical and intentional conceptualisations of the world.” (2006, p. 2) As part to the results of this program we proposed, in line with the “involved” perspective, a framework for analysing justified function ascriptions to artefacts, where this framework explicitly includes an

action-theoretical description of artefact use and designing (Houkes and Vermaas 2004; Vermaas and Houkes 2006).

Recent developments

The three features described in the previous section characterise most of the existing accounts of artefacts in analytic philosophy. But there are signs that the situation is changing. Both in philosophy and in closely related, more empirical disciplines, recent work diverges from tradition.

In analytic metaphysics, recent work does not only show an increasing attention for artefacts, but also a shifting away from the situation described above. This does not constitute a radical break with the questions and notions used in this discipline. However, there is a gradual admixture of notions that are particular to artefacts and to the “involved” perspective.

One line of work that is quickly gaining prominence concerns the defence of artefacts as “mind-dependent” objects (Baker 2004; Thomasson 2003; 2006), and the discussion that ensues from this defence. This “artefact-apologetic” work questions the central assumption of metaphysical realism, that real objects exist, persist, and can be classified independently of human experience and knowledge. Specifically, it focuses on the way in which artefacts and their classification depend on human intentions, without automatically taking this dependence as a metaphysical deficiency. The main reason for this leniency is that artefacts are indispensable in everyday life.³ Moreover, the way in which they both make sense of and defend the metaphysical status of artefacts is by appealing to and analysing human attitudes and activities.

This transition is controversial and far from complete. Certainly not all recent metaphysical work on artefacts emphasises their mind-dependence. Elder (2004), for instance, develops a metaphysics of what he calls “copied kinds” – comprising both biological items and artefacts. These kinds are characterised by a common shape, a proper function and a set of normal circumstances, not by any type of dependence on mental states. Thus, Elder’s (2004, p. 140) description of the “nature of the copying process” for household screwdrivers scrupulously avoids all references to activities such as designing, manufacturing, or using. He even goes so far as claiming that “the essential properties that [the artisan’s] product will inherit stem from a history of function and of copying that began well before the artisan undertook his work. This history reaches forward through the artisan’s motions – it shapes his shaping.” (Elder 2004, p. 142).

Even more recent papers by Thomasson (2006; 2007a) and Elder (2006; 2007) show some of the problems and promises in emphasising the mind-dependent nature of artefacts. Most interestingly, perhaps, it shows how a metaphysics of everyday objects, like artefacts, should not and need not “[borrow] an idea suitable for realism about natural objects” (Thomasson 2007a, p. 72). Books by Baker (2007) and Thomasson (2007b) further explore how a metaphysics that is specific to artefacts may be constructed.

Their emphasis on mind-dependence brings to light interesting connections between the metaphysics of artefacts and some slightly older work in philosophy, as well as recent empirical studies on artefact representation and categorisation.

Existing definitions of the notion of “artefact”, or proposals to distinguish conceptually various types of artefacts, typically appeal to human intentions or activities – even though such

definitions are few and far between. To cite some of the more well-known attempts, an artefact is “an intentionally modified tool whose modified properties were intended by the agent to be recognised by an agent at a later time as having been intentionally altered for that, or some other purpose” (Dipert 1993, pp. 29-30); “An object o made by an agent Ag is an artifact only if it satisfies some type-description D included in the intention I_A which brings about the existence of o ” (Hilpinen 1992); or an artefact is “[a]ny object produced to design by skilled action” (Simons 1995). By the central place of intentional actions such as design, production and modification, all definitions appear to be constructed from the involved perspective. And of those who proposed definitions, Randall Dipert has developed a more encompassing analysis of artefacts that combines action-theoretical, epistemological and ontological elements.

Those who seek a more specific, more “involved” metaphysics of artefacts may not just look to existing definitions for support, but also to recent empirical studies. In the last decade, the representation and categorisation of artefacts have become a topic of considerable interest in, for instance, cognitive psychology. Much of this work is aimed at testing and developing general theories of concept formation, but specific experiments have been performed for the case of artefacts. The hypotheses tested in these experiments show some confluence with philosophical work on artefacts, in that the experimental hypotheses typically share the function focus described in the previous section. Some researchers, most notably Paul Bloom (1996; 1998), have even developed a function-essentialist view on artefact categorisation that is a psychological counterpart of the more metaphysical function focus mentioned above. More generally, this strand of empirical work has concerned the importance of recognising the intentions of authors and users in artefact categorisation – a concern that clearly conforms to the questions and notions developed in the involved perspective on artefacts.⁴ The changing focus in the metaphysics of artefacts may bring this work sufficiently close to cognitive studies to promote a fruitful interaction. A very recent example of this interaction are the papers collected in *Creations of the Mind* (Margolis and Laurence 2007).

Themes identified in this current surge of attention for artefacts include many features that are particular to them. One is the specific way in which artefacts may be said to be mind-dependent. Another is that artefacts are *used* and *designed*. Analysing these activities necessarily means taking an involved perspective on artefacts, even if one’s ultimate goal remains to determine the nature of artefacts. Despite their metaphysical goals, authors like Baker, Elder and Thomasson cannot avoid characterising design and its role in determining the function and nature of artefacts. Frequently, their characterisations appeal to designer’s (or user’s) *intentions*, both to explain mind-dependence and to show how characterising artefacts in terms of design would undermine realism about artefacts. And, finally, much current work retains the *function* focus of older work on artefacts, although there is more attention for the way in which artefact functions might differ from the functions of natural objects, and for theories of function ascriptions.

The papers making up this special issue also inquire into design, intentions, functions and the nature of artefacts. Some take a decidedly involved perspective, even on metaphysical issues that have traditionally been studied from a detached perspective. Others resist this tendency. In both ways, the papers continue and strengthen an exciting new movement in analytic philosophy: instead of treating artefacts as marginal objects, interesting only for the most encompassing metaphysical projects, they put artefacts into the centre of attention.

Overview of the contributions

In the first paper, Lynne Rudder Baker focuses on one important aspect of artefacts, namely their normativity – which manifests in the all-too-common phenomenon of artefact malfunctioning. According to Baker, malfunctioning should be regarded as an aspect of reality, and she rejects various “Deflationary” views that discard artefacts and malfunctioning simply on the basis of their mind-dependence. She offers her own Constitution View, which allows for mind-dependent objects such as artefacts, as a more adequate alternative.

The reality of artefacts and artefact kinds is also defended, on different grounds than Baker’s, by Marzia Soavi. She distinguishes various arguments – metaphysical, epistemological and semantic – that have been presented for the claim that there is a radical distinction between artefact kinds and natural kinds. After detailed scrutiny, Soavi concludes that none of these arguments is sound: they do not indicate a distinction between artefact kinds and natural kinds that is sufficiently large to support anti-realist claims regarding the former.

In the third paper, Massimiliano Carrara focuses on another aspect of artefacts that has drawn the attention of metaphysicians, namely their identity. Carrara considers a view on which, following Geach’s more general analysis, the identity of artefacts is relative to some general term. He finds wanting one type of support for this claim, based on considerations of cardinality; but he admits another that is based on the idea that “artefact” is not a sortal concept.

Wybo Houkes and Pieter Vermaas examine which limitations are set for an ontology of artefacts by the intuition that artefacts are non-natural objects. In the course of these examinations, they criticise the function focus of most existing accounts of artefacts, and they confront – and attempt to harmonise – two conceptions of artefacts: one in which they are instruments, and another in which they are intentionally produced objects. The authors conclude that, no matter the results of this confrontation, the basis for an ontology of artefacts is epistemological or action-theoretical.

Pawel Garbacz presents an account that may be described as an “ontologisation” of designing. Building upon Van Ingarden’s phenomenological work, Garbacz introduces the notion of intentional states of affairs, and analyses artefact designs in terms of this notion. The result, which accommodates a possible multiplicity of designs and a distinction between artefact types and tokens, puts an apparently epistemological notion at the heart of a metaphysics of artefacts – in a way that is both different from and markedly similar to that presented by Baker in her contribution.

In the final paper, Ulrich Krohs presents an account of technical artefacts in which they are described by means of two supplementary models – one physicalist and the other functional. Krohs argues that coherence between the two models can be provided through what he calls two-sorted theory elements, which map elements of one model on that of the other. Functions retain their central importance for artefacts, because they play this coherence-providing role, as Krohs argues and illustrates by means of an elaborate example.

Acknowledgements

Research by Wybo Houkes and research by Pieter Vermaas was supported by the Netherlands Organisation for Scientific Research (NWO).

References

Baird, D. 2004. *Thing Knowledge*, Berkeley: University of California Press.

- Baker, L.R. 2000. *Persons and Bodies*, Cambridge: Cambridge University Press.
- Baker, L.R. 2004. "The Ontology of Artefacts," *Philosophical Explorations*, 7: 1-14.
- Baker, L.R. 2007. *The Metaphysics of Everyday Life: An Essay in Practical Realism*, Cambridge: Cambridge University Press.
- Bloom, P. 1996. "Intention, History, and Artifact Concepts," *Cognition*, 60: 1-29.
- Bloom, P. 1998. "Theories of Artifact Categorization," *Cognition*, 66: 87-93.
- Costall, A. and O. Dreier, eds. 2006. *Doing Things with Things: The Design and Use of Everyday Objects*, Aldershot: Ashgate.
- Cummins, R. 1975. "Functional Analysis," *The Journal of Philosophy*, 77: 741-765.
- Dipert, R.R. 1993. *Artifacts, Art Works, and Agency*, Philadelphia, PA: Temple University Press.
- Elder, C.C. 2004. *Real Natures and Familiar Objects*, Cambridge, MA: The MIT Press.
- Elder, C.C. 2006. "Conventionalism and Realism-Imitating Counterfactuals," *Philosophical Quarterly*, 56: 2-15.
- Elder, C.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*, Oxford: Oxford University Press, 33-51.
- Gelman, S.A. and P. Bloom. 2000. "Young Children are Sensitive to How an Object was Created When Deciding How To Name It," *Cognition*, 76: 91-103.
- Hilpinen, R. 1992. "Artifacts and Works of Art," *Theoria*, 58: 58-82.
- Houkes, W. and P.E. Vermaas. 2004. "Actions versus Functions: A Plea For an Alternative Metaphysics of Artifacts," *Monist*, 87: 52-71.
- Ihde, D. 1990. *Technology and the Lifeworld*, Bloomington/Minneapolis: Indiana University Press.
- Kornblith, H. 1980. "Referring to Artifacts," *The Philosophical Review*, 89: 109-114.
- Kroes, P.A. and A.W.M. Meijers. 2000. *The Empirical Turn in the Philosophy of Technology, Research in Philosophy and Technology*, Vol. 20, 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 (an introduction to a special issue on the *Dual Nature of Technical Artifacts* program). Available [Online]: http://scholar.lib.vt.edu/ejournals/SPT/v6n2/kroes_meijers.pdf
- Kroes, P.A. and A.W.M. Meijers. 2006. "The Dual Nature of Technical Artifacts," *Studies in History and Philosophy of Science*, 37(1): 1-4. (an introduction to a special issue on results of the *Dual Nature of Technical Artifacts* program)
- Krohs, U. and P. Kroes, eds. 2009. *Functions in Biological and Artificial Worlds: Comparative Philosophical Perspectives*, Vienna Series in Theoretical Biology, Cambridge, MA: The MIT Press.
- Lewens, T. 2004. *Organisms and Artifacts: Design in Nature and Elsewhere*, Cambridge, MA: The MIT Press.
- Malt, B.C. and E.C. Johnson. 1992. "Do Artifact Concepts Have Cores?," *Journal of Memory and Language*, 31: 1-23.
- Margolis, E. and S. Laurence, eds. 2007. *Creations of the Mind: Theories of Artifacts and Their Representation*, Oxford: Oxford University Press.
- Matan, A. and S. Carey. 2001. "Developmental Changes within the Core of Artifact Concepts," *Cognition*, 78: 1-26.
- Merricks, T. 2001. *Objects and Persons*, Oxford: Clarendon.
- Millikan, R.G. 1984. *Language, Thought and Other Biological Categories*. Cambridge, MA: The MIT Press.
- Millikan, R.G. 2000. *On Clear and Confused Ideas*, Cambridge: Cambridge University Press.
- Neander, K. 1991. "The Teleological Notion of 'Function'," *Australasian Journal of Philosophy*, 69: 454-468.

- Preston, B. 1998. "Why is a Wing like a Spoon? A Pluralist Theory of Function," *The Journal of Philosophy*, 95: 215-254.
- Radder, H., ed. 2003. *The Philosophy of Scientific Experimentation*, Pittsburgh: University of Pittsburgh Press.
- Simons, P. 1995. "Artefact," in: J. Kim and E. Sosa, eds., *A Companion to Metaphysics*, Oxford: Blackwell, 33-34.
- Thomasson, A. 2003. "Realism and Human Kinds," *Philosophy and Phenomenological Research*, 67: 580-609.
- Thomasson, A. 2006. "Metaphysical Arguments against Ordinary Objects," *Philosophical Quarterly*, 56: 340-359.
- Thomasson, A. 2007a. "Artifacts and Human Concepts," in: E. Margolis and S. Laurence, eds., *Creations of the Mind: Theories of Artifacts and Their Representation*, Oxford: Oxford University Press, 52-71.
- Thomasson, A. 2007b. *Ordinary Objects*, Oxford: Oxford University Press.
- Van Inwagen, P. 1990. *Material Beings*, Ithaca: Cornell University Press.
- Verbeek, P.-P. 2005. *What Things Do: Philosophical Reflections on Technology, Agency, and Design*, University Park, PA: Penn State University Press.
- Vermaas, P.E. and W. Houkes. 2006. "Technical Functions: A Drawbridge between the Intentional and Structural Nature of Technical Artefacts," *Studies in History and Philosophy of Science*, 37: 5-18.
- Vermaas, P.E., P. Kroes, A. Light and S.A. Moore, eds. 2008. *Philosophy and Design: From Engineering to Architecture*, Dordrecht: Springer.
- Wiggins, D. 2001. *Sameness and Substance Renewed*, Cambridge: Cambridge University Press.

Endnotes

- 1 Another argument for the Denial Thesis is presented by Trenton Merricks (2001).
- 2 Ruth Garrett Millikan (2000) defends the claim that artefact kinds are functional and historical, but real.
- 3 Baker describes the deficiency assumption in traditional realist work as "bizarre" (2004, p. 14), given the enormous impact that artefacts have upon the world. Similarly, Thomasson claims that, without accepting mind-dependent or "human" kinds, it is impossible "to make sense of the human world" (2003, p. 607).
- 4 Relevant studies include those of Malt and Johnson (1992), Gelman and Bloom (2000), Matan and Carey (2001) and several contributions to Margolis and Laurence (2007).

The Metaphysics of Malfunction

Lynne Rudder Baker
 Department of Philosophy
 University of Massachusetts

Abstract

Any artefact – a hammer, a telescope, an artificial hip – may malfunction. Conceptually speaking, artefacts have an inherent normative aspect. I argue that the normativity of artefacts should be understood as part of reality, and not just “in our concepts.” I first set out Deflationary Views of artefacts, according to which there are no artefactual properties, just artefactual concepts. According to my contrasting view – the Constitution View – there are artefactual *properties* that things in the world really have. For example, there is a property of being a telephone per se; we apply our concept *telephone* to things that have that property. Things that have the property of being a *telephone* are constituted by, but not identical to, aggregates of particles. To be an artefact, an object must have an intended function, among other things. Telephones – in virtue of being the *kind* of objects that they are – are always subject to malfunction. And malfunctions, when they occur, are just as much part of the world as telephones are. The example of artefacts shows that what is in the world – what really exists – need not be “mind-independent” nor independent of our concepts.

Keywords: malfunction, artefacts, function, normativity, intention-dependent items, reductionism, eliminativism, ontological significance, constitution view, practical realism, mind-dependence

Artefacts are ubiquitous in the world that we encounter. Most broadly, artefacts include everything that is produced intentionally – paintings and sculptures as well as scissors and microscopes. Our concern here is with an important subclass of artefacts – technical artefacts, characterized by the organizers of this conference as “the material products of our endeavour to attain our practical goals.” Since goals are the sorts of things that we attain or fail to attain, a distinction between proper performance and malfunction is built into the very idea of a technical artefact. From now on, when I say ‘artefact,’ I mean ‘technical artefact.’ Any such artefact – a hammer, a telescope, an artificial hip – may malfunction.

The concepts of function and malfunction, as they apply to artefacts, are normative.¹ Artefacts have *intended* functions, which are obviously normative. To carry out an intended function is what an artefact *is supposed to* do; to fail to carry out the function in certain circumstances is a kind of error, a malfunction. Where there is room for error or mistake, there is normativity. Normativity pervades the *Lebenswelt*: There is no intention without the possibility of its being thwarted, no desire without the possibility of its being frustrated, no function without the possibility of malfunction. We simply cannot understand the world we live in without presupposing normativity. Unfortunately, like most other philosophers, I have no theory of normativity. But if we take the world as we encounter it as our starting point (as I do), then normativity is part of the price of admission. Nowhere is normativity more glaring than in the behavior of artefacts – from the trivial (people get wet when umbrellas blow inside-out) to the significant (combatants get killed when guns jam).

Elsewhere, I have argued that artefacts have ontological status: they are genuine objects in the world.² Here I want to consider the malfunction. After some general introductory remarks and a brief discussion of the notion of malfunction, I'll set out a kind of view of artefacts that I think many philosophers would find attractive – I'll call such views 'Deflationary Views.' I hope to supplant Deflationary Views with a view of artefacts, which I call 'the Constitution View' and show how the Constitution View can treat the phenomenon of malfunction. After contrasting Deflationary Views and the Constitution View, I want to turn to some metaphysical issues concerning the nature of reality and to challenge the view of many metaphysicians that there is a sharp and important distinction between what is really in the world and what is only a matter of our concepts – a distinction sometimes formulated as a distinction between what is mind-independent and what is mind-dependent. The example of artefacts shows that what is in the world – what really exists – need be neither mind-independent nor independent of our concepts.

1. The Idea of Malfunction

Artefacts, by definition, have intended functions. Anything that has an intended function is subject to malfunction. Thus, for technical artefacts, the concepts *artefact*, *function*, and *malfunction* are conceptually linked: None is intelligible without the others.

The concept of an artefact's function – along with the concept of malfunction – is one of a huge and important class of concepts that has been overlooked by philosophers. This class includes nonmental concepts that entail mental concepts – e.g., being in debt, being a driver's licence, being a delegate. Nothing can be in debt or be a driver's license in a world without beings with propositional attitudes. Being a driver's license is not itself a mental concept; it is not a concept that is applied to minds, or to things that have minds; but it is a concept that would have no application in a world without minds. I shall coin the term 'intention-dependent' or, for short, 'ID' for such concepts.

An ID concept is any concept that either is a propositional-attitude concept (like believing, desiring or intending) or entails that there are beings with beliefs, desires and/or intentions an 'ID concept,' an 'intention-dependent' concept. ID concepts are concepts whose applicability depends on intentionality. ID phenomena are phenomena that fall under ID concepts. Such phenomena include being a wedding, being a carrot peeler, being a treaty, and so on. Many, if not most, social, economic, political, and legal concepts are ID concepts. For example, the concept of writing a check is an ID concept, because there would be no such thing as writing a check in a world lacking the social and economic conventions that presuppose that people have beliefs, desires and intentions. ID concepts apply to most human activities – both individual (getting a job, going out to dinner, designing a house) and collective (manufacturing automobiles, changing the government, etc.). They could not exist or occur in a world without beliefs, desires, and intentions.

Other communities may be familiar with other kinds of ID concepts; but all communities recognize many kinds of ID concepts—as well as other ID objects like pianos and paychecks, and ID phenomena like conventions and obligations.³ ID concepts stand in contrast to nonID concepts – e.g., being a promise as opposed to an audible emission, being a signature as opposed to a mark on paper, being a dance step as opposed to a bodily motion. The audible emission, the mark on paper, the bodily motion could all exist or occur in a world lacking beings with propositional attitudes, but the promise, the signature, and the dance step could not.⁴

Indeed, many different kinds of things are ID phenomena in the sense just stipulated: events (e.g., a baseball game), objects (e.g., a passport), actions (e.g., voting), dispositions (e.g., being honest), activities (e.g., reading your mail), institutions (e.g., a national bank), medical procedures (e.g., transplanting a heart), business dealings (manufacturing new medications and marketing them) – all these are ID phenomena.⁵ Intentional language contains terms (e.g., ‘wants to buy milk,’ ‘was elected president,’ ‘paid her taxes’) whose application presupposes that there are beings with beliefs, desires, intentions. So, actions – like buying a car, sending an email, or washing the dishes – are ID events whose occurrence entails that there are beings with beliefs, desires and intentions. ID phenomena encompass a huge range of phenomena that characterize the world as we know it.

What is important about ID phenomena for our purposes is that all artefacts and their associated properties – in particular, properties of function and malfunction – are ID phenomena. Artefacts are defined by their intended functions: The function of the brakes in a car is to reduce its speed; if someone wants to slow down and applies pressure to the brake pedal and the car maintains its speed, then the brakes have malfunctioned.

Not all cases in which something fails to perform its intended function seem to be malfunctions. For centuries, people tried to build perpetual motion machines. Of course, they all failed. Should we say that each of the machines malfunctioned? Or: Suppose that someone had an amulet whose intended function was to protect its user and to cause harm to her enemies. (An amulet is a paradigm case of a technical artefact – “a material product of our endeavor to attain our practical goals.”) The amulet was supposed to produce a desired effect when its user uttered certain incantations. It is plausible to suppose that no such causal connections are physically possible. Did the amulet malfunction? There seems to be a difference between a flaw in a design in which the mechanism did not operate as expected (e.g., the designer had overlooked the fact that the gas would be under so much pressure that the device would explode when operated for more than a few seconds), and a flaw in which the mechanism operated as planned, but did not accomplish the intended function (e.g., a perpetual motion machine or the amulet).

The examples of the perpetual motion machine and the amulet raise questions about the concept of intended function. Can an artefact have a function that is it is physically impossible for it to perform? My suggestion is to take terms like ‘amulet’ and ‘perpetual motion machine’ to mean, respectively, ‘item intended to protect its user and to harm her enemies’ and ‘machine intended to produce perpetual motion.’ Then, we can say that there are such artefacts, and that they have functions that it is physically impossible for them to perform. But I would reserve the term ‘malfunction’ for artefacts that have functions that are physically possible to be performed. Hence, the failure of a perpetual motion machine to produce perpetual motion and the failure of the amulet to cause mishaps should not count as malfunctions.

Other cases of failure to perform the intended function that should not be considered to be malfunctions include these: A car that does not start because it is out of gas. (A car is not intended to run in conditions in which it lacks gas.) A computer that does not operate because its operator is incompetent (say a two-year-old). In general, failure to perform an intended function is not a malfunction unless there is an attempt by a competent operator to perform the intended function in conditions for which the artefact was designed. So, here is an initial stab at a pretheoretical characterization for an occurrence to be a malfunction:

- (M) x is a malfunction of an artefact a if and only if:
- (a) x is a failure to perform the intended function of a , where the intended function of a is such that it is physically possible to be performed, and
 - (b) x occurs when a competent operator tries to use a to perform its intended function under conditions for which a was designed.

There are a variety of sources of malfunction: The materials used may be poorly chosen (as when soft metal is used in the manufacture of a key); the materials may themselves be defective (as when too much sand is used in mortar holding up the bricks on the UMass library); or the design may be defective (as when gas tanks in Pintos explode on impact); or there may be damage to the structure (as when the surface of the space shuttle Columbia was punctured during take-off). Although there is much more to be said about the concept of malfunction, let us move on to the theories.

2. Deflationary Views of Artefacts

I made up the term ‘Deflationary Views,’ and I am not wedded to it; but I want a label for some views associated with most prominent metaphysicians today.⁶ What the disparate philosophers that I take to be proponents of ‘Deflationary Views’ have in common is that they hold, roughly, that, ontologically, there is no more to being an artefact (as opposed to being a collection or particles) than our talk about artefacts. Such philosophers hold that there is a sharp distinction between our concepts, our language, our interests, on the one hand, and what really exists on the other. Not only do such philosophers suppose that can we study each side of the divide independently of the other, but they also suppose that the business of metaphysics is exclusively on the side of what exists independently of our concepts, our language, our interests.

Let me illustrate two versions of this view by considering an actual event. On February 1, 2003, the space shuttle, Columbia, broke up during a seemingly routine reentry into the Earth’s atmosphere. It was a spectacular disaster, leaving myriad pieces from the shuttle scattered over several U.S. states. (It was later determined that the malfunction was caused by damage to the left wing during launch; during the flight of the space shuttle, the damage had seemed slight.) How might Deflationary-Viewers interpret this event? Here are two versions of Deflationary Views:

(1) Eliminativism: Strictly speaking, no space shuttle ever existed: the words ‘space shuttle’ do not refer. All that existed were simples arranged space-shuttle-wise; there is no object that is a space shuttle. Sentences like ‘The space shuttle broke up’ are rephrased to eliminate the apparent reference to an object. When speaking in the “strict and philosophical sense,” we may mention simples-arranged-space-shuttle-wise, instead of space shuttles. When the space shuttle broke up (as we say), the only change in reality was in the arrangement of certain simples. But nothing went out of existence. I associate this view with Peter van Inwagen, according to whom the only (finite, concrete) objects that exist are simples and living organisms.⁷ There exist no artefacts, though we can find true paraphrases of sentences putatively about artefacts: For ‘This is the house that Jack built,’ we may substitute ‘These are simples that were arranged housewise by Jack.’

(2) Reductionism: There are space shuttles; the words ‘space shuttle’ do refer, but what they refer to are aggregates of matter that occupy spacetime points arranged space-shuttle-wise. The Columbia was nothing more or less than a mereological sum of bits of matter at those spacetime points. Indeed, every aggregate of matter-filled spacetime points have mereological sums; we have names (e.g., ‘space shuttle’) for a few of the sums that exist, but no names for most of the

sums. (Indeed, we couldn't possibly name them all; there's a nondenumerable infinity of objects.) The only concrete objects that really exist are bits of matter at spacetime points and their sums arranged in various ways. I associate this view with David Lewis.⁸

Ontologically, the eliminativist and reductionist views are alike with respect to artefacts. On both views, strictly speaking, nothing literally went out of existence when the space shuttle broke up; there was only a change in the arrangement of particles (or simples – from now on, I'll use 'particles' as the all-purpose term). There was no change in what exists – it just became inappropriate to apply our concept of 'space shuttle' to the particles in their new arrangement. All the objects [or, in the case of van Inwagen, nonliving objects] that exist, according to both views, are particles (or simples) arranged in certain ways. On both the eliminativist and reductionist views, there is no *ontological* difference between the space shuttle and the little pieces scattered over several U.S. states. When the particles are arranged in a certain way (space-shuttle-wise), we call them a 'space shuttle,' but nothing actually went out of existence when the space shuttle broke up.

Both reductionists and (some) eliminativists take the sentence, 'The space shuttle broke up' to be true.⁹ The eliminativist takes that sentence to have a paraphrase that does not mention space shuttles: 'There are some simples arranged space-shuttle-wise at one time, and not arranged space-shuttle-wise at a later time.'¹⁰ The paraphrase (putatively) expresses what we want to say in using the original sentence, but without seeming to refer to space-shuttles. The reductionist does not need a paraphrase that avoids mention of space shuttles. Unlike the eliminativist, the reductionist holds that there are space shuttles, but what a space shuttle *is* is just an arrangement of particles.¹¹ The semantic difference is that the reductionist takes 'space shuttle' to be a referring word (that refers to a certain mereological sum of particles), but the eliminativist does not take 'space shuttle' to refer to anything (because, on the eliminativist's view, those particles have no mereological sum). But the aggregate of particles – which the reductionist says really is a space shuttle and the eliminativist says really is no thing – is the same in both cases.

That is, the reductionist and eliminativist agree that what actually exists is only the particles arranged in a certain way. The difference between them is only whether they consider such an arrangement of particles to be an entity (i.e., to have a mereological sum, as they would put it). If we take seriously Lewis's comment that mereology is "ontologically innocent,"¹² – i.e., that mereological sums do not introduce new objects over and above their parts – then it seems that the difference between reductionism and eliminativism is not ontological, but purely semantic.

In any case, neither eliminativism nor reductionism can take discourse about artefacts at face value. The eliminativist cannot suppose that the sentence 'the space shuttle broke up' is both true and literally expresses the proposition that the space shuttle broke up. For the eliminativist, common nouns in everyday discourse disappear under analysis. So, eliminativists cannot take discourse about artefacts at face value. The reductionist, on the other hand, does suppose that our talk about space shuttles really is about space shuttles, but takes talk about space shuttles to be just talk about aggregates of particles. However, if talk about the malfunction of Columbia were just talk about re-arrangement of particles, then certain rearrangements of particles should suffice for a malfunction. But there is nothing about any arrangement of particles independently of our concepts and interests that makes it the case that the space shuttle malfunctioned. It is only in virtue of our concepts and interests that the dispersal of particles (say) is a malfunction. So, reductionists cannot take statements like "An object went out of existence when Columbia was destroyed," at face value any more than eliminativists can. Literally, on Deflationary Views,

when Columbia was destroyed, no object went out of existence. The upshot is that neither eliminativism nor reductionism takes our discourse about artefacts at face value.

According to the Deflationary Views, there is nothing in reality that makes an ontological difference between a hammer and a pillow – or, for that matter, between a hammer and an aggregate of your left eyeball and my right shoe. All are just aggregates of particles, to some of which we apply our artefactual (and other) concepts. (Again: according to the reductionist, the aggregate itself is an entity; according to the eliminativist, the aggregate is not an entity. In both cases, there is no more to things that apparently exist than the existence of particles.) According to these views, something is a hammer in virtue of the fact that we apply our concept ‘hammer’ to certain aggregates of particles. A malfunction of a hammer – say, its head flies off its handle – is likewise just a change in arrangement of the particles. The normativity of artefacts, on the Deflationary Views, is wholly in our language or concepts, and not in the world at all. Function and malfunction are a product of our concepts; what are in the world are just aggregates of particles that could exist in worlds that lack our concepts. The laws of physics apply equally to machines that function properly and to machines that malfunction. So, on the Deflationary Views, malfunction is wholly a matter of our language; it is not to be found in the world. What happened to the space shuttle Columbia has no ontological significance whatever.

Indeed, strictly speaking, on the Deflationary Views, there is no metaphysics of artefacts, and no metaphysics of malfunction. As Peter van Inwagen remarked, if we confine our discussion to a canonical language that “refers to nothing besides simples and living organisms and abstract objects,” – the only objects that van Inwagen countenances – “we shall be able to formulate no *philosophical* questions about the identities of artifacts at all.”¹³ The activities of engineers are of no philosophical interest. If what I’ve called ‘Deflationary Views’ are correct, then the expression ‘metaphysics of malfunction’ is simply an oxymoron.

3. The Constitution View of Artefacts

I want to propose an alternative, according to which the destruction of the space shuttle Columbia does have ontological significance: What happened when Columbia broke up was that something went out of existence, not just that particles changed arrangements. On my alternative – I call it the ‘Constitution View’ – all macrophysical objects are constituted, ultimately, by aggregates of particles; but macrophysical objects are not identical to their constituters.¹⁴

According to the Constitution View, reality comes in fundamentally different kinds. Each existing thing is of a primary kind. An entity’s primary kind is given by the answer to the Aristotelian question: What is x most fundamentally? There is no “mere thing” behind or underlying the instance of a primary kind. Entities are of their primary kinds essentially: an entity cannot survive loss of its primary-kind property. Entities of different primary kinds have different causal powers as well as different persistence conditions. Constitution is a relation between things of different primary kinds.

Primary kinds include not only kinds determined by structure or by material constituent, or by underlying essence; but also there are primary kinds determined by function. Underlying the Constitution View is the idea that what something is most fundamentally is often determined by what it can do – its abilities and capacities – rather than by what it is made of. This is obvious in the case of artefacts: What makes something a clock is its function of telling time, no matter what it is made of.

Consider a hammer, constituted by an aggregate consisting of two pieces of wood (one for the handle, one for the wedge) and a piece of steel for the head. When the pieces of wood and steel in the aggregate are in hammer-favorable circumstances (including the right shapes and the intention to be used for pounding), the aggregate comes to constitute a hammer. The primary kind of the constituting aggregate is *wood/steel*; the primary kind of the artefact is *hammer*. The constituting aggregate is itself constituted by more fine-grained aggregates, down all the way to aggregates of sub-atomic particles.

The hammer has all kinds of properties – some nonderivatively (because it is a hammer) and others derivatively (because it is constituted by the wood/steel aggregate. E.g., it has the property of being worth 20 Euros nonderivatively, but of weighing a half a kilogram derivatively. The aggregate weighs half a kilogram nonderivatively (because it would weigh half a kilogram whether it constituted anything or not; the weight of the particles adds up to half a kilogram), and is worth 20 Euros derivatively (because its worth is determined by the fact that it constitutes a hammer). Properties that may be had derivatively are shared by both the constituter (the pieces of wood and steel) and the constituted thing (the hammer).

What kinds of materials are suitable for various kinds of artefacts is an engineering question, not a philosophical one. But wrong choice of material may be a source of malfunction. E.g., using a soft material like rubber for the head of a hammer intended to be used on a hard material like stone will destroy the head and render the hammer unable to perform its function. The hammer does not cease to exist when the rubber head deteriorates. It just malfunctions, but there is still an ‘it’ that has an intended function – perhaps never to be carried out again.

What exactly is the line, someone may ask, between having a hammer that is broken, and having something that is not a hammer at all? There is no sharp line. In the absence of a clear boundary between a malfunctioning F and a nonF, one may either take a Deflationary View or acknowledge that there is vagueness in reality. Elsewhere, I take and defend the latter position: there is vagueness in reality. I believe that recognition of ontic vagueness is required for a realistic view of the special sciences. Indeed, every science that recognizes things that evolve – things like species in biology and solar systems in astronomy – assumes that there is vagueness in reality. I cannot argue for this position here. I just want to acknowledge this consequence of the Constitution View.

Now apply the Constitution View to the example of the space shuttle Columbia. The malfunction in the space-shuttle case put an end to the existence of Columbia. But according to the Constitution View, Columbia really existed in its own right, so to speak. It was constituted by a vast aggregate of a complex primary kind, which itself was constituted by further aggregates, until finally there is a constituting aggregate of subatomic particles.¹⁵ Let *P* be an aggregate that is a subatomic constituter of Columbia at *t*. Columbia was essentially a space shuttle; *P* was only derivatively a space shuttle at *t* – while *P* constituted Columbia. Recall that an aggregate exists as long as the items in it exist, no matter where they are. We cannot say, “*P* is identical with Columbia at *t*.” We cannot say this, because we are assuming classical identity and three-dimensionalism: identity is necessary identity, not relative to time; and on three-dimensionalism, ‘Columbia at *t*’ does not denote an entity, but an ordered pair $\langle \text{Columbia}, t \rangle$. So, although *P* constituted Columbia at *t*, *P* was not identical with Columbia – at *t* or any other time.

According to the Constitution View, it is not just that we found it convenient to stop referring to *P* as ‘Columbia’ (à la Lewis). It is rather that Columbia went out of existence altogether, but *P* did not. Nor is it just that there was no such entity as Columbia at all (à la van Inwagen). By

contrast, on the Constitution View, the break-up of Columbia was a loss to reality, ontologically speaking. It is rather that there was an entity Columbia and there was an aggregate, *P*, and at the break-up, the former ceased to exist but the latter did not. The change was more than a change in the arrangement of particles. The contents of the world changed when Columbia was destroyed; complete inventories of the world would include different objects before and after the break-up.

The Constitution View, in contrast to the Deflationary Views, allows us to be realists about artefacts: Artefacts exist in their own right. Since part of what it is to be an artefact is to have an intended function, artefacts are always liable to malfunction. Proponents of Deflationary Views can allow that statements about malfunction – e.g., ‘The space shuttle malfunctioned’ – are true. But they cannot take the sentence at face value to state what it seems to state. On a Deflationary View, such a statement is either about a change in arrangement of particles, or about no thing at all. The normativity drains away. By contrast, the Constitution View easily accepts the characterization of malfunction on its face-value interpretation, without having to reinterpret it (as van Inwagen does) or to suppose that talk about malfunction is really just talk about concepts (as Lewis does).

4. Practical Realism

Attention to artefacts, I think, will shed light on an old metaphysical issue – namely, realism. Many philosophers take realism to depend on a distinction between what is mind-independent and what is mind-dependent, where they think of quarks, rocks and stars as mind-independent and of after-images, raw feels and thoughts as mind-dependent. (They usually do not think of the ID phenomena that I discussed at the outset at all.) The Constitution View is a challenge to this way of understanding realism.

This distinction between what is mind-independent and what is mind-dependent is coherent, but I believe that its philosophical significance has been vastly overrated. In particular, it does not demarcate what is genuinely real.¹⁶ If it did, then artefacts would be found wanting. Yet, many philosophers who consider themselves to be realists take the distinction between mind-independence and mind-dependence to be the foundation of their view. For example, Ernest Sosa has reported:

What the metaphysical realist is committed to holding is that there is an in-itself reality independent of our minds and even of our existence, and that we can talk about such reality and its constituents by virtue of correspondence relations between our language (and/or our minds), on the one hand, and things-in-themselves and their intrinsic properties (including their relations), on the other. (Sosa 1993, p. 609)

I suspect that ‘mind-independent’ is an example of what J.L. Austin called a ‘trouser word.’ It wears the pants in the family, and ‘mind-dependent’ must be defined in terms of it – as what is not mind-independent.

All ID phenomena are thus mind-dependent by definition, and as we have seen, all artefacts are ID objects: they are not mental items, but they can not exist in a world without minds. Artefacts are not part of in-itself reality independent of our minds and even of our existence. Nothing would be a carburetor in a world without intentional activity.¹⁷ So restricting reality to what is mind-independent will not only eliminate from reality everything that depends on language, but also all artefacts.

A distinction between mind-independence and mind-dependence puts carburetors and dreams, statues and imaginings, and other subjective phenomena on the same side of the ontological divide. I am confident that it is basically wrong-headed to put artefacts and after-images in the same ontological category, and hence I am also confident that the mind-independence/mind-dependence distinction is itself misguided as a basis for metaphysics.

To reject the mind-independence/mind-dependence distinction as the basis of metaphysics is to reject the idea that there is a sharp division between language and “the world.”¹⁸ But, of course, language is not isolatable from the world.¹⁹ The world as we know it is infected with language through and through. The significance of discarding the mind-independence/mind-dependence distinction is this: What exists in reality need not be wholly independent of language. The world as encountered is full of examples. To take one example almost at random: The existence of credit cards depends on social and economic practices that require language, and *de re* features of credit cards inherit that dependence on language.

By rejecting the mind-independence/mind-dependence distinction as a constraint on theorizing, a practical realist opens the door to an integration of metaphysical and epistemological approaches to artefacts. We no longer have to seal off metaphysics from “contaminants” like what we already know from scientific, engineering, or even commonsense sources. The fact that artifacts are intention-dependent in no way counts against their being objects of metaphysical inquiry.

Metaphysical realists standardly think of reality in terms of mind-independence. As I have emphasized, I do not. Hence, I do not call myself a metaphysical realist, but a practical realist: “Realist” because I believe that there may exist objects and properties beyond our ability to recognize them; “practical” because I believe that the world as encountered – that part of reality that includes us, our language, and the things that we interact with – is ontologically significant. We shall make no headway on a philosophical understanding of the world as encountered if we frame our investigation globally in terms of mind-independence vs. mind-dependence. Instead of starting with a priori metaphysical commitments, I prefer to start with what is at hand – for example, with artefacts about which what we know a lot and whose existence we cannot seriously doubt – and try to think clearly about such things as unencumbered with antecedent metaphysics as possible. I want the metaphysics to emerge from the reflection on the world, rather than the world to be squeezed into a preconceived metaphysical strait jacket.

Conclusion

Our concepts of artefacts are interwoven with concepts of function and malfunction. According to the Deflationary Views of artefacts, however, our artefactual concepts tell us nothing about reality. (Indeed, it is a mystery how we could have come up with such concepts that swing so free of reality in the first place.) According to the Constitution View, our artefactual concepts are a good guide to reality.

Artefacts are ubiquitous and are part of the fabric of human life. According to the Constitution View, a telephone has the property of being a telephone essentially. The property of being a telephone entails a certain intended function (communicating with people remotely situated in space), which, in turn, entails the possibility of malfunction. So, telephones – in virtue of being the *kind* of objects that they are – are always subject to malfunction. When you pick up or try to activate the telephone and do not get a dial tone, the malfunction is as much in the world as telephones are. There are no technical artefacts without functions; there are no functions without the possibility of malfunction. If artefacts are in the world, so are malfunctions.

Acknowledgements

This paper was presented at a conference (Artefacts in Philosophy) at the Delft University of Technology, Oct. 3-4, 2004. Thanks to participants in that conference, especially to Anthonie Meijers, and to Katherine Sonderegger for commenting on a draft. Parts of this paper have been published, in revised form, in Chapter 3 of *The Metaphysics of Everyday Life: An Essay in Practical Realism* (Cambridge: Cambridge University Press, 2007).

References

- Baker, L.R. 1995. *Explaining Attitudes: A Practical Approach to the Mind*, Cambridge: Cambridge University Press.
- Baker, L.R. 2000. *Persons and Bodies: A Constitution View*, Cambridge: Cambridge University Press.
- Baker, L.R. 2002. "On Making Things Up: Constitution and its Critics," *Philosophical Topics*, 30: 31-51.
- Baker, L.R. 2004. "The Ontology of Artifacts," *Philosophical Explorations*, 7: 99-111.
- Lewis, D. 1991. *Parts of Classes*, Oxford: Basil Blackwell.
- Merricks, T. 2001. *Objects and Persons*, Oxford: Clarendon Press.
- Schaffer, J. 2003. "Is There a Fundamental Level?" *Noûs*, 37: 498-517.
- Sosa, E. 1993. "Putnam's Pragmatic Realism," *Journal of Philosophy*, 90: 605-626. Reprinted in J. Kim and E. Sosa, eds. 1999. *Metaphysics: An Anthology*, Oxford: Blackwell, 607-619.
- Thomasson, A.L. 1999. *Fiction and Metaphysics*, Cambridge: Cambridge University Press.
- van Inwagen, P. 1990. *Material Beings*, Ithaca, NY: Cornell University Press.
- Wiggins, D. 2001. *Sameness and Substance Renewed*, Cambridge: Cambridge University Press.

Endnotes

- 1 I am putting aside here consideration of biological function.
- 2 See Baker (2004).
- 3 In other places, I've used the expression 'intentional object' to refer to ID objects. Although I characterized what I meant by 'intentional object' carefully, I am now resorting to the technical term 'ID object' in order to avoid confusion with uses of 'intentional object' associated with Brentano and Meinong.
- 4 As we shall see, I repudiate that notion that the promise is identical to the audible emission, and hence that what is a promise in this world could exist in another world – a world without minds – without being a promise. The relation between the promise and the audible emission is constitution, not identity.
- 5 Amie L. Thomasson discusses varieties of existential dependence in her (1999).
- 6 Peter van Inwagen discusses artefacts in (1990, section 13). David Lewis, as far as I know, never explicitly discusses artefacts, but he is one of the most influential metaphysicians of the day, and I have applied his views (as I understand them) to artefacts.
- 7 van Inwagen (1990).
- 8 Lewis (1991). David Lewis is a four-dimensionalist; it is more accurate to say that on his view the Columbia was a spacetime worm made up of a mereological sum of four-dimensional parts.
- 9 Some eliminativists do not even allow that statements putatively about artefacts are true at all. See, e.g., Merricks (2001).
- 10 cf. van Inwagen (1990, p. 109).
- 11 cf. Lewis (1991, p. 87). I am trying to avoid the language of mereology, because Lewis and van Inwagen differ on whether the particles arranged space-shuttle-wise have a mereological sum. Lewis says yes; van Inwagen says no. I am calling the particles arranged space-shuttle-wise 'an arrangement of particles' in order to be neutral between Lewis and van Inwagen. Neither would quantify over *arrangements* of particles. I think that, metaphysically speaking, Lewis and van Inwagen are on the same side with respect to artefacts.
- 12 See Lewis (1991, p. 81).

- 13 van Inwagen (1990, p. 130). [Emphasis his.]
- 14 For greater detail, see Baker (2000). See also the Book Symposium on Baker (2000) in *Philosophy and Phenomenological Research* 64: 592-635 (2002), and Baker (2002).
- 15 I think that it is an empirical question whether there is an ultimate constituter; but if there is not, then there are still subatomic constituters. See Schaffer (2003).
- 16 For a similar line of thought see Thomasson's (1999).
- 17 See a lengthy discussion of artefacts (specifically, carburetors) in Baker (1995).
- 18 Without such a sharp division, the thesis that all vagueness is linguistic, and hence not de re, becomes problematic. The thesis that all vagueness is linguistic, and hence not de re, requires that language be isolable from the world, from genuine reality.
- 19 I cannot resist an appeal to authority here. "Let us forget once and for all," said David Wiggins, "the very idea of some knowledge of language or meaning that is not knowledge of the world itself." (2001, p. 12).

Antirealism and Artefact Kinds

Marzia Soavi
Department of Philosophy
University of Padova

Abstract

Many realists on kinds deem it highly controversial to consider artefact kinds real kinds on a par with natural ones. There is a built-in tendency in realism to conceive of artefact kinds as merely a conventional classification used for practical purposes. One can individuate three main different approaches characterizing real kinds and accordingly three different types of arguments against viewing artefact kinds as real kinds: the metaphysical, the epistemological and the semantic arguments. The aim of this contribution is to undermine the thesis that it is possible to trace a clear distinction between artefacts and natural kinds in each of these approaches. As a consequence there are no metaphysical, epistemological and semantic bases for claiming that artefact kinds as opposed to natural ones are not real kinds.

1. Realism and artefacts

The realist perspective to which I refer in the present discussion is characterised by the following theses:

- (1) There is a world existing *independently* from human thought and language.
- (2) This world is divided into kinds existing *independently* from human thought and language – these are called “real kinds” or “sortal kinds”.
- (3) An individual object O is a real entity if and only if there is a real kind S such that O belongs to S.

The central role is played by the independence theses (1) and (2). These theses are not meant to exclude the trivial possibility of real entities being products of human actions and in that sense dependent on human thought; what realists want to exclude are those entities that are nothing more than projections of our thoughts and which thus lack an independent nature. The fact that something depends, for its existence, on human thoughts channeled via human intentional actions, cannot be considered a sufficient basis for taking such an entity to be a non-real entity. In Michael Devitt’s words:

Finally, in asserting the independence and objectivity of the world, the realist does not mean to deny certain familiar causal relations involving minds. Beliefs, desires, sensations, and so forth cause behaviour which affects external reality, even creating some items. (1997, p. 16).¹

Many of the authors who accept (1), (2) and (3), and try to draw the line between sortal (real) kinds and nominal (non-real) kinds, consider kinds of artefacts to be non-real nominal kinds. Let (N) be their thesis:

- (N) Kinds of artefacts are not real kinds.²

What may “artefact” mean in (N)? Unfortunately the definition of “artefact” is not straightforward and many problems arise when we try to distinguish natural from artificial objects. Nonetheless, there seem to be objects that we clearly consider to be artefacts, like chairs, cars, cakes etc. The following is the classical characterization:

- (A) An artefact is an object or a substance that is the intentional product of intentional actions.³

Consider, for instance, the case of an artist carving a statue from a piece of wood. The artist intends to create a statue but when carving the piece of wood he will also produce a lot of shavings. According to (A) the statue is an artefact while the shavings are not, because even if both the shavings and the statue are products of the same intentional action, only the statue was meant by the artist to be the final product of his work. Despite some problems concerning agricultural products and artificial substances this distinction between intentional products – the statue – and mere results of intentional actions – the shavings – seems to correspond quite well to our intuitions.

Nonetheless, (A) is not universally accepted by antirealists. David Wiggins, for example, does not adopt it. In his words:

[...] it is not the question of whether a thing was fabricated but rather the difference between satisfying and not satisfying this *condition* that makes the fundamental distinction. (2001, pp. 89-90).

Here Wiggins is speaking about the condition of having a principle of activity founded in law-like dispositions. Unfortunately, as we will see, Wiggins also uses this condition to trace the distinction between real and non-real kinds. Clearly someone aiming at arguing in favour of (N), as Wiggins does, cannot simply use this condition to draw the line between artefacts and natural objects without rendering (N) trivially true by definition. Of course, we could decide to apply ‘artefact’ and ‘natural’ according to such a distinction but here the problem at stake is clearly not that of being coherent in the use of the terms but rather of making an inquiry into the differences between the nature of artefacts and natural objects.

If (N) indeed holds, then it follows with (3) that objects such as tables, chairs, cars, cakes and cities are not real objects *qua* tables, chairs etc. The kinds to which these objects belong are not real kinds but mere conventional classifications and what they allow to individuate are not real objects. When speaking of tables, chairs, etc., we are really referring only to quantities of matter shaped in certain ways.

In the next three sections I will consider three main arguments in favour of (N): a metaphysical argument, an epistemological argument and a semantic argument. These arguments correspond to three main different approaches to characterizing real kinds. My aim is to undermine the thesis that it is possible to trace a clear distinction between artefact and natural kinds in each of these approaches. As a consequence there are no metaphysical, epistemological and semantic bases for claiming that artefact kinds as opposed to natural ones are not real kinds.

2. Metaphysical argument

Metaphysical arguments in favour of (N) are based on Aristotle's idea that there is not a real principle of unity for artefacts;⁴ they do not have their own nature or form, that is to say, they are not substances.

Wiggins, for instance, gives such an argument in *Sameness and Substance* (1980; 2001), which can be analysed as follows:

- (M)
- (i) If a kind S is a real kind, then there are clear principles of individuation for objects belonging to S;
 - (ii) there are no clear principles of individuation for artefacts;
 - (iii) artefact kinds are not real kinds.

In my discussion of this argument I shall accept (i) as an expression of a fundamental metaphysical thesis of the kind of realism under discussion. A real kind – “sortal kind” in Wiggins's terminology – collects objects that share a common nature, and that can be traced in time and space according to some common principles of individuation. These principles of individuation are based on what Wiggins calls “principles of activity” that specify the typical way in which objects of the same kind behave, interact with the environment and change. In Wiggins's words “they are law-like norms of starting to exist, existing, and ceasing to exist by reference to which questions of the identity and persistence [...] can be arbitrated.” (2001, p. 83). Such principles correspond to regularities of behaviour that could be known or unknown to us, so such regularities can either be described in terms of law-like norms or they still have to be discovered.

Any problem of identity for objects belonging to real kinds is founded only in our ignorance about fundamental facts concerning the nature of such objects. Our knowledge of the principles of activity of a real kind can be incomplete or even wrong but we can always obtain more scientific facts. Disputes concerning the identity of real objects can be resolved by new achievements in scientific inquiry. A distinctive mark of real objects is that it is never up to us to decide between conflicting statements concerning their identity.

The truth of the second premise is based on the fact that it is not possible to formulate principles of activity for artefacts analogous to those for natural objects. Therefore, while natural kinds satisfy the metaphysical requirements for being real kinds, artefact kinds do not. Typical problems concerning the identity of artefacts are due to the fact that artefacts can persist through radical mereological changes, interrupting their functioning, and completely dismantling and rebuilding. The result is that the principles of identity for artefacts are so weak that it seems there is no fact of the matter at all about identity claims for artefacts. But even if we can in some way improve the principles of persistence for artefacts, we cannot avoid identity puzzles like the Theseus' ship puzzle, because for many artefacts it is easy to individuate circumstances in which we are compelled to simultaneously apply two different principles of persistence, thus arriving at the unpleasant result of identifying one object with two different objects.

A first principle of continuity that seems to be specific to artefacts is that of continuity of matter or mereological continuity. According to this principle, an artefact that is dismantled and rebuilt, using the same original parts arranged in the same original structure, is still the same artefact. Despite the intuitive appeal of this principle, some problems arise when we try to apply it. It is

not clear if the artefact's existence persists during the whole process of dismantling and rebuilding, whether it can exist even in a dismantled state, or whether we have to admit that there is intermittent existence and, in the last case, in which state exactly does it cease to exist and when does it again start to exist. A second principle of continuity that can be applied to artefacts is the principle of continuity of form or functional continuity. According to this principle, an artefact can undergo the gradual substitution of all its parts and still continue to exist. If we do not accept the application of such a principle, we face the following dilemma: either we deny – against well established common practice – that an artefact could survive the loss of even the smallest of its parts, or we allow the object to persist only through a certain number of substitutions and in such a case we are again stuck with the problem of finding the threshold of its survival. All these well known problems seem to be irresolvable in terms of scientific research; they seem to permit only conventional or even arbitrary solutions.

Situations in which rebuilding and substitution occur simultaneously give rise to identity puzzles. In such cases both the principles of mereological and form continuity can be applied and this leads to an identification of the original object with two different objects, thus leading to contradiction. The conclusion drawn is that there seems to be no fact of the matter concerning the identity of artefacts; principles are so undemanding that it is simply up to us to decide when an artefact starts or stops its existence.

According to Wiggins such a despairing situation for artefact identity principles derives from the fact that there are no principles of activity, no law-like sentences describing the form and behaviour of artefacts. This is not due to a lack of knowledge, but rather to metaphysical matters, that is to say, there are no common laws governing the behaviour of artefacts belonging to the same kind. Consider the example of clocks: a clock is simply, in Wiggins's words, "any time-keeping device". There are many different devices that can perform such a function, devices with different structures that function in different ways so no regularity in behaviour and form can be individuated for all clocks.

A key role is played by the principle of classification that is adopted; it is often said that artefact kinds are mere functional kinds which means that for artefact kind S it holds that:

(F) An object O belongs to S if and only if O has the function F.

Unfortunately, we are not told what a function of an artefact is, or what the truth conditions of a sentence like "O has the function F" are. Does that mean that O can perform F? Or that O can be used for F? Or does it mean that O has a certain selection history? Or that O has been designed for F?

Indeed, this is not the only reason for denying the existence of regular behaviour for artefacts. This would simply amount to the problem of finding a sufficiently fine-grained classification for artefacts. It is easy to develop a way to specify functional criteria of classification so that the only artefacts that perform the same function according to the same principle of functioning belong to the same kind. According to Wiggins this would be sufficient: all available solutions to problems concerning identity of artefacts would remain and would have an arbitrary or conventional character. In order to solve such problems, we cannot appeal to any fact concerning the artefacts themselves for the simple reason that there are no such facts.

2.1. Criticism

Wiggins presents some evidence in support of (ii). The idea is that it is not possible to individuate persistence conditions for artefacts unless we appeal to conventional decisions because artefacts do not have their own nature. Tables are simply quantities of matter that we decide to trace in time and space as continuous existing objects in line with our own interests.

Any property referred to for the purposes of explaining why natural objects are real objects while artefacts are not, has to determine a direct ontological difference between artefacts and natural objects. That is to say: the property has to be one that artefacts definitely have and natural objects definitely lack, or *vice versa*. If we take seriously the ontological task of recognizing when our categories are able to capture real objects and when they are not, we have to be able to detect clear and sharp differences; it seems difficult to allow vagueness to exist in the distinction between real and non-real entities.

As illustrated above, there is a wide range of mereological and structural changes that artefacts can undergo. They can be dismantled piece by piece in a long process or all at once in a single unity of time. They can be rebuilt in the same fashion, their components can be partly or completely substituted by other components, or even by slightly different components, while the whole object still persists in time.

On intuitive grounds it is clear that our chance to control changes with the intention of preserving or destroying artefacts is quite complete, while in the case of natural objects we only have limited opportunities to enter into and control the autonomous processes that determine persistence in time. Moreover while for natural entities we can individuate autonomous regular processes of transformation, we cannot do that most of the time for artefacts. The conclusion is that the persistence of artefacts seems to depend on our decisions to an extent that the persistence of natural objects does not. Is this sufficient for making a distinction between the nature of artefacts and the nature of natural objects that is substantial enough to justify the alleged radical ontological difference?

The fact that artefacts can undergo changes like those described above without ceasing to exist does not qualify as evidence because, in view of our knowledge, we cannot tell whether this depends on the nature of the objects themselves or on the skills and techniques we have at our disposal. It is not at all clear if, even from a biological point of view, there is any general veto on the possibility of comparable changes of parts in the case of living beings – it is true that in the case of simple organisms it is already experimentally possible.⁵

According to Wiggins what makes the difference is the ‘activity’ or, in a more Aristotelian vein, the ‘internal principle of change’. The *activity* of natural objects is something so closely related to their existence that, given a law-like description of such an *activity*, we are able to determine the condition of persistence of natural objects. For artefacts it is not possible to individuate any *activity*, at most we can individuate a function, but whatever a function is supposed to be it is not intimately related to the persistence of artefacts in the same way in which *activity* is related to the persistence of natural objects. An artefact can cease performing its function and even lose the capacity to perform it for a considerably long period of time and nonetheless retain its identity.

Wiggins is not clear as to what precisely an *activity* is, what he is clear about is the relation that there must be between the principle of activity and principles of persistence. As he states:

All the doctrine implies is that the determination of a natural kind stands or falls with the existence of law-like principles, known or unknown, that will collect together the

extension of the kind around two or three representatives of the kind [...] to be something of (a) kind is to exemplify the distinctive mode of activity that they determine. (2001, p. 80).

Wiggins explicitly refers to Leibnizian⁶ and Aristotelian doctrines. The following are the relevant references:

Things which exist by nature [...] such as animals and the organs of these or plants and the elementary stuff [...] have in them a principle of change or rest (in respect of place or growth and decline or alteration generally) [...] the nature of a thing being the source or cause of non-accidental change or rest [...] ⁷

An activity is a chain of internal and/or external causal interactions describable through law-like claims, a kind of process able to determine the persistence of the object. The prototypical examples of such an activity can be all the metabolic processes of the human body. Alternatively life itself can be described as a single complex process resulting from a synergy of different processes. No doubt it is a kind of process describable through law-like sentences. If we adopt such an interpretation of “activity” then what seems to have a central role is the notion of *internal change*. An object endowed with activity has the capacity to change its parts while retaining its proper form and identity. If this is the intended meaning of “activity”, can we really use this notion to trace the distinction between natural objects and artefacts?

Before trying to answer this question, it is worth considering a further specification of what an activity is deemed to be. Wiggins explains in a note:

The Leibnizian echo made by ‘activity’ is deliberate but, outside the monadological framework, it does not have to import anything very different from ‘way of being, acting and reacting’ – something a stone might have. (2001, footnote on page 72).

With respect to our previous interpretation, this is definitely a less demanding notion of activity facilitating the inclusion of real objects of living beings and all other natural objects. Wiggins explicitly refers to the following natural entities: lakes, rivers, volcanoes, springs, seas, and glaciers. Indeed, it is possible to describe in law-like sentences how a stone or a river behave in certain circumstances, but it is not clear at all if such principles are really closely allied to the persistence of those objects to the degree that the doctrine of activity seems to require. Let us consider the case of the volcano mentioned by Wiggins. Of course there is, even literally, an activity of volcanoes describable in law-like sentences but a volcano can suspend its activity for hundreds of years and then start again without ceasing to exist, in much the same way that a clock can stop and restart again, thus fulfilling its function. Rivers and lakes can dry up completely and then be replenished with water without this causing new rivers or lakes to come into existence. It is inappropriate in such situations to even say that a new river or a new lake has come into existence. What is relevant is that such problems cannot be settled by scientific research so that even in these cases we seem to face identity statements that can only be decided in conventional or arbitrary ways. Hence the reason that the activities of lakes, volcanoes and rivers, if they are activities at all, are not able to provide principles of persistence for these natural objects. Furthermore, such a broad notion of activity would enable us to individuate activities for artefacts as well.

We are left with only two possibilities: either we take into account what Wiggins says in the footnote mentioned above or we ignore it. In the second case we are left without a clear idea of

what, in general, an activity is supposed to be in the case of natural beings. The reason for this is that only the processes involved in life seem to match the Leibnizian-Aristotelian descriptions. This narrow interpretation thus excludes from real existence all natural non-living things. Instead, if we take into account a broader interpretation of *activity*, we are able to attribute activity even to non-living entities but in such cases we cannot establish any criterion of persistence for them and, above all else, in exactly the same manner we can also ascribe an activity to artefacts. If displaying activity, according to the less demanding notion of activity, is the only property that anchors the difference between instances of real kinds and instances of non-real kinds, then there seems to be no good reason to maintain that artefact kinds are not real while lakes or volcanoes do qualify as real kinds.

One more string to Wiggins's bow is the Theseus's ship puzzle. Wiggins seems to see this kind of puzzle as a symptom of the particular weakness of the identity conditions of artefacts: in the end it is through the weak and undemanding constraints placed on artefact identity that puzzles like that of Theseus's ship become possible. So the existence of these kinds of puzzles could be seen as a further way of detecting the difference between real and not real kinds.

In biology one may find similar puzzles. Let us consider the case of tubers. Tubers are the parts of roots of some kinds of vegetables from which new plants can grow. In some cases it is even possible to obtain new plants from just part of the tuber. Suppose one takes a tuber *T* and plants it in the ground so that a new plant *P* grows. Suppose that one then takes the very same tuber *T*, cuts half of the upper part away, destroys it and puts the other part in the ground so that *P*^I grows. Can we then say that $P=P^I$? We may have reasons to consider *T* and its upper part as two distinct entities, so the plants that grow from them would be considered to be distinct plants.⁸ This would lead to the conclusion that the tuber did not survive the loss of one of its parts. This is not, of course, acceptable if we consider a tuber from the point of view of what may reasonably be considered to be its *activity*, in fact even when deprived of half of its constituent tuber part, it can still produce a new plant. On the other hand, if we accept that a tuber does not survive the loss of one of its parts, we are left with the problem of accepting a sort of mereological essentialism for tubers or of establishing the threshold of mutilations so as not to compromise their persistence. If, to break this impasse, we accept that *T* is the same tuber in both situations we are bound to say that $P=P^I$. Then imagine a third situation in which a third plant *P*^{II} grows from the upper part of the tuber while the bottom section is destroyed. Can we say that $P=P^{II}$? If we have accepted the identification of *P* and *P*^I, there seem to be no valid grounds for not identifying *P* and *P*^{II}. In a fourth situation in which both the halves of *P* are planted in the ground and two distinct plants result, *P*^{III} and *P*^{IV}, the identification of *P*^{III} and *P*^{IV} with *P*^I and *P*^{II}, respectively, remains straightforward. Hence, for the transitivity of identity, we are bound to identify *P* with both *P*^{III} and *P*^{IV}. The common problem underlying the Tuber puzzle and the Theseus's ship puzzle is that we cannot solve either of them without paying a high price in terms of intuition. The weak point of this puzzle is clearly the first step, once we accept identification of *P* and *P*^I, it is difficult to find reasons to reject the other identifications. Can we really find clear, unequivocal reasons to avoid that first identification on the basis of our botanical knowledge? Does biology tell us what to think about such identity? Can we really think in terms of a kind of empirical research that allows us to obtain new scientific data in order to settle such a problem? The point seems to be that botany and biology are perfectly insensitive to the problem of identity between tubers and parts of tubers and this could suggest that there is no fact of the matter basis to such identities at all. Ultimately, these are the very sciences that are supposed to find the solutions to such problems concerning tubers. So even in this case it seems that a solution may merely depend on conventional or arbitrary decisions. But we do not use puzzles like this one to infer anything concerning the nature of tubers. In any case, as is well known, double identification puzzles do

not confirm any clear distinction between natural and artefact beings so they cannot be used to assert any ontological difference between artefacts and natural beings.

What seems to affect metaphysical arguments of this sort is both the tendentious notion of *natural beings* and the adoption of naive notions of *artefact*. While the explicit intent behind discussing the nature of artefacts is to confront artefacts with natural beings, the principle adopted for the distinction does, in many cases, clearly lend itself better to living beings than to natural objects in general. This is what happens with the doctrine of the principle of activity discussed above. It seems to be perfectly in line with the Aristotelian doctrine, but this doctrine has to prove its own reliability and it offers no alternative than to accept the ultimate conclusion that only biological entities exist.⁹

On the other hand, the discussion on artefact kinds is based on a too naive conception of artefact types. Let us return to the example presented by Wiggins: a clock is any time-keeping device, a pen is any rigid ink-applying implement. These cannot be considered, even from the point of view of everyday language, adequate characterizations of what clocks or pens actually are; a lot of objects that satisfy such descriptions would never be considered as candidate items for the categories of clocks or pens.

The point is that while in the case of natural entities we are zealous in admitting that science provides us with the best conceptual instruments for sorting objects into real kinds, in the case of artefacts we seem to be perfectly happy with a classification directly drawn from common terminology like “pen” and “clock”. We seem to forget that artefacts are products of scientific research as well: that they rely on the different possible solutions to what we can call functional problems. Different solutions can give rise to different kinds of artefacts. There is, for artefacts as well as for natural objects, the aspect of a taxonomy based on a scientific technical approach and not simply on the loose common sense categorization.

3. Epistemological argument

The following statement is illustrative of the reasons provided in favour of (N) in epistemological arguments:

Members of nominal kind do not share a common hidden nature, and we can give an analytic specification in terms of form and function of what it is to be a member of the nominal kind. One reason for distinguishing nominal kinds is that they do not support inductions in the following sense: the fact that several examined chairs are upholstered, say, does not support the claim that all the chairs are upholstered. The fact that several chairs are wooden does not support the claim that all the chairs are wooden and so on. In fact if a scientist were interested in chairs as a subject of scientific study and got himself a good specimen and started to examine it closely in order to discover the nature of chairs, we would think that he was crazy. (Schwartz 1980, p. 189).

The main reason here adduced for (N) is that artefact kinds do not support induction. As Schwartz put it, we cannot infer any truth about other chairs (*qua* chairs) from the observation of some exemplars. This thesis seems to have two corollaries:

- (a) The only acceptable inductions on artefact kinds are those founded on the nature of the material composing the artefacts, so it is the natural kind that corresponds to the material supporting the induction and not the artefact kind.

- (b) There is nothing new to discover about artefact kinds. That is to say, no new law-like generalizations are possible.

It is important to bear in mind that the epistemological argument – as well as the semantic argument given below – is aimed at proving the metaphysical thesis according to which there is no common nature for objects that fall under the same artefact kind. So the argument would take the following form:

- (E)
- (i) If a kind is a real kind, it can be trusted for induction;
 - (ii) artefact kinds cannot be trusted for induction;
 - (iii) artefact kinds are not real kinds.

Under the assumption that kinds are either real or nominal, it follows that artefact kinds are nominal kinds.

3.1. Criticism

I will address the following two questions concerning the epistemic argument: in what respect is (ii) true if it is true at all? Can it really be taken to show that artefact kinds are not real kinds?

At face value (ii) seems to be convincing, but I maintain that its intuitive appeal is misleading and merely due to the examples chosen for supporting it. We all agree that it is not possible to infer that all chairs are made of wood simply from the fact that some chairs are made of wood. Schwartz claims that in general we do not trust categories such as *chair*, *computer* or *telephone* in terms of universal induction because we know that the objects that fall into such categories do not share a common structure.

Perhaps the intuitive appeal of Schwartz's example lies in the particular properties and categories mentioned. Let us first consider the following examples: we would never take the quality white-skinned to be a projectable property of the kind *human being*; we would never presume, taking the tiger as the basis for our induction, that all mammals have stripes. *Human beings* and *mammals* are kinds of a higher level with respect to those upon which *being white* and *having stripes* can be projected. In much the same way, the examples concerning artefacts could be misleading because the material of which a chair is made or other details of its structure could be mere incidental properties with respect to the nature of chairs. A classical position in the literature is to take the nature of artefacts and allow that to coincide with their function. Indeed, it is a stance that Schwartz seems to adopt. If we accept such a thesis, it is not surprising that a property such as *being constituted of wood* cannot be projected onto such categories. What these examples therefore prove is that either induction of artefact kinds is generally not possible or that the strategy of characterizing artefacts through such a generic functional description (by, for instance, stating that a chair is an object that has the function of being used for sitting upon) is not the right strategy.

A further possibility is to accept a generic functional characterization for artefact kinds but deny that such physical properties can be projected onto such kinds. We may need to seek different projectable features. Speaking on purely intuitive grounds, it seems possible, for example, to project the minimal physical requirements for an object to function as a chair: such as having a structure that allows human beings to maintain a certain posture and having certain dimensions and certain proportions in its component parts. Naturally this strategy has to face the problem of

the relation between the function and physical structure of an object. Nonetheless, promising as it is, this strategy may fail for another reason. Artefact kinds also support malfunctioning statements which implies, as is widely conceded, that the function attributions involved in the sorting of artefacts have to be normative attributions. In other words, even if the function of pens is to write, there might be pens that cannot be successfully used for that purpose.

Concerning (a) we all agree on the impossibility of inferring from the fact that some chairs are made of wood that all chairs must be made of wood. But it seems possible to draw conclusions about more restricted kinds, such as *wooden chairs*. For example, we can infer from the fact that a wooden chair burns in certain conditions that all wooden chairs will burn in the same conditions. Indeed this does not seem to be an induction concerning the kind *wooden chair*. The fact that a wooden chair burns in certain conditions seems rather to be directly deduced from what we know in general about the properties of wood and all wooden objects. So the real induction seems to be supported by the natural kind *wood* and not by the artefact kind *wooden chair*.

If all the possible inductions on artefact kinds can be rewritten as deductions from the properties of the materials they are made of, then there are no properties that can be exclusively attributed to chairs as such or to wood as such. This seems to be plainly false. We frequently seem to accept inductions for structural properties on artefacts that have the same source of design and also inductions of functional properties on artefacts that have the same structure. Obviously the structural properties of an object cannot depend exclusively on the materials of which it is composed whilst functional properties¹⁰ depend on the material in question as well as on the structural properties. For example, the fact that the chair I am sitting on at this moment is able to hold my weight, partly depends on the properties of its materials and partly on its structure. The belief that all objects that have a structure similar to that of a certain chair can hold the same weight and the belief that all objects intentionally produced according to the same design can also support the same weight seems sufficiently warranted.

It may be that the whole idea behind this thesis is that once we know what materials an object is made of and once we know its structure, we can then explain all the physical and functional properties of the whole artefact. All the relevant properties of artefacts, including their functional properties, do not emerge with their structural and material properties which means that they do not give rise to a new ontological level. The reducibility of artefact properties to their materials and structure leads to the thesis that artefacts are indeed ontologically superfluous. The argument thus takes on the form of an application of Ockham's Razor principle on the basis of epistemological considerations.

As far as the metaphysical argument is concerned it cannot establish a difference between artefacts and natural kinds that would support the thesis that the latter but not the former are real kinds. Even in the case of many of the biological functions characterizing biological entities – that of pumping blood, for example – it is possible to explain them on the basis of their structural and material properties.

Ultimately I think that it is possible to claim that there are properties that play a major role in artefact classifications into kinds but are not so easily reducible, namely normative functions. A major feature of the notion of *artefact* adopted here is that there is something that artefacts of certain kinds are expected to perform, that is their normative function or the use to which they are put.

Many different criteria can be adopted to account for these notions and we can roughly divide them into two different groups: those confirming the existence of a selective history for artefacts and those that pertain to human intentions. In both cases it is clear that a straightforward reduction of the kind described above is not possible.

Concerning (b), the idea is that with existing artefacts there is nothing new to be discovered. Yet tests may actually have to be done on artefacts to verify their behaviour in particular circumstances and often new unpredicted properties will be discovered in the process. A simple knowledge of the properties of materials and their structure does not provide us with a suitable epistemic basis to deduce all the properties of artefacts.

4. Semantic argument

By “semantic argument” I mean all those arguments that set out to demonstrate (N) on the basis of claims about the semantics of general artefact terms. The idea is that while terms for natural kinds “refer” according to the direct theory of reference, artefacts terms do not, and this is taken to be a clear indication of the nominal nature of artefact kinds.

According to a specific version of realism, one of the main metaphysical points of the direct theory of reference is that it can accommodate the fact that general terms for natural kinds continue to refer to the same kinds of natural entities, even if our knowledge and conception of them radically changes in the light of new scientific achievements. We need natural kind terms to refer, according to this theory, so as to guarantee that reference to kind terms remains the same despite possible radical changes in our conceptions.

We can depict a naively standard process of improvement of our knowledge and classification of the things existing in the outside world. At the beginning they are only collected because of a certain similarity in their exterior and superficial qualities. Then, through more precise and experimental examinations of their inner structure, we improve our knowledge by discovering the common causes responsible for their separate similarity, that is to say, by discovering the common nature of objects classified as being objects of the same kind. Of course, things are not always so straightforward, and our initial categorization can be completely misleading as we can collect objects that appear similar but are of a completely different nature. The discovery of such differences leads to a rearrangement and improvement of our classification which is such that the boundaries of our kinds match those of real kinds.

I shall distinguish two main theses concerning the semantics and the use of natural kind terms which are taken as prototypical examples of real kind terms:

- (A) The reference to a natural kind term is not determined through a description that specifies the meaning of that same term.
- (B) The use of natural kind terms presupposes the existence of an underlying nature that makes something the kind of object that it is.

Artefact kind terms satisfy neither (A) nor (B). Here there are three theses that seem to come together: the thesis of the direct theory of reference for real kind terms, that of the common inner structure of objects belonging to the same real kind, and that of the necessity of scientific inquiry if new knowledge is to be acquired on the nature of these objects. The fact that the real nature of a thing does not depend on our conventions implies that we have to get to know it through study and inquiry. In this case a direct theory of reference guarantees that throughout all these inquiries

the reference of the term we use remains constant. If artefact kind terms do not function according to a direct theory of reference we will not need such continuity because we do not need any inquiry into the nature of artefacts. We already know all that there is to know about artefacts because their nature depends on our own decisions.

To provide an argument for the metaphysical thesis we not only need to say that artefact kinds do not refer, according to a direct theory of reference, but that they inherently do not do that. Otherwise, either we admit that it is possible to have artefact kinds that are real kinds, or we have to admit that the semantics of general kind terms is not a reliable indication of the metaphysical status of the corresponding kind. The first conclusion seems to be absurd: the fact that artefact kinds are not real kinds seems to be a necessary truth, if indeed it is a truth at all. The second option would be self-defeating for antirealists using the semantic argument. Hence the thesis that artefact kind terms do not actually refer according to a direct theory of reference but rather according to a descriptive theory of reference which is not sufficiently strong to show that artefact kinds are not real kinds. I will try to show that it is false to maintain that artefact kind terms necessarily refer according to a descriptive theory of reference. The point is that if the analysis given at the beginning of the paragraph is correct, the semantic behaviour of artefact terms will seem sensible within the context of our epistemic position with regard to artefact kinds.

4.1. Criticism

It is commonly held that once we have a description that enables us to select a certain class of artefacts, for example “objects used for drinking, normally made of a sufficiently light material and in a shape and dimension that allows us to handle them”, namely glasses, we do not question the inner structure or real nature of these objects and we do not try to advance our knowledge on such objects. That is to say, we do not need artefact kind terms to function in the way that natural terms do. The reason for that, it is argued, is because there is not an inner structure or real nature to discover. Nonetheless, it is not clear if the reason why artefact kind terms normally behave like abbreviations of descriptions is metaphysical at all. More precisely, is not clear if this depends either on the nature of artefacts or on the status of our knowledge at the time of introduction of new artefact kind terms.

Schematically, the structure of the argument is as follows:

- (S)
- (i) Natural kinds are prototypical examples of real kinds.
- (ii) In order to give an account of the improvement of our knowledge of natural kinds we need a direct theory of reference for natural kind terms.
- (iii) In order to give an account of the knowledge of artefact kinds we do not need a direct theory of reference for artefact kind terms.
- (iv) Artefact kinds are not real kinds.

There are general terms such as ‘bachelor’ or ‘widow’ that are traditionally taken to refer, in line with a descriptive theory of reference. A bachelor is simply any human adult male who is not married, a widow is any human female whose husband is dead. There is nothing more or different to discover about bachelors and widows. We introduce these terms as abbreviations of certain descriptions which facilitate the formation of true analytic statements. According to some authors artefact terms are terms like ‘bachelor’ and ‘widow’, that is to say, they are used as abbreviations for descriptions. I will try to demonstrate that it is perfectly possible to conceive a use of artefact kind terms which remains analogous to the use of natural kind terms.

Let us consider the case of manufactured items from an ancient civilization which have been discovered in a grave. There is no doubt about the fact that these are artefacts. Archaeologists do not have any idea about the possible function of these objects, but by observing their physical structure and on the basis of information derived from other civilizations they conclude that the objects were used for medical ends. They are more or less oval, one inch square, flat, thin, made of metal and have a blade on one side. It becomes a thesis that is accepted by the entire community of competent archaeologists and historians. These objects are named 'glabre'. Now we have a new term for artefact kinds and, according to the previous theory, we are supposed to use 'glabre' as an abbreviation for the description of the function and at least some of the main features relating to the objects found. Years pass and some new graves are discovered in the same area. New glabres emerge from these graves, along with some written documents. From these documents it becomes apparent that these objects do not have the function previously attributed to them and also that there are many different types of glabres. For example, they can be devices used for sacrifice in certain ceremonies and they can be of different shapes and dimensions. What would archaeologists conclude in such a case? That glabres do not exist? Or that glabres have a different function to that previously supposed? Would they not say that they had been wrong about the glabre? It is unlikely that archaeologists would collectively conclude that glabres do not exist, instead they would probably announce that they have discovered what glabres really are, what was their real function and what are their typical features.¹¹

It would be possible, if the previous theory were right, for 'glabre' to start functioning as an abbreviation of a description but the glabre would almost certainly not have behaved in such a way in the circumstances described in the example. What is clear is that with respect to glabres archaeologists will have found themselves in the same epistemic situation that we were once in with respect to gold.¹² What is thus clear is that what determines whether 'glabre' behave according to a direct theory of reference or not are epistemic more than metaphysical facts. The semantic behaviour of 'glabre' does not depend on the nature of the glabres but on the initial state of ignorance of the archaeologists.¹³ We could say that we normally know in conjunction with artefacts precisely what we do not know about natural objects, that is to say, we know about their inner structure, their common nature, what makes them the kind of artefacts they are.

Let us now compare water and polyethylene. We know that the molecular structure of water is H₂O but 'water' is not an abbreviation for 'substance with the molecular structure H₂O' because we could just as well have discovered that the molecular structure of water was XYZ. Putnam called this "epistemic possibility". Given our initial knowledge about water it could have turned out that water has a completely different molecular structure had we perhaps been ignorant about its molecular structure. Where artificial substances such as polyethylene are concerned – granted that we know from the beginning everything that is essential to its chemical structure – it seems that we lack the same type of epistemic possibility we possess for water. Nonetheless, I cannot see how this difference between water and polyethylene could depend on the nature of these substances rather than the difference between our epistemic position with respect to them. I am sure that no chemist would draw the line between real chemical kinds and nominal chemical kinds and I can see no metaphysical reason to substantiate such a claim.

5. Conclusion

The epistemological and semantic arguments do not seem to be real autonomous arguments. They tend rather to be used by antirealists as tools for persuasion, persuasions that I have claimed can be misleading. Only by distinguishing them from the main metaphysical argument, can one fully

reveal their weaknesses. The metaphysical argument is thus ultimately presumed to carry the burden of the distinction between real natural kinds and nominal artefact kinds but it is not able to individuate the criteria for a clear-cut difference between the two. I think the conclusion ought to be that the distinction between artefacts as intentionally-produced objects and natural entities as non intentionally-made objects is metaphysically irrelevant with respect to the realist versus the antirealist debate on artefacts, while all the other attempts considered here that have aimed at providing an alternative and relevant way of distinguishing between artefacts and natural entities fail.

References

- Dipert, R. 1993. *Artifacts, Art Works and Agency*, Philadelphia: Temple University Press.
- Devitt, M. 1997. *Realism and Truth*, 2nd edition, Princeton: Princeton University Press.
- Elder, C.L. 1989. "Realism, Naturalism, and Culturally Generated Kinds," *The Philosophical Quarterly*, 39: 425-444.
- Hilpinen, R. 1992. "On Artifacts and Works of Arts," *Theoria*, 58: 58-82.
- Hilpinen, R. 1993. "Authors and Artifacts," *Proceedings of Aristotelian Society*, 93: 155-78.
- Katayama, E.G. 1999. *Aristotle on Artifacts: A Metaphysical Puzzle*, Albany: State University of New York Press.
- Kornblith, H. 1980. "Referring to Artefacts," *Philosophical Review*, 89: 109-114.
- Leibniz, G.W. 1696. "On Nature Itself, or on the Inherent Force and Actions of Created Things," in: L.E. Loemker, ed., *Philosophical Papers and Letters*, Dordrecht: Reidel.
- Merricks, T. 2001. *Objects and Persons*, Oxford: Oxford University Press.
- Schwartz, S.P. 1980. "Natural Kinds and Nominal Kinds," *Mind*, 89: 182-195.
- Thomasson, A. 2003, "Realism and Human kinds," *Philosophy and Phenomenological Research*, LXVII: 580-609.
- Van Inwagen, P. 1990. *Material Beings*, Ithaca: Cornell University Press.
- Wiggins, D. 1980. *Sameness and Substance*, Cambridge, MA: Harvard University Press.
- Wiggins, D. 2001. *Sameness and Substance Renewed*, Cambridge: Cambridge University Press.

Endnotes

- 1 This position is only acceptable if we assume a certain privileged position for entities such as minds, thoughts, language and persons. Such an assumption may be problematic for realism but this is ignored in the arguments provided for this contribution.
- 2 From now on 'antirealist' will be used to refer to those authors and theories that endorse (N). I take (N) to be equivalent to the thesis that artefact kinds are nominal kinds.
- 3 This definition corresponds to the definition offered by Dipert (1993) and Hilpinen (1992; 1993).
- 4 Aristotle's position with regard to artefact substances is controversial. For a detailed examination see Katayama (1999). There are other important metaphysical arguments that I do not take into account here: for example van Inwagen's (1990) argument dealing with problems of composition and Merrick's (2001) argument dealing with problems of causal supervenience. They are based on general approaches that are not sensitive to the natural/artefact distinction nor to the problems of kinds.
- 5 It is sufficient to consider the well-known technique of nuclear transfer in cells, widely used in cloning experiments, that involves taking the nucleus out of an unfertilised egg and replacing it with one from an adult cell.
- 6 Wiggins (2001, p. 81) quotes Leibniz *On Nature Itself, or on the Inherent Force and Actions of Created Things*, (translated by Loemker (1969)).
- 7 Aristotle Physics Book II, Wiggins (2001, p. 81).
- 8 The idea that non-identification of the two halves forces us to deny the identity of the resultant plants seems to rely on essentialism concerning origin. That is to say, two objects are identical if and only if they originate exactly under the same conditions. Obviously we could question such a thesis and claim that even if the two halves are not identified we would still have the same plant. It is clear that in both cases the argument holds and so I am not concerned about the choice of the right position to be endorsed on this point.

- 9 This thesis is defended, for example, by van Inwagen (1990) on the basis of mereological arguments but I think that the belief that only objects involved in the process of life can create new objects combined with the ideas surrounding activity and internal change adopted here by Wiggins, have more in common than one might at first sight think.
- 10 When referring to functional properties I do not use a normative notion of function, but a dispositional one that we might call capacity.
- 11 A similar example is discussed by Kornblith (1980).
- 12 The same point is made by Thomasson when she discusses the semantic and epistemic dependence relative to a social group. “But one must be cautious, for properly speaking, the issue of what epistemic relation or theory of reference is appropriate for a given kind is relative not only to the kind, K, of entities, but also to a certain group, G, of people” (2003, pp. 583).
- 13 A similar point of view is defended by Elder (1989) but he adopts the inverse strategy. He tries to show that it is possible to conceive of terms such as ‘gold’ that behave according to a descriptive theory but do simultaneously refer to real natural kinds such as a kind of gold.

Relative Identity and the Number of Artifacts

Massimiliano Carrara
Department of Philosophy
University of Padova

Abstract

Relativists maintain that identity is always relative to a general term (*RI*). According to them, the notion of absolute identity has to be abandoned and replaced by a multiplicity of relative identity relations for which Leibniz's Law does not hold. For relativists *RI* is at least as good as the Fregean cardinality thesis (*FC*), which contends that an ascription of cardinality is always relative to a concept specifying what, in any specific case, counts as a unit. The same train of thought on cardinality and identity is apparent among those – *Artifactualists* – who take relative identity sentences for artifacts as the norm. The aim of this paper is (i) to criticize the thesis (T1) that from *FC* it is possible to derive *RI*, and (ii) to explain why Artifactualists mistakenly believe that *RI* can be derived from *FC*. The misunderstanding derives from their assumption that the concept of *artifact* – like the concept of *object* – is not a sortal concept.

Keywords: Relative identity, cardinality, sortal concept.

1. Introduction

Let *a* and *b* be any two objects and consider the claim that one cannot judge whether *a* is identical to *b*, or whether *a*, for example, remains “the same” unless one specifies some kinds of *things F*. Or, in other words, let identity be a relation which is always relative to some general term of an appropriate kind and accept, moreover, that *a* and *b* can stand in the relation “same *F*” (formally ‘ $a =_F b$ ’) but simultaneously not be in the relation “same *G*” (‘ $a \neq_G b$ ’), where ‘*G*’ is, again, like ‘*F*’, a symbol for a common noun standing for a kind of thing – even though ‘*G*’ stands for a property that *a* and *b* possess. This claim is the *relative identity thesis (RI)*.

An easy way to illustrate *RI* is on the basis of the following example. Consider a gold ingot that is first used to make a ring and is afterwards melted down to make a brooch. One and the same gold ingot can then – at different times – be different jewelry. According to P.T. Geach (1967/68) – one of the first supporters of relative identity – examples such as these are by no means an exception to the rule; they are rather the norm.

As the example shows, it is usual to find sentences implicitly committed to this relative identity thesis *RI* for artifacts:

- (1) *a*, the *Ford Fiesta* I saw yesterday, is the same car as *b*, the *Ford Fiesta* I see now, but *a* is not the same sheet of metal as *b*.
- (2) The *Goldberg Variations* as played by Glenn Gould is the same piece of music as in Murray Perahia's rendition, but the two do not constitute the same interpretation of Bach's masterpiece.
- (3) Theseus' ship is the same collection of planks as the reassembled ship, but the two entities are not the same ship.
- (4) The inscription ‘identity’ is the same word (i.e., a so-called type-word) as the inscription ‘identity’, but it is not the same inscription (token-word).¹

So, we can – at least *prima facie* – claim that *there exist linguistic phenomena* of relative identity. And one can add that they are very common, and that they are furthermore *prima facie*, plausible, in particular for artifacts, as the above examples indicate.

Let us call *Artifactualists* those who take relative identity sentences for artifacts to be the norm. For Artifactualists identity is a relation that is always relative to some general term of an appropriate artifact kind. Moreover, Artifactualists claim that there are, or that there could be, cases in which *a* and *b* – where *a* and *b* are artifacts – stand in the relation “same *F*” but not in the relation “same *G*”, where ‘*F*’ and ‘*G*’ are count nouns for specific artifact kinds such as ‘car’ and ‘fork’, even though ‘*G*’ stands for a property that both *a* and *b* possess. (1) – (4) are examples of the thesis just mentioned. According to Artifactualists, examples such as these are by no means the exception, but rather the norm.

Relativists (*RI supporters*) add that the reasons supporting the notion of relative identity are the same as those that bear out what they take to be a strictly connected thesis, namely: a Fregean cardinality (*FC*) thesis. Fregeans (*FC supporters*) claim that any numerical ascription underscores a concept whose role is to specify the kinds of objects to be counted, i.e., what, in any given case, has to be taken as *a unity*. The train of thought which underpins Relativists’ argument is the following: if the Fregeans are right in claiming that it makes no sense to talk of counting objects in general, because what are counted are always objects of a specific kind then, given the strict connection that there is between cardinality and identity, one is also right to claim that it makes no sense to talk of individuating objects in general, because what is individuated is always an object of a specific kind.²

The Relativists’ general idea is that *RI* is at least as good as the *FC* thesis: if the latter holds, then the former has to hold as well because it is very similar.

The same argument could be attributed to Artifactualists. To them it does not make sense to talk of counting artifacts in general, because what are counted are always artifacts of a specific kind. But, given the strict connection that there is between cardinality and identity, one is also right to claim that it does not make sense to talk of individuating artifacts in general, because what are individuated are always artifacts of a specific kind. “When does the modification of an object (or objects) by an agent lead to the existence of a new object?” – Hilpinen asks in *Authors and Artifacts*. “This depends on concepts [...] used for describing objects, that is, on the ways we choose to divide the world into objects”.³

The aim of this paper is (i) to criticize the thesis that:

(T1) from *FC* it is possible to derive *RI*,

and (ii) to explain why Artifactualists mistakenly believe (T1) that *RI* can be derived from *FC*. Their reason for this belief is related to their assumption that the word “artifact” is equivalent to the word “object”: “artifact” is not a *sortal* term, or, in other words, the concept of an *artifact* – just like the concept of an *object* – is not a *sortal* one. *Sortal* concepts possess an identity criterion; Artifactualists assume that *object* and *artifact* do not possess an identity criterion.

2. Some relevant consequences of the Relative Identity thesis

The doctrine of relative identity includes three claims. Claim (A) is that:

$$(a =_F b)$$

is not equivalent to:

$$Fa \wedge Fb \wedge a = b$$

Or, in other words:

$$(P) \quad (a =_F b) \leftrightarrow (Fa \wedge Fb \wedge a = b)$$

is not true. One argument against (P) is the following one.⁴ Let us consider the following two sentences:

- (5) Lord Newriche discussed armorial bearings with a herald yesterday and discussed armorial bearings with the same herald again today.

and:

- (6) Lord Newriche discussed armorial bearings with a man yesterday and discussed armorial bearings with the same man again today.

Both (5) and (6) contain expressions of the form “the same F ” where F is a sortal term, namely, “the same herald” and “the same man”. Now, according to Geach, if (P) held, then (5) and (6) would be logically equivalent to the following two sentences, respectively:

- (5') For some x , x is a herald and Lord Newriche discussed armorial bearings with x yesterday and discussed armorial bearings with x again today.
 (6') For some x , x is a man and Lord Newriche discussed armorial bearings with x yesterday and discussed armorial bearings with x again today.

But this equivalence does not hold. In fact:

- (7) Whatever is a herald is a man

is a true sentence. Moreover (7) is equivalent to:

- (7') For any x , if x is a herald then x is a man.

However, (5') and (7') entail (6'), whereas (5) and (7) do not entail (6). In fact, (6) may be false even if (5) is true (there is an overnight change of staff in the College of Heralds). Geach's conclusion is that (5) and (6) are not equivalent to (5') and (6'). It then implies in turn that:

$$(PL) \quad (a =_F b) \rightarrow (Fa \wedge Fb \wedge a = b)$$

is not true.

The second claim (B) is that:

$$(8) \quad (a =_F b) \wedge (a \neq_G b)$$

is compatible with the fact that *both* Ga and Gb are true: so, for example, a is the same car as b , but a is *not* the same sheet of metal as b even if a is a sheet of metal and b is a sheet of metal.

(A) is the central claim of a Relativist, (B) is just evidence of it. Moving from (A), a Relativist concludes that (C) no one object is absolutely identical to or distinct from another object. In particular, there is no such thing as being just “the same”. The absolute relation of identity needs to be replaced by a multitude of relative identity relations for which the logical principle of the identity of the indiscernibles:

$$(InId) \quad \forall x \forall y (x=y \rightarrow \forall F (Fx \leftrightarrow Fy)),$$

does not hold.

3. Relative Identity and Frege Cardinality

For Relativists *RI* is *similar* to the Fregean cardinality thesis *FC* that there is no such thing as counting or numbering *simpliciter*; there is only counting or numbering according to a concept F , or G .

The *FC thesis* features as the conclusion to an argument often known in the literature as the *relativity argument*.⁵ The argument aims at showing that the real bearers of numbers are concepts and not ordinary objects or ordinary external events. Frege’s strategy is to prove that thesis through a *reductio ad absurdum* of the opposed thesis. It runs roughly like this. If the real bearers of numbers were ordinary objects, cars and spoons for example, then there would be no absolute sense in which a given number could be said to belong to its bearers. The reason is that a given object – or a given artifact – can be conceived of in many different ways. Take, for example, an artifact such as the *Iliad*. One could think of it as one poem, as twenty-four books, or as a large number of verses. The ascription of a number to something would therefore be relative and not absolute. But what would it be relative to?

Frege’s answer is that any ascription of a number to something is always relative to a concept, introduced by a general term (or as he says: “The content of a statement of number is an assertion about a concept”). The role of the concept is to make counting possible by specifying, in each case, the nature of the task to be performed, or – as Frege puts it – the object of investigation. For example, if we say:

$$(9) \quad La\ Rotonda \text{ has zero rooms,}$$

we are ascribing a certain property to the concept *Room of La Rotonda*, i.e. the property of having an empty extension. Instead, if we say:

$$(10) \quad Palazzo\ Barberini \text{ has four rooms,}$$

we are stating that the number which belongs to the concept *Room of Palazzo Barberini* is four, or that the concept has four unities in its extension.

In general, for Frege, a sentence like:

$$(11) \quad x \text{ is one (object)}$$

is always an incomplete way of saying:

$$(11') \quad x \text{ is one } A$$

where A is a specific concept and *one* a certain property we are ascribing to it.

There are – at least – two close similarities between *RI* and *FC*. Firstly, for a Relativist/Artifactualist there is always a monadic predicate involved in an identity statement, i.e., one always says that x and y are *the same something*; and for a Fregean there is always a concept involved in any numerical statement. Thus, for Geach we cannot say that:

$$(12) \quad a \text{ and } b \text{ are the same}$$

but we must say, instead:

$$(13) \quad a \text{ is the same } F \text{ as } b$$

for an appropriate F . And for a Fregean we cannot say of a certain collection that it simply numbers *two*, but we must say instead that it numbers *two* F s for an appropriate F , *room in this building* for example.

The second similarity is that both Fregeans and Relativists/Artifactualists would agree that there are unquestionable logical relations between the notions of identity and counting in:

$$(14) \quad \text{If } x \text{ is not } y, \text{ then they are two.}$$

The antecedent and the consequence of (14) are clearly connected and, given this connection, it seems to be impossible for the relativization to concern only the consequence of (14) and not its antecedent. All things being equal, what is at stake is whether these similarities are sufficient to back the Relativists/Artifactualists thesis that whoever claims that F is essentially involved in any cardinality statement is thereby committed to claiming that F is essentially involved in any identity statement as well.

4. If Frege Cardinality holds, does Relative Identity then hold as well?

The answer to this question would be positive if the reasons justifying the use of the general term F in the first case also justified parallel use in the second case. Since those reasons have to do with the fact that a lack of specification of the general term F in a cardinality statement would signal incompleteness in that very statement, we can conclude that *FC* would justify *RI* if the reasons why a sentence of the form:

$$(15) \quad x \text{ is } n \text{ (where 'n' is a numeral)}$$

is incomplete also justified the incompleteness of an identity statement such as:

$$(16) \quad x=y.$$

The notion that this is how things are is precisely what Geach – a Relativist – believes. Even though Geach does not actually put forward any explicit argument in defense of that thesis, the

train of thought that supports it seems to be the following. If a sentence such as (15) is incomplete because any ascription of cardinality to an object or collection is always relative to a general term which specifies the kind of objects to be counted then, given that the introduction of different completing general terms can determine different ascriptions of cardinality to the same object or collection, it follows that identity must also be relative.

Unfortunately, Frege's grounds for the incompleteness of a cardinality statement such as (15) do not justify the incompleteness of (16) as *relative identity* theorists hold. "One" is, according to Frege, the name of an object – notably a particular kind of logical object – and an object, for Frege, is a saturated entity which, by its very nature, is unsuited to playing a predicative role. The fact that "one" – or any other numeral for that matter – cannot express a property of an object emerges very clearly if one considers the outstanding differences that there are between a sentence such as:

(17) x is one

and:

(18) x is strong.

Take two hammers a and b . While we can combine:

(19) a is strong,

and

(20) b is strong,

to obtain the sentence:

(21) a and b are strong,

we cannot in the same way combine:

(22) a is one,

and

(23) b is one,

to obtain:

(24) a and b are one.

Of course, we can see (22) as an elliptical sentence for:

(22') a is a strong hammer,

as happens when, for example, a person is asked to refer to a strong hammer and, for sake of brevity, she answers by uttering (22).

Even though it is possible for a numeral to figure alone in a predicative position, this does not mean that these cases can be properly described as cases in which the property of the uniqueness of an object is predicated. Otherwise, as Fregeans show, we would have a property that, unlike any other property, would not allow an inference such as:

$$\begin{array}{l} a \text{ has the property } P \\ b \text{ has the property } P \\ \hline a \text{ and } b \text{ have the property } P \end{array}$$

We can thus say that, according to a Fregean, a cardinality statement of the form:

$$(17) \quad x \text{ is one}$$

is incomplete because ‘one’ is an *Eigenname*; it can never function as a predicate but, at most, as a predicate constituent. But which predicate can it be a part of? To answer this question one has to consider the role played by the general term which is introduced to complete the sentence. Now, according to a Fregean, such a term plays the completion role, not because it specifies the sentence predicate – so that the predicate would be ‘... is one *F*’ as a Relativist/Artifactualist maintains – but because it specifies the *logical subject* of the sentence, the object about which something is said when a numerical judgment is made.⁶

Now that we have shed light on the real subject of cardinality statements this puts us in a position to understand what kind of completion is appropriate in such cases, i.e., the predicates of which ‘is one’ is a part. The open sentence “... is one” must be completed in such a way as to express a property which is ascribable not to objects but to concepts. The property in question is that of *having a given number of exemplifications*, in this particular case of *having exactly one such exemplification*.

The Fregean thesis is that what is said in sentences like:

$$(17) \quad x \text{ is one}$$

is that the concept *being identical with x* has the property of *having a singular exemplification*. So, for example, the sentence:

$$(24) \quad a \text{ and } b \text{ are one}$$

(where ‘ a ’ and ‘ b ’ are two different names of my hammer) predicates the property of *having a singular exemplification* to the concept of *being identical with b* . According to this analysis (24) turns out to be equivalent to:

$$(26) \quad a = b.$$

But (26), as one can see, is not equivalent to any relative identity sentence! In fact, for Relativists:

$$(P) \quad (a =_F b) \leftrightarrow (Fa \wedge Fb \wedge a = b)$$

is not true. The claim (C) of Relativists/Artifactualists is that there is no absolute relation of identity: no object is absolutely identical or distinct from another object; there is no such thing as being just “the same”. But, if there is no absolute sense in which a certain object a differs from b , then there is no absolute sense in which a set containing two objects a and b has two objects. So, in Geach’s cardinality picture, for example, if Tom and Bob are *two human beings* but *the same herald*, then the set with the two individuals as its members:

$$\{\text{Tom, Bob}\}$$

will have cardinality $2_{\text{human being}}$ and 1_{herald} because:

$$(27) \quad \text{Tom} \neq_{\text{human being}} \text{Bob}$$

and

$$(28) \quad \text{Tom} =_{\text{herald}} \text{Bob}.$$
⁷

So, for Geach (for Relativists and Artifactualists) the predicate F in a sentence such as:

$$(13) \quad a \text{ is the same } F \text{ as } b$$

tells us which relative identity relation is being questioned, and, similarly, in cardinality claims F ’s role is to determine which relation we are to determine: because the question of x and y and their identity has no absolute sense, nor does the question of whether x and y are one. For Frege, on the other hand, the concept F is essential in cardinality statements because without it there is no specification of what is to be counted. The statement, pointing at a pile of cards:

$$(29) \quad \text{that is one}$$

is ambiguous. We might mean to claim that there is one pack or that there is one card. But the connection with absolute identity is straightforward:

If there is one pack on the table, then for any pack x and y on the table, $x = y$.

If there is one card on the table, then for any cards p and q on the table, $p = q$.

Once it is clear what are the entities in question, there is nothing left to be specified; there is no variety of identity-like relations between which to choose.

To sum up: I have rejected Geach’s (and Relativists/Artifactualists) claim that Frege’s cardinality thesis is analogous to the relative identity thesis by showing that the role played by the completing general term is different.

3. Why Artifactualists mistakenly believe that Frege Cardinality implies Relative Identity?

I propose that the answer to this question is simply that Artifactualists think that the concept of *artifact* – just like the concept of *object* – is not a sortal one.⁸ If ‘artifact’, as ‘object’, is not a sortal word, then the expression ‘the number of artifacts’, like the expression ‘the number of objects’, is meaningful only when supplemented by a sortal term. For Frege the sentence:

$$(11) \quad x \text{ is one object}$$

is meaningful only when supplemented by a sortal term. In the same way, for an Artifictualist:

(30) x is one artifact

is meaningful only when supplemented by a sortal term. Due to the fact that *artifact* is not a sortal word Artifictualists think that it does not make sense to talk of counting artifacts in general; what are counted are always artifacts of a specific kind. But – as with Relativists – given the close similarities that there are between cardinality and identity, they also think that it does not make sense to talk of individuating artifacts in general, because what are individuated are always artifacts of a specific kind. That is one reason why Artifictualists hold that Frege Cardinality (*FC*) implies Relative Identity (*RI*), i.e., it is why they think that the entailment is true.

Putnam, for example, could be viewed as a leading exponent of Artifictualists.⁹ He supports the idea that it is nonsensical to speak of the number of objects with the help of the following example. I bring a friend into a room. There is a table and chair with a book and a spoon on the table. Nothing else. I ask:

(31) How many objects are there?

If the friend's answer is: "Four" I ask again:

(32) Which objects are there?

Answer:

(33) A table, a chair, a book and a spoon.

A reply could be:

(34) What about the pages of the book? And what about the chair's legs?

And so on.¹⁰

One way to stop this chain of queries is to argue that (31) does not hold a determinate meaning. It would be rightly raised if there was a specification concerning the kind or sort of objects to be counted.¹¹ Then Putnam, talking about the fact that (31) does not hold a determinate meaning, argues that also "certain identity statements exhibit the same phenomenon". The examples of identity sentences he proposes are relative identity sentences, along the lines of (1) – (4).¹²

If *artifact* – as *object* – is not a sortal concept, then nothing prevents us from recapitulating Putnam's argument using, instead of (31), something like:

(35) How many artifacts are there?

Why is 'artifact' not a sortal word? According to Frege, if a concept is to be ascribed a finite number, the following two conditions have to be met:

- (i) The concept must determine in a precise way, which objects belong to its extension.
- (ii) The concept should not permit an arbitrary division into parts of the objects which belong to its extension.

A good illustration of the Fregean standpoint is provided by the following quotation:

The concept “letters in the word ‘three’” isolates the ‘t’ from the ‘h’, from the ‘r’, and so on. The concept “syllables in the word ‘three’” picks out the word as a whole, and as indivisible in the sense that no part of it falls any longer under that same concept (Frege 1884, §54).

By contrast, a concept such as *red*, for example, does not isolate what it applies to because it permits an arbitrary division into parts of the objects belonging to its extension (contravening (ii)). As Frege says “we can [...] divide up something falling under the concept ‘red’ into parts in a variety of ways, without the parts thereby ceasing to fall under the same concept ‘red’” (Frege 1884, §54).

For Wiggins condition (i) “could be naturally developed to cover precisely that which we have intended by our conditions upon being a sortal”.¹³ a predicate “with which we articulate or segment the reality of our experience”.¹⁴ As Rumfitt observes, talking of articulation or segmentation is no more metaphorical than when Frege’s talks of delimitation. However, Wiggins explains the metaphor by making explicit the connection between a sortal term and a claim of identity; a sortal term articulates reality because it provides a basis for answering question such as:

- (36) Is this *F* identical to that *F*?
- (37) Is the *F* that is *G* identical to the *F* that is *H*?

Like Frege we can say that a concept/term is sortal if and only if it carries an identity criterion. Conversely, the concepts that do not satisfy that condition are called *characterizing concepts* or *non-sortal concepts*. Strawson claimed that a sortal concept, or in his words a “sortal universal”, supplies a principle for *distinguishing* and *counting* the individual particulars which it collects. On the other hand, a characterizing universal can only be applied to particulars already distinguished, or distinguishable, in accordance with an antecedent principle or method. In rough terms, and with some reservations, we can therefore assert that certain common nouns for particulars introduce sortal concepts, whereas verbs and adjectives applicable to particulars introduce characterizing concepts.¹⁵ Whatever the value of the distinction between *sortal concepts* and *characterizing concepts* may be, we believe that an appeal to this may prove useful in capturing some of Frege’s insights concerning concepts.

If that is so, then asking if ‘artifact’ is a sortal word corresponds to asking if ‘artifact’ extends an identity criterion. In a recent paper written with various others (Carrara et al. 2004) I argued that ‘artifact’ is not a sortal term.¹⁶ Consider the usually adopted philosophical definition for ‘artifact’:

An artifact is a concrete object intentionally produced by human beings.¹⁷

Given this characterization of ‘artifact’ it seems natural to look for conditions concerning the origin of the objects as identity criteria for ‘artifact’. Consider, as plausible, elements of that same origin:

- (A) the matter that constitutes the object at its origin
- (B) the identity of the author.

Can (A) and (B) be accepted as a general necessary and sufficient condition of identity for artefacts? No.

In fact, firstly for (A), it is easy to imagine some circumstances in which there are two artifacts constituted of the same amount of original matter: a chair could be built from the same wood as a previously made table.¹⁸

Secondly, arguing for (B) is equivalent to arguing that if an artifact has been produced by a different author, then it is different. The identity of the author seems to be of great importance to artworks but almost completely irrelevant to other kinds of artifacts such as mass-produced industrial products.

Notice that many other elements of origin, like for example spatio-temporal location or the specific intentions of the author or the instruments used for making the artifact etc., could be seen as good candidates for being considered to be identity conditions for artifacts.

Unfortunately, all these kind of tentative identity criteria for artifacts suffer from the same general problem faced by reductive conceptions of identity criteria. In fact, from a reductive point of view, identity criteria are conceived as principles that reduce issues of identity among objects of a given kind to relations among objects of a more basic kind. Kripke formulates the above notion of identity criteria in this way:

$\bar{x} = \bar{y}$, but \bar{x} is the entity of the new kind associated with x , and \bar{y} is the entity of the new kind associated with y if and only if x and y , which are admittedly distinct objects (or can at least be distinct objects; of course they could be the same object) stand in the relation R . R will in general be some equivalence relation in the unbarred entities (Kripke 1978, p. 36).

Formally:

$$\bar{x} = \bar{y} \leftrightarrow R(x, y)$$

Kripke speaks of a reductionist conception of identity criteria just because identity between objects of a certain kind depends on relations between more basic objects.

There is a fundamental criticism to this kind of reductive conception of identity criteria: if identity criteria have to provide an analysis of identity even if sortally determined, we have to admit objects for which there are no reductive criteria of identity and from which we move on in order to give identity criteria to less basic objects. Otherwise we run into an infinite regress. Consider this example of identity criterion:

- (M=) Material objects are identical if and only if they occupy the same place at all times.

One could ask for a criterion of identity for the notion of place, a criterion that has to be given in terms of entities different from material objects and places. It is not clear what these entities could

possibly be, but there should be some entities if identity between places is to be reduced to identities between more basic entities. It is obvious that continuing to apply the same kind of demand produces an infinite regress.

In order to stop this infinite regress, a reductivistic philosopher could introduce some scientific standards and suppose that a criterion of identity is adequate if and only if the right-hand side of the criterion is an ontological reduction of the left-hand side in terms of the selected scientific standard. This seems to us to be, more or less, for example, Sellars' answer.¹⁹ He argues that "chairs" do not really exist. There are objects that really exist and which correspond to what the layman calls "chairs", but the objects called "chairs" by the layman are part of a pre-scientific, intuitive, picture of the world. "Chairs really are ...", and here the reduction follows the basis of the scientific standard adopted. For example, if the scientific standard adopted is a physical theory, the reduction will be in terms of a bundle of particles and so on. Such a kind of explanation forces the whole question of the adequacy of identity criteria to depend on the reference standard adopted. The problem then becomes: what is the standard in the case of artifacts?

A third reason for maintaining that 'artifact' is not a sortal term relates to certain notorious *puzzles* on artifact identity. Consider the well-known problem of the Ship of Theseus. This is an example of a problem that concerns ordinary artifacts which cannot be decided on the basis of the relevant information. Let v be the old Ship of Theseus that has been restored and n the new one resulting from the replacement of all the old planks. Of course, v is different from n . But, let t be the ship that was sailing in Theseus' time. The relevant information is known and does not allow us to decide whether $t=v$ or $t=n$. This is a question of identity concerning ordinary artifacts.²⁰

Obviously, this does not commit us to the idea that it would not even be possible to find identity criteria for *specific* kinds of artifacts, like *cars*, *forks*, *hammers* etc. A good attempt would be to specify identity criteria based on the function and structure of the objects.

If there is no identity criterion available for *artifact* we can conclude that it is not a sortal concept, and if *artifact* – as *object* – is not a sortal concept then the expression 'the number of artifacts', like the expression 'the number of objects', is meaningful only when supplemented by a sortal term. This is the first step for an Artifactualist.

The next step for an Artifactualist is to argue that – along the same lines as a Relativist – since it does not make sense to talk of counting artifacts in general because what are counted are always artifacts of a specific kind (given the strict connection between cardinality and identity) it does not make sense to talk of individuating artifacts in general, because what are individuated are always artifacts of a specific kind. Hence, from *FC* and the position that artifact is not a sortal concept, an Artifactualist mistakenly believes it is possible to hold *RI*.

4. Conclusions and final remarks on Artifactualists

In this paper I have argued that one reason for Artifactualists to hold that the cardinality thesis is connected to the thesis of relative identity is that they think – like Frege for *object* – that *artifact* is not a sortal concept.

A final remark on Artifactualists concerns the costs and advantages of taking the concept of artifact to not be a sortal concept. According to Quine, identity criteria are required for ontological respectability: only entities that have clearly determined identity criteria are

ontologically acceptable. Think, for example, of the case of properties: they would not be ontologically acceptable because they do not have any suitable identity criterion. If *artifact* is not a sortal concept/term because there is not an identity criterion for artifacts, then *artifacts* are not entities that are ontologically acceptable.

But ordinary language describes a world inhabited by entities of different sorts: people, tables, trees and one could say, more generally, *artifacts*. We utter sentences such as:

(38) There is an artifact on the table

which contain explicit existential idioms and which therefore seem to commit us to the existence of the corresponding entities or artifacts. Even without explicit quantification, the very use of a term – singular or general – naturally suggests the existence of a corresponding entity, as in:

(39) This artifact is heavier than that one.

Some would say that sentences such as these imply the existence of the entities named. Others would say that they presuppose the existence of those entities. Either way, the existential import can hardly be questioned.

If artifacts are not ontologically acceptable (38) becomes misleading for Artifactualists; they should then argue that its grammatical form is not ontologically transparent and that only a suitable reformulation would exhibit its proper truth conditions. For example, (38) could be paraphrased as:

(38') There are *xs* on the table, and these *xs* are arranged artifact-wise,

where the bound variable ranges over accepted entities. (38') would be true even if the original sentence (38) were, strictly speaking, false.

Thus, another consequence of the Artifactualists' thesis to the effect that *artifact* is not a sortal concept is that natural language has to be considered to be ontologically opaque: ordinary sentences must be suitably rewritten or paraphrased before questions of ontological commitment may be raised.

The problem which then ensues is this: according to what criteria do Artifactualists feel entitled to change the meaning of what one says in (38)? Why do we have to accept that – in this case – there is, on the one hand, a language in use that is highly idiomatic but ontologically deceptive while on the other hand there is regimented language that is hardly utterable but ontologically transparent or “intrinsically non-misleading”, as Ryle put it?²¹ On what grounds should we accept that?

One solution to the above questions is that the revisions seems to be necessary if we are to avoid the traps of grammatical form: in such cases we have to accept that the grammatical form of a sentence such as (38) turns out to be deceptive in terms of its semantic analysis.

In general, one familiar form of argument when rejecting a grammatical form or a certain piece of language in use is that, in rejecting it, we avoid certain problems or inconsistencies associated with it: if some well-confirmed logical or epistemological principles become false using a certain grammatical form then that same grammatical form will be deceptive in its semantic analysis. For Artifactualists, in rejecting that *artifact* is a sortal term and by holding *RI*, we avoid the notorious

puzzles of coincidence like, for example, the statue/lump and the Ship of Theseus puzzles. A revision of the grammatical form of a sentence such as (38) – which is highly idiomatic but ontologically deceptive – then seems to become necessary.

The cost of denying that artifact is a sortal concept is thus that one cannot thereafter take the ontological commitment of ordinary language at face value.

Acknowledgments

Thanks to Gillian Davies, Wybo Houkes, Vittorio Morato, Alex Orenstein, and Pieter Vermaas for their comments on a previous version of this paper.

References

- Baker, L.R. 1997. "Why Constitution is not Identity," *Journal of Philosophy*, 94: 599-621.
- Blanchette, P. 1999. "Relative Identity and Cardinality," *Canadian Journal of Philosophy*, 29: 205-224.
- Bloom, P. 1996. "Intention, History and Artifact Concepts," *Cognition*, 60: 1-29.
- Carrara, M. and P. Giaretta. 2004. "The many facets of identity criteria," *Dialectica*, 58: 221-232.
- Carrara, M., P. Giaretta, V. Morato, M. Soavi, and G. Spolaore. 2004. "Identity and Modality in OntoClean," in: A.C. Varzi and L. Vieu, eds., *Formal Ontology in Information Systems*, Amsterdam: IOS Press, 128-139.
- Frege, G. 1884. *Die Grundlagen der Arithmetik. Eine logisch-matematische Untersuchung Über den Begriff der Zahl*, Breslavia: Kábner.
- Garbacz, P. 2004. "Subsumption and Relative Identity," *Axiomates*, 14: 341-360.
- Geach, P.T. 1967/68. "Identity," *Review of Metaphysics*, 21: 3-12.
- Geach, P.T. 1973. "Ontological Relativity and Relative Identity," in: M.K. Munitz, ed., *Logic and Ontology*, New York: New York University Press: 287-302.
- Gibbard, A. 1975. "Contingent Identity," *Journal of Philosophical Logic*, 4: 187-221
- Hilpinen, R. 1993. "Authors and Artifacts," *Proceedings of the Aristotelian Society*, 93: 155-178.
- Kripke, S. 1978. *Time and Identity*, manuscript.
- Lowe E.J. 1989. *Kinds of Being. A Study of Individuation, Identity and the Logic of Sortal Terms*, Oxford: Basil Blackwell.
- Putnam, H. 1987. *The Many Faces of Realism*, La Salle, IL: Open Court.
- Putnam, H. 1988. *Representation and Reality*, Cambridge, MA: MIT Press.
- Putnam, H. 2004. *Ethics without Ontology*, London: Harvard University Press.
- Ryle, G. 1931/32. "Systematically Misleading Expressions," *Proceedings of the Aristotelian Society*, 32: 139-170.
- Rumfitt, I. 2002. "Concepts and Counting," *Proceedings of the Aristotelian Society*, 112: 403-430.
- Sellars, W. 1930, *Science, Perception and Reality*, London: Routledge & Kegan Paul.
- Sloman, S. and B.C. Malt. 2003. "Artifacts are not Ascribed Essences, nor are they Treated as Belonging to Kinds," *Language and Cognitive Processes*, 18: 563-582.
- Strawson, P. 1959. *Individuals. An Essay in Descriptive Metaphysics*, London: Methuen.
- Thomasson, A.L. 2003, "Realism and Human Kinds," *Philosophy and Phenomenological Research*, 67: 580-608.
- Van Inwagen, P. 2002. "The Number of Things," *Philosophical Issues*, 12: 176-196.
- Wiggins, D. 2001. *Sameness and Substance Renewed*, Cambridge: Cambridge University Press.
- Yourgrau, P. 1997. "What is Frege's Relativity Argument?" *Canadian Journal of Philosophy*, 27: 131-171.

Endnotes

- 1 The examples (2)–(4) are from Garbacz (2004, p. 348). Garbacz observes that *RI* is very convenient in applied ontologies, such as for example *stratified ontologies*. By “stratified ontology” Garbacz means “an ontology that splits everyday objects into sets of objects”. He proposes considering, for example, a cylindrical brass paperweight. From a functional point of view it will lose its identity when used as a missile, even if from a morphological point of view it retains its identity. “The paperweight used as a paperweight has the same morphological properties as the paper-weight used as a missile, but the cylindrical paper-weight has different morphological properties than the cube paper-weight since they are not congruent. From the topological point of view it will still retain its identity, which it will lose if one of its parts will be detached from it” (2004, p. 352).
- 2 See, for example, this quotation from Geach: “Frege emphasized that “*x* is one” is an incomplete way of saying “*x* is one *A*” [...] or else it has no clear sense since the connection of the concepts one and the same comes out as much in the German ‘*ein und dasselbe*’ as in the English ‘one and the same’ it has always surprised me that Frege did not similarly maintain the parallel doctrine of relativized identity” (1967/68, p. 3).
- 3 Hilpinen (1993, p. 166).
- 4 This argument does not concern artifacts, but it seems to be rather easy for an Artificalist to reproduce the same kind of argument for artifacts along the same lines as (1) – (4). The argument is taken from (Lowe 1989, pp. 65-66).
- 5 For an analysis of the argument see Yourgrau (1997).
- 6 On the same topic see the important paper of Blanchette (1999).
- 7 An Artificalist can supply the same kind of example with gold ingots and jewelry.
- 8 Recently, the same thesis has been discussed by Bloom (1996) and Sloman and Malt (2003) from a cognitive point of view, and furthermore by Thomasson (2003) from a philosophical point of view.
- 9 In Putnam (1987; p. 2004).
- 10 The example is taken from (Putnam 1988, ch. 7).
- 11 Putnam (1987, p. 19). For Hilpinen “[c]haritably interpreted, this should be regarded merely as a somewhat misleading formulation of Frege’s old point that the word ‘object’ is not a sortal expression” (Hilpinen 1993, p. 166). For a response to Putnam see Van Inwagen (2002).
- 12 Putnam (2004, p. 47).
- 13 Wiggins (2001, pp. 75-76).
- 14 Rumfitt (2002, p. 56).
- 15 Strawson (1959, p. 168).
- 16 I repeat here an idea outlined in Carrara et al. (2004).
- 17 In Carrara et al. (2004) we stipulate that the author has to be a human being. Obviously, if we want to include among artifacts all intentionally produced objects, the realm of artifacts can be expanded in relation to the adopted notion of intentional action and intentional agent. So, for example, we could include sticks used by monkeys for catching ants, or paintings created by elephants, etc.
- 18 On this topic see, for example, Gibbard (1975) and Baker (1997).
- 19 For example in Sellars (1930).
- 20 The topic of identity criteria is discussed in more detail in Carrara and Giaretta (2004).
- 21 Ryle (1931/32).

Produced to Use: Combining Two Key Intuitions on the Nature of Artefacts

Wybo Houkes
Department of Philosophy and Ethics of Technology,
Eindhoven University of Technology
and
Pieter E. Vermaas
Department of Philosophy,
Delft University of Technology

Abstract

In this paper we examine the possibilities of combining two central intuitions about artefacts: that they are functional objects, and that they are non-natural objects. We do so in four steps. First we argue that, contrary to common opinion, functions cannot be the cornerstone of a characterisation of artefacts. Our argument suggests an alternative view, which characterises artefacts as objects embedded in what we call use plans. Second, we show that this plan-centred successor of the function-focused view is at odds with the non-naturalness intuition. Third, we show that this intuition can be developed by defining artefacts as produced or human-made objects, but that the resulting definition might collapse into the plan-centred view, and has trouble distinguishing artefact types or kinds. Finally, we propose a division of labour between production and use plans: among objects in general, artefacts are distinguished as human-made objects; within the domain of artefacts, types or kinds are characterised by the use plans in which artefacts are embedded.

Keywords: artefacts; artefact kinds; production; use plans; functions

Introduction

Intuitively, artefacts are non-natural objects. When we see a wren flying over a deserted heath, we infer that a small population of birds lives there; when we find a working watch with a name inscribed, we conclude that someone produced it and that someone lost it. These judgements may be contested, but one can hardly deny that we make them: we routinely and perhaps irrepressibly distinguish artefacts from natural objects, conceptualising the former in terms of productive, primarily human activities and the latter in terms of natural processes such as reproduction.

People not only tend to distinguish artefacts from natural objects, but also make different descriptive and evaluative claims about the distinguished objects. This leads, among other things, to debates about the status of objects that cannot be unambiguously classified as artificial or natural. May, for instance, genetically modified organisms be taken as the products of design, and thus be copyrighted? Do restored landscapes count as parts of our natural environment, to be preserved from further human interference? Intuitions regarding these specific types of objects differ, but all parties in these debates appear to share the opinion that there is something at stake in distinguishing natural objects from artefacts.

A philosophical analysis that explicates the intuition that artefacts are non-natural objects may illuminate the grey area between artefacts and natural objects, and thus resolve the debates mentioned above. Likewise, it may show what exactly is at stake, descriptively or evaluatively, in

distinguishing artefacts and natural objects. In short, it makes sense to introduce the intuitive distinction between artefacts and natural objects as an important constraint and challenge for a philosophy of artefacts.

In some philosophical analyses this distinction has been contested. It has been argued that biological organisms such as domesticated animals and cultivated plants, which are typically taken as natural objects, are artefacts as well (Sperber 2007), and that natural kinds in chemistry such as purified iron actually may be positioned somewhere at a continuum between artefacts and natural objects, a continuum at which there are no principled points for drawing metaphysical distinctions (Grandy 2007). We do not directly address these arguments in this paper. Instead, we note that the distinction between artefacts and natural objects is sufficiently common in philosophy and elsewhere to warrant proposals to explicate it, and to put the burden of proof on detractors of the distinction to refute the proposals. In particular, there is a strong intuitive distinction between objects that are the products of engineering¹ and those that exist independently of human interference. This distinction is regularly taken as a starting point of philosophical characterisations of artefacts, as illustrated by the following encyclopaedia entries, one of which refers to the Aristotelian origin of the distinction:

Any object produced to design by skilled action. (...) Artefacts contrast with natural objects. (Simons 1995, p. 33)

Aristotle divided things into those that “exist by nature” and “products of art” or “artificial products” (*Physics*, Book II, 126b). Artifacts are contrasted to natural objects; they are products of human actions. (Hilpinen 2004. sect. 1)

Yet most philosophers who examine the nature of artefacts start from a different perspective, namely that artefacts are primarily functional objects. The intuitive appeal of this starting point is readily apparent: in most languages, a large number of artefacts is characterised in functional terms, such as ‘screwdriver’, ‘computer’, and ‘paperclip’. In some cases, this ‘function focus’ has led to the view that functions are the essences of artefacts (Kornblith 1980; Wiggins 2001). This function essentialism plays a major role in the debate whether artefact kinds are real or nominal (e.g., Millikan 2000; Elder 2004; 2007),² and in attempts to determine the persistence conditions of artefacts (Baker 2000; 2004; 2007). Furthermore, attempts to bring artefact functions into the fold of general function theories carry a strong suggestion of capturing a central feature of artefacts in doing so (Millikan 1984; Neander 1991; Preston 1998; Krohs 2009; Longy 2009).

The focus on functions in philosophical analyses of artefacts is not necessarily at odds with the equally intuitive characterisation of artefacts as non-natural. However, since both have been used to carve out the domain of artefacts within that of objects in general, an analysis of their relation seems in order: do the intuitions lead to co-extensive characterisations, do they conflict, or – perhaps most interesting – are they complementary?

In this paper, we analyse the relation between the non-natural and functionality intuitions regarding artefacts. We argue for the complementarity of these intuitions, in four steps. First, we consider the philosophical focus on artefact functions. We draw upon the results of some of our earlier papers to show that function essentialism – the extreme form of function focus – is unstable, in the sense that it is undermined by a phenomenologically adequate notion of artefact function. These results suggest an alternative view, which characterises artefacts as embedded in what we call use plans, and which salvages the functionality intuition. The second step in our argument is that a plan-centred view of artefacts, just like the original function-centred view,

seems at odds with the intuitive characterisation of artefacts as non-natural. In the third step, we explore this apparent conflict by considering one plausible way of developing the non-naturalness intuition. We argue that the resulting characterisation of artefacts as intentionally produced or human-made objects has problems of its own. It easily collapses into the plan-centred view, which would mean that it is equally useless in developing the non-naturalness intuition; and it does not lead to a sufficiently flexible and discriminative characterisation of artefact types or kinds. In the final step, we therefore propose a division of labour between the non-naturalness intuition and our use-plan analysis: among objects in general, artefacts are distinguished as human-made objects; within the domain of artefacts, types or kinds are characterised by the use plans in which artificial objects are embedded. Only the combination of these two views provides a solid basis for an analysis of the nature of artefacts.

1. Function essentialism

The view that artefacts are primarily or even essentially functional objects is almost a commonplace among philosophers who attempt to characterise artefacts. Still, this high regard for functions has not led to detailed interest: just like artefacts themselves, artefact functions are not often considered to warrant specific analysis. Authors may, on the one hand, commit themselves to an apparently strong claim, namely essentialism with regard to artefact functions, but may on the other hand characterise these essences in a fast and loose way,³ or not at all. Philosophers who seek to analyse the notion of function typically focus on the biological domain, and treat the domain of artefacts as a relatively unproblematic spin-off.⁴

Detailed attention to artefact functions shows that they are neither unproblematic nor essential. The line of argumentation that we have developed in some earlier publications can be summarised as follows. First, if one wants to account for some basic aspects of artefact use and design in terms of functions, a small set of specifications for a successful theory of functions can be derived.⁵ Most etiological theories of functions – in fact, all such theories currently in existence – fail to satisfy these specifications.⁶ Hence, the most popular general theory of functions does not, at the moment, apply to artefact functions. Second, it is possible to construct a characterisation of artefact functions that does satisfy the small set of specifications. This construction, which we have proposed under the heading of the *ICE-function theory* of artefact functions,⁷ involves action-theoretical, physical, epistemological, and social notions. In concise form, this ICE-theory reads as follows:

An agent a justifiably ascribes the physicochemical capacity to ϕ as a function to an item x , relative to a use plan p for x and relative to an account A , iff:

- I. a believes that x has the capacity to ϕ ;
 a believes that p leads to its goals due to, in part, x 's capacity to ϕ ;
- C. a can justify these beliefs on the basis of A ; and
- E. a communicated p and testified these beliefs to other agents, or
 a received p and testimony that the designer d has these beliefs.

Finally and, for our present purposes, most importantly, this characterisation of artefact functions by means of the ICE-theory undermines function essentialism. The reason is that the ICE-theory subordinates functions to what we called 'use plans'. These plans are series of (considered) actions, which someone takes in order to realise a goal, and which include manipulations of other objects than the body of the plan-executing agent. Thus, a use plan may consist of putting water into a coffee pot, filling a filter with coffee grounds, and other actions, which taken consecutively

realise the goal of making a fresh cup of coffee;⁸ and this plan is a use plan for all the objects involved, such as the percolator and the grounds. On our characterisation, use plans constitute the background of function ascriptions in two important ways. Most obviously, they set the context in which agents ascribe ‘practically relevant’ capacities to artefacts, i.e., capacities that are supposed to contribute to the realisation of a practical goal. Without such a practical context, there are no artefact functions. In addition, the use plan provides a historical context for function ascriptions; some people, the designers of the plan, have selected objects to play a role in that plan, and they have communicated the plan, including the role of the objects, to other people – the potential users of artefacts. This historical context distinguishes artefact functions from, to put it colloquially, ‘other things that the artefact can do’; to give one example, it distinguishes a car’s function to transport people from its polluting and noise-making features.

In short, we have shown that some basic aspects of artefact use and design may be adequately described in terms of functions. But this phenomenological adequacy comes at a price, namely that artefact functions require a background of use plans. Therefore, we can derive a dilemma regarding function essentialism: either functions are not the key concepts in describing basic aspects of use and design, making them phenomenologically non-fundamental; or functions are themselves characterised in other terms and thus conceptually non-fundamental. In other words, staying true to the phenomenology of artefact use and design in a theory of artefacts and their functions compromises one’s commitment to function essentialism.⁹ Thus, grounding one’s characterisation of artefacts on functions only seems wise as long as the issue is not closely considered; otherwise, it soon becomes clear that the chosen cornerstone does not, in fact, uphold the structure.

The phenomenological advantages of the use-plan analysis mainly concern evaluative aspects of use and design. Once functions are defined in terms of use plans, one can start distinguishing between proper and alternative use, and accounting for the possibility of malfunctioning in terms of functions; alternative function theories do not do this job. Accommodating these evaluative aspects requires a reconstruction of artefact use and design and the introduction of explicit standards for these activities. Here, use plans and the possibility of evaluating plans with respect to their (practical) rationality do all the work. Artefact use may be reconstructed as carrying out a use plan. Designing is, primarily, the construction and communication of such a plan and perhaps, secondarily, the description of currently non-existing items that are manipulated while executing this plan; if such a description is an additional goal of the design process, we call this process ‘product designing’ to distinguish it from designing in general.¹⁰

On the basis of these two characterisations of agent activities, several evaluative notions can be defined. Rational or effective use amounts to the execution of a practically rational plan; proper or standard use is the execution of a use plan that has been socially institutionalised, typically because it is constructed and communicated by socially acknowledged designers; and malfunctioning is accounted for by, roughly speaking, the fact that effective and proper use are not necessarily co-extensive: designers may make mistakes that undermine both the effectiveness of a use plan and their status as socially acknowledged experts. The notion of artefact function is redundant in analysing all these evaluative aspects of use and design.

2. Plan relativism

In the previous section we argued that function essentialism is either phenomenologically inadequate or self-defeating, and we defined artefact functions in terms of use plans. This suggests that function essentialism can be replaced with a view that characterises artefacts in

terms of plans. In this section, we test this suggestion, mainly by looking whether a use-plan characterisation can distinguish between artefact types or kinds, and between artefacts and natural objects.

This plan-based characterisation takes artefacts to be objects embedded in use plans, i.e., objects that are justifiably believed to be useful. More precisely, it takes artefacts to be described implicitly in our reconstructions of use and designing. The resulting ‘useful-material’ characterisation of artefacts reads:

An object is an artefact (more specifically, a ϕ -er) if and only if manipulation of the object is part of (executing) a use plan, which is designed, communicated and evaluated in accordance with the use-plan analysis, and in which the object is justifiably ascribed the capacity to ϕ as a function.

Hence, fresh, running water is an artefact, more specifically a cooling agent, in the context of generating nuclear electricity, just as it is an artefact, more specifically a cleaning agent, in the context of washing one’s hair. In the same way, 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 view, everything to which our use-plan analysis applies is an artefact by definition.¹¹ One may at this point object that fresh water is clearly not an artefact. We acknowledge this when we confront the plan-based characterisation with the non-naturalness intuition later.

On this characterisation, artefacts have a plan-relative nature. Aspirin, for example, is nowadays produced, marketed and used for two different purposes: to alleviate pain by taking an incidental, high dosage, and to prevent cardiovascular problems by taking a daily, lower dosage. These ways of using Aspirin involve different use plans, because they have different goal states. Consequently, on our useful-material characterisation, a tablet of Aspirin is a painkiller when it is swallowed in the context of executing the more traditional use plan to alleviate pain. But the very same object is a blood thinner when it is swallowed in the context of executing the more recently designed and communicated plan to prevent blood clots.

In this way, the useful-material definition can be used to determine the persistence conditions for artefacts – one of the standard tasks of the ontology of a domain. The persistence conditions of useful-material artefacts 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. Consequently, determining artefact persistence becomes as context-sensitive as determining the differences between plans. One might, for example, be inclined to say that an object ceases to be a ϕ -er as soon as it becomes impossible, or at least generally irrational, to execute the use plan to which the object contributed by ϕ -ing. A car that is wrecked beyond repair in a crash is then no longer a car, but a twisted aggregate of steel and plastic. Similarly, when the white, fresh-smelling stuff 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 use it as such – and becomes filler instead. In both cases, the persistence of the artefacts is intimately connected to realising goals by means of material objects.

This indicates how a useful-material ontology of artefacts may be developed. We now explore some further features of our characterisation. It turns out that, although all features can be

presented as problems, none of them shows that the useful-material characterisation is indefensible or unstable.

One feature concerns the identity criteria of artefacts. Like their persistence conditions, these criteria must derive mainly from the use plans for artefacts. Although such plans might be told apart on the basis of having different goal states or different actions in different orderings, these criteria are far from exhaustive. Artefacts, when defined in terms of use plans in which they are manipulated, would be rather indeterminate entities. Suppose, for instance, 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 a tea bag 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; the answer would depend on intuitions about proper use, personal beliefs about the quality of the tea produced and other context-sensitive factors.

Quine's slogan 'No entity without identity', which is widely adopted as a barrier for candidate abstract objects,¹² may 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, one must seek another definition or admit that the sought category is ontologically disreputable. In these lights, plan-relative artefacts can be discredited because they are inherently vague – just like mental properties and events. Those who sympathise with Quine's physicalism and general distrust of abstract terms¹³ may argue that this metaphysical Puritanism is as bloodless as trimming a very unkempt beard. Others might conclude that the program sacrifices too many *bona fide* entities to an elusive ideal of metaphysical rigor and parsimony, and might therefore not object to such vague objects as plan-relative artefacts. We just note that this feature of our useful-material definition is not necessarily an objection.

A second feature is that the plan-relativity of artefacts may be mitigated by considerations of proper use. Some use plans are socially privileged and play a role in a network of responsibilities and requirements; other use plans are merely recommendations for use that feature in useful knowledge. Categorising an object as token of a functional artefact type or kind (i.e., as token of the ϕ -ers) may be one way of expressing that, among all known plans, a specific use plan is privileged; the useful-material characterisation is easily modified to accommodate this suggestion.

This appeal to proper use, which is part and parcel of our evaluative approach to artefacts, limits the relativity of useful-material artefacts. It may, for instance, be said that a screwdriver is not a paint-can opener, although it is occasionally used as such. For other items and in other situations, however, relativism still looms large. The river Rhine, for instance, is in some places properly used as an industrial 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, 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 for recreational purposes in the other. Similarly, items that are equipped with ever more functionalities, such as cell-phones and organisers, would be many things simultaneously: they play a role in many determinate use plans, or none clearly, and their context of use is very open-ended.

This remaining plan relativism of artefacts may be taken as a curious consequence rather than a *reductio* of the useful-material definition. 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 ϕ -er without being the same ψ -er (e.g., the same cooling agent, but not the same recreational artefact), is sustainable, and whether it is needed to account for other language use. It would be fair to say that relative identity is believed to be innocent only by a minority, although the jury still has to reach unanimity.¹⁴

Although both of the features discussed raise problems, these are the topics of current and unresolved debates in general metaphysics. They are not internal problems such as those that we uncovered for function essentialism in the previous section. More specifically, they do not offer grounds to suppose that the useful-material characterisation is either conceptually or phenomenologically non-fundamental. Hence, we may tentatively conclude that it is a defensible, albeit problematic successor of function essentialism.

This means that we can turn to our original question of combining the two intuitive starting points of characterisations of artefacts: functionality and non-naturalness. Now that the former has been included in the useful-material characterisation of artefacts, the tension with the latter intuition becomes clear. For it is impossible to make a fixed and principled distinction between useful-material artefacts and natural objects. As soon as there is a designed, communicated, and justifiable use plan in which an object is manipulated, that object is an artefact. Hence, there is at best a fluid, agent-dependent or community-dependent line between natural objects and artefacts. Natural objects become artefacts as soon as an agent constructs and communicates a use plan for them, and artefacts cease to exist once an agent (or a community of users) discards or forgets this plan.¹⁵

Defining artefacts as useful materials leads to a kind of global instrumentalism, since there is hardly any object, natural or non-natural, which we cannot imagine using deliberately for some purpose. Just like most of us know how to use staplers to join stacks of paper, some of us know how to use planets to accelerate interplanetary probes. Others can use their fingers for playing 'Für Elise'. We may even be said to use taxi drivers to get from the airport to a hotel. This view is not entirely without precedent in philosophy. Heidegger's (1977) question concerning technology, or at least his apparent answer, leads in a similar direction: he uses power plants on 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*). But despite this precedent, pan-instrumentalism is clearly at odds with the intuitive understanding of artefacts as non-natural, which places them in a more or less fixed and limited domain.

One may conclude from this that the non-naturalness intuition is incompatible with the intuitions about use and design that the useful-material definition can accommodate. Or one might consider the possibility to devise other categories within the use-plan analysis to make relevant distinctions within the all-encompassing class of artefacts (e.g., 'improvised', 'permanent', 'goal-enabling', 'self-producing'), and thus save all intuitions simultaneously, in a roundabout way. We choose to regard the conflict between functionality and non-naturalness as a challenge for further analysis. Apparently, the non-naturalness intuition sets effective boundaries for understanding the nature of artefacts, in the sense that a characterisation that seems promising in other respects fails to meet it. However, further analysis is needed to go beyond this negative conclusion. After all, the match is not even: the non-naturalness of artefacts must be developed beyond the intuitive stage to find out exactly how this feature conflicts with the useful-material definition, and to use this conflict more productively.

3. Human-made material

It may not come as a surprise that the plan-based characterisation of artefacts conflicts with a key intuition. In fact, the problem appears to precede the useful-material definition at the start of the previous section. The notion of designing that we introduced at the end of section 1 is already at odds with an intuitive characterisation of this activity – a characterisation that is intimately related to the understanding of artefacts as non-natural objects.

Our notion of designing puts the spotlight on the construction and communication of use plans, and turns the description of new objects into a secondary, optional activity. We emphasise the instrumental side of designing, its contribution to practical purposes by providing means, over its productive side. This shift of focus has many advantages. Yet these are bought at the price of the intuition that designing is primarily productive. This intuition is easily converted into one about artefacts by defining artefacts as the products of designing. The intuitive understanding of designing then yields the characterisation of artefacts as non-natural, i.e., intentionally produced objects.

This means that the non-naturalness intuition can, like the useful-material definition above, be developed in terms of intentional actions, as follows:

An object is an artefact (more specifically, a ϕ -er) if and only if the object has been intentionally produced for the purpose of ϕ -ing by an agent.

Let us call this the ‘human-made’ characterisation of artefacts, although it leaves open the possibility that bird nests and beaver dams are artefacts, in case non-human animals can be called ‘agents’.

Non-naturalness is now rephrased in terms of an action, namely production. In itself, this does not clarify the relation between the functionality and non-naturalness intuitions. It does, however, shift it entirely to the realm of actions. This shift turns our main question from one concerning a possible conflict between two central intuitions about artefacts to one concerning the best way of bringing out the importance of two types of activities – instrumental and productive – in a characterisation of artefacts.

Furthermore, the human-made definition has three problems: it is insufficiently flexible and insufficiently discriminative as a characterisation of artefact types or kinds, and it easily collapses into the plan-centred view of the previous section, once further inquiry is made into the meaning of ‘production’.

The inflexibility problem is closely related to the considerations about proper use discussed earlier. Like function essentialism, a characterisation of artefacts that focuses exclusively on their history of production cannot accommodate the dynamics of (proper) artefact use. Many artefacts, such as Aspirin, acquire new uses that co-exist peacefully with the old ones; others, such as plate armour, acquire uses that replace their traditional purpose. For each of these artefacts, there is a highly specific story to be told about the establishment of its new use in a small or larger group of users. The use-plan analysis, which divorces designing from the production of artefacts and therefore does not identify designers with creators, allows for such changes in the characterisation of artefacts, artefact functions, and artefact use. However, the human-made definition of artefacts

rules out in advance such changes in the nature of artefacts, because the original production of an artefact determines it to be a ϕ -er once and for everyone.¹⁶

The discrimination issue shows up once one looks into the notion of intentional production. Developing an appropriately hands-on notion of producing or making is no mean feat.¹⁷ If one requires an artefact to be produced from raw materials, then products of assembly, such as the stereotypical driftwood raft, are not artefacts. Alternatively, if producing implies that something is physically changed to serve a practical purpose, a prototypical natural object such as Jupiter is an artefact on the human-made definition: using this planet to accelerate probes changes its orbit ever so slightly. Furthermore, every episode of use results in wear and tear, leading to an enormous variety of use-related physical modifications. These physical changes are, of course, side effects rather than the goal of production or use. Yet spelling out this intuitive judgment requires a distinction between intentions that are relevant and those that are irrelevant to the production process. Such a distinction between relevant and irrelevant intentions is also needed to develop the human-made definition as an account of artefact types or kinds. Some cell-phones may have been produced to enlarge a company's share of the cell-phone market, but this should not lead us to classify such cell-phones as market-share-enlargers, in one category with cars and laptop computers that were produced for the same purpose. The use-plan analysis is the only account currently on offer that avoids this problem while retaining the reference to agents' intentions.

The collapse problem is an immediate corollary of the last point. It turns out to be difficult to analyse the notion of 'production' in such a way that its difference from the useful-material definition is retained. Suppose we develop the useful-material definition as sketched in the previous section. Then, an artefact comes into existence as soon as an agent constructs and communicates a use plan in which an object is to be manipulated. It does not matter whether the object existed before this manipulation: the artefact, if not the material that constitutes it, is created by its inclusion in a use plan. The Aspirin case, for example, can be analysed in terms of a quantity of acetylsalicylic acid, which once exclusively constituted a painkiller, but may now also constitute a blood thinner, as soon as another use plan is designed and executed. Hence, as long as the human-made definition does not get its hands dirty, i.e., as long as it says nothing about physical modifications of an object and the intentions involved in that activity, it may collapse into the useful-material definition.

To illustrate this point, consider the following condition for being an artefact, proposed by Hilpinen (1992): 'An object o is an artefact 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 '. At first glance, this condition seems to express the human-made definition. However, it can also be interpreted as expressing the useful-material definition: D may be identified with the goal-contribution term ' ϕ -er' (e.g., 'cooling agent'), the productive intention I_A with the intention that is central to our general notion of designing, i.e., to contribute to other agents realising their goals, and the 'bringing into existence' may be explained by means of the constitution view.

Despite these problems, the human-made definition has one marked advantage over the useful-material definition: it provides a basis for distinguishing between artificial and natural objects. Determining whether something is an artefact or not, i.e., identifying something as an artefact, rather than as a token of an artefact type or kind, 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 the abilities of an agent or group of agents (e.g., trained archaeologists). If it is possible to develop

the notion of production in such a way that it is clearly distinct from the use-plan notion of design, characterising artefacts as intentionally produced provides the requisite absolutism: on this characterisation, an object becomes an artefact on its original production, once and for everyone.

4. Producing a useful combination

The advantages and drawbacks of the useful-material and human-made definitions, and their mutual focus on intentional actions, suggest that a combination of both may provide a satisfactory characterisation of artefacts. The human-made definition serves, to some extent, to make an absolute distinction between artefacts and natural objects. However, it cannot be developed to distinguish types or kinds of artefacts without collapsing into the useful-material definition. This latter definition can be used to characterise types or kinds of artefacts, albeit in a highly context-sensitive manner; yet it performs poorly in distinguishing artificial and natural materials. Although the human-made definition still has to be developed into one that makes a fully plausible distinction – the driftwood-raft and Jupiter examples indicate some of the problems awaiting in this process – it does show considerably more promise than the useful-material characterisation. Thus, the two views may be complementary, and may account for both the functionality and non-naturalness intuitions simultaneously. The characterisation of artefacts as human-made objects might be used to single out within the domain of application of the use-plan analysis the domain of artefacts (understood as non-natural objects), in which artefact types or kinds can then be defined by means of the useful-material definition (which develops the functionality intuition). This retains relativism at the level of artefact types or kinds, without undermining an absolute distinction between artefacts and natural objects.

The human-made and useful-material definitions may be integrated into the following ‘produced-to-use’ definition:

An object is an artefact, more specifically a ϕ -er, if and only if: (1) the object has been intentionally produced for the purpose of ψ -ing¹⁸ by an agent; and (2) manipulating of the object is part of a use plan, which is constructed and communicated by an agent¹⁹ and in which the object is justifiably ascribed the capacity to ϕ as a function.

In this definition, the productive and instrumental activities appear to play different roles in characterising artefacts, but in practice they may still coincide. The artefact may, but need not have been produced for a different use than that to which it is put later. Thus, Aspirin may have been produced as a painkiller, but it may also be used as a blood thinner – and thus, on the view developed in section 2, constitute a blood thinner. This case requires two groups of designers, who justifiably believe that two artefact capacities contribute to two constructed and communicated use plans. However, the definition allows for cases in which the mentioned agents, as well as the mentioned capacities, are identical. Aspirin, to return to the example, may be both produced and used as a painkiller.

The ‘dual-activity’ characterisation of artefacts still conceals one major analytical challenge. As we argued in the previous section, developing the notion of intentional production requires a distinction between relevant and irrelevant intentions. There is reason to suppose that this distinction must rely on the notions used in the use-plan analysis, but this reliance should not be so great that production becomes equivalent to design. This leads to significant constraints on developing an appropriate notion of production – constraints that we can just note in this paper.

Whatever the results of analysing production may be, questions about the nature and classification of artefacts are, on our proposed definition, intimately connected to questions about agents and their beliefs and activities. Productive activities may be followed by different use-oriented activities: artefacts may be redesigned by engineers or by users, who might promote uses that supplement or replace those envisaged by the original designers. This may lead to reclassification of the artefacts, or to classifying them as hybrid objects. These issues can, moreover, not be resolved by armchair metaphysics: one needs to consider, among other things, how a community of designers and users standardises artefact use. The conflict between the non-naturalness intuition and the plan-centred characterisation – which succeeds the functionality intuition – is thereby moved to the area of intentional actions, namely production, design and use.

In sum, it is possible to combine the human-made and useful-material definitions and their underlying activities into an encompassing, produced-to-use definition. This definition may be regarded as the successor of plan relativism, which was itself a successor of function essentialism. The produced-to-use definition is phenomenologically adequate in the sense that it accommodates both the intuitions about artefact use and design that proved fatal to function essentialism, and the non-naturalness intuition that caused problems for plan relativism. As a third-generation definition, it provides a broad and diverse basis for a characterisation of artefacts; its basic elements are proper and effective use, the physical structure supporting this use, and the history of design and/or production of the item.²⁰ It does not, however, contain artefact functions as an essential element: unlike the use-plan analysis, which is still an integral part of the produced-to-use view of artefacts, the focus has shifted from functions to actions in the course of our reflections.

These reflections have shown a number of things. We argued that the intuition that artefacts are non-natural objects conflicts with the plan-relativist characterisation of the nature and classification of artefacts, which is the natural successor of function essentialism. On a more positive note, we argued that plan relativism and the non-naturalness intuition may be combined in what we called the ‘produced-to-use’ definition. On this definition, productive and instrumental activities may jointly provide a characterisation of artefacts. There is a natural division of labour between these two activities: production distinguishes artefacts from other types of objects, and instrumental actions carve out artefact types or kinds within the domain set by production. Finally, we showed that there are significant constraints on developing the appeal to production in this dual-activity characterisation. Philosophers may have examined the nature of artefacts in some detail, but a satisfactory analysis of production is still in its infancy.

Acknowledgements

Research by Wybo Houkes and research by Pieter Vermaas was supported by the Netherlands Organisation for Scientific Research (NWO).

References

- Ariew, A., R. Cummins and M. Perlman, eds. 2002. *Functions: New Essays in the Philosophy of Psychology and Biology*, Oxford: Oxford University Press.
- Baker, L.R. 2000. *Persons and Bodies*, Cambridge: Cambridge University Press.
- Baker, L.R. 2004. “The Ontology of Artefacts,” *Philosophical Explorations*, 7: 1-14.
- 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*, Oxford: Oxford University Press, 131-148.
- Bloom, P. 1996. "Intention, History, and Artefact Concepts," *Cognition*, 60: 1-29.
- Bloom, P. 1998. "Theories of Artefact Categorization," *Cognition*, 66: 87-93.
- Elder, C.C. 2004. *Real Natures and Familiar Objects*, Cambridge, MA: The MIT Press.
- Elder, C.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*, Oxford: Oxford University Press, 33-51.
- 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.
- Grandy, R. 2007. "Artifacts: Parts and Principles," in: E. Margolis and S. Laurence, eds., *Creations of the Mind: Theories of Artifacts and their Representation*, Oxford: Oxford University Press, 18-32.
- Heidegger, M. 1977. "The Question Concerning Technology," in: *The Question Concerning Technology and Other Essays*, New York: Harper and Row.
- Hilpinen, R. 1992. "Artefacts and Works of Art," *Theoria*, 58: 58-82.
- Hilpinen, R. 2004. "Artifact," in: E.N. Zalta, ed., *The Stanford Encyclopedia of Philosophy* (Fall 2004 Edition), Available [Online]: <http://plato.stanford.edu/archives/fall2004/entries/artifact>
- Houkes, W. and A.W.M. Meijers. 2006. "The Ontology of Artefacts: The Hard Problem," *Studies in the 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 Artefacts," *The Monist*, 87: 52-71.
- Houkes, W. and P.E. Vermaas. 2009. *Useful Material: An Action-Theory of Artefacts and Their Functions*, Manuscript.
- 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.
- Ingold, T. 2000. *The Perception of the Environment: Essays in Livelihood, Dwelling and Skill*, London: Routledge.
- Kornblith, H. 1980. "Referring to Artefacts," *The Philosophical Review*, 89: 109-114.
- Krohs, U. 2009. "Functions as Based on a Concept of General Design," *Synthese*, online first.
- Longy, F. 2009. "How Biological, Cultural and Intended Functions Combine," in: U. Krohs and P. Kroes, eds., *Functions in Biological and Artificial Worlds: Comparative Philosophical Perspectives*, Vienna Series in Theoretical Biology, Cambridge, MA: MIT Press: 51-67.
- Losonsky, M. 1990. "The Nature of Artefacts," *Philosophy*, 65: 81-88.
- Lowe, E.J. 1995. "The Metaphysics of Abstract Objects," *The Journal of Philosophy*, 92: 509-525.
- Millikan, R.G. 1984. *Language, Thought and Other Biological Categories*. Cambridge, MA: The MIT Press.
- Millikan, R.G. 2000. *On Clear and Confused Ideas*. Cambridge: Cambridge University Press.
- Neander, K. 1991. "The Teleological Notion of "Function"," *Australasian Journal of Philosophy*, 69: 454-468.
- Preston, B. 1998. "Why is a Wing like a Spoon? A Pluralist Theory of Function," *The Journal of Philosophy*, 95: 215-254.
- Preston, B. 2000. "The Functions of Things: A Philosophical Perspective on Material Culture," in: P. Graves-Brown, ed., *Matter, Materiality and Modern Culture*, London: Routledge, 22-49

- Preston, B. 2003. "Of Marigold Beer: A Reply to Vermaas and Houkes," *British Journal for the Philosophy of Science*, 54: 601-612.
- Quine, W.V.O. 1960. *Word and Object*, Cambridge, MA: The MIT Press.
- Quine, W.V.O. 1969. "Speaking of Objects" (original: 1958), reprinted in: *Ontological Relativity and Other Essays*, New York: Columbia University Press, 1-25.
- Simons, P. 1995. "Artefact," in: J. Kim and E. Sosa, eds., *A Companion to Metaphysics*, Oxford: Blackwell, 33-35.
- Sperber, D. 2007. "Seedless Grapes: Nature and Culture," in: E. Margolis and S. Laurence, eds., *Creations of the Mind: Theories of Artifacts and their Representation*, Oxford: Oxford University Press, 124-137.
- Vermaas, P.E. and W. Houkes. 2003. "Ascribing Functions to Technical Artefacts: A Challenge for Etiological Theories of Functions," *British Journal for the Philosophy of Science*, 54: 261-289.
- Vermaas, P.E. and W. Houkes. 2006. "Technical Functions: A Drawbridge between the Intentional and Structural Natures of Technical Artefacts," *Studies in the History and Philosophy of Science*, 37: 5-18.
- Wiggins, D. 2001. *Sameness and Substance Renewed*, Cambridge: Cambridge University Press.

Endnotes

- 1 Our focus is on *technical* artefacts, material objects that are typically designed by engineers and used for immediate practical purposes. Hence, our analysis is meant to apply to screwdrivers and cell-phones; we leave it to others to determine whether or not art works such as Michelangelo's David and social artefacts such as the First Amendment fall inside its scope.
- 2 Function essentialism is not exclusive to philosophers: in cognitive science, where artefact categorisation has recently become a topic of interest, it is maintained by Paul Bloom (1996; 1998).
- 3 David Wiggins claims, for instance, that '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.' (2001, p. 87). The first full sentence expresses a commitment to function essentialism, the second loosely (and not too consistently, since rigidity describes the constitution of objects) characterises the notion of artefact function via two examples.
- 4 In, for instance, the introduction to (Ariew, Cummins and Perlman 2002), a volume on the topic that deals by its subtitle explicitly with biology and psychology, functions are introduced as entities that typically and unproblematically pertain to artefacts.
- 5 The specifications, derived in Vermaas and Houkes (2003), are: (1) that the theory distinguishes proper functions from accidental features; (2) that the theory leaves open the possibility that artefacts malfunction, i.e., do not fulfil their function; (3) that the theory makes explicit the relation between artefact functions and underlying physical characteristics; (4) that the theory leaves open the possibility for radically innovative functionalities.
- 6 Beth Preston (2003), in defence of her pluralistic etiological approach, argues that one might give up on (one of) the specifications.
- 7 An early version of this ICE-theory is presented in the final section of Vermaas and Houkes (2003). An improved version is given in Houkes and Vermaas (2004) and, in increasing detail, in Vermaas and Houkes (2006) and Houkes and Vermaas (2009).
- 8 Some more details of this use plan, along with many other examples, can be found on 'How To' websites such as www.wikihow.com.
- 9 This result is conditional: if there would be theories of artefact functions that satisfy our specifications and in which functions are primitive features (i.e., not relative to use plans or other items), function essentialism may be upheld.
- 10 An early version of the use-plan analysis of artefact use and design can be found in Houkes *et al.* (2002); a modified and less schematic account is presented in Houkes and Vermaas (2004).
- 11 The useful-material definition of artefacts may be developed into a full-fledged ontology by means of Lynne Baker's (2000; 2004; 2007) constitution view. One of us has argued elsewhere that Baker's own application of the constitution view to artefacts rests on function essentialism, and has problems dealing with the phenomenology of artefact use (Houkes and Meijers, 2006).
- 12 The slogan is originally found in Quine (1969, p. 23). For a recent defence of its application, see Lowe (1995).
- 13 An expression of this distrust is: 'Once we start admitting abstract objects, there is no end.' (Quine 1960, p. 123)

- 14 The thesis of relative identity is primarily associated with Geach (1967), and is vehemently attacked by, among others, Wiggins (2001, ch. 1). Garbacz (2002) has recently presented a proposal for the logical representation of statements of relative identity.
- 15 This result is not specific to the plan-based view; it also holds on a view in which artefacts are nothing but functional items. There, natural objects – e.g., water in the river Rhine – become artefacts as soon as they acquire functions.
- 16 Preston (2003) presents a similar argument to criticise theories that characterise artefact functions exclusively in terms of author's intentions.
- 17 The lack of appropriate notions of production, as well as the problems in developing them, are discussed by, e.g., Preston (2000) and Ingold (2000; chs. 18 and 19).
- 18 The capacities to ϕ and to ψ may, but need not be, identical.
- 19 The agents mentioned in the characterisation may, but need not, be identical.
- 20 In philosophy, a similar starting point is suggested 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 *et al.* (2004) have proposed the HIPE-theory. On this theory, design History, Intended use, Physical structure, and the Events resulting from use are the relevant elements in artefact categorisation. Again, production (and therefore a way of accommodating the non-naturalness intuition) is lacking.

What is an Artefact Design?

Pawel Garbacz
Faculty of Philosophy
The John Paul II Catholic University of Lublin

Abstract

The paper contains a first order formal theory pertaining to artefact designs, designs which are construed as the results of designing activities. The theory is based on a minimal ontology of states of affairs and it is inspired by the ideas of the Polish philosopher Roman Ingarden. After differentiating the philosophical notion of design from the engineering notion of design specifications, I then go on to argue that the philosophical category of artefact designs may be compared with Ingarden's category of intentional states of affairs. At least some artefacts are found to be determined by more than one design. I also show how this ontological framework allows for the distinction between artefact tokens and artefact types. That leads to a proposal on how to define a criterion of identity for artefact types. The proposed theory serves as a basis both for a better understanding of what artefacts are and for the construction of computer-readable models of design specifications.

Keywords: artefact, design, ontology, representation

The category known as design is a crucial notion in engineering science. In engineering and in analyses of this notion the focus tends to lie predominantly on the process of designing, rather than on the results or products of design. Engineering theories of design inform us about the stages and principles of design, but they are not so explicit when it comes to the nature of the design specification itself.¹ Needless to say, philosophical reflections on the ontic structure of designing and its products are even scarcer. It might be argued that the lack of analyses of design products hinders ontological investigation into the nature of artefacts both in engineering and philosophy since it is the design specifications that determines, or at least co-determines, the intentional aspects thereof. Given the suitable level of granularity of our descriptions of artefacts it may be asserted that they are what they are due to their design specifications. In this paper I will analyse the products of designing in order to contribute to the resolving of philosophical puzzles about the identity of artefacts and to aid the engineering search for adequate representations of design data. What I have in mind is the approaches in engineering design that resort to the idea of engineering ontologies that can provide the foundations for computer-readable representations of design specifications (e.g. Alberts (1994), Gruber and Olsen (1994), Chandrasekaran et al. (1999), Kitamura et al. (2002), to name but a few).

The principal aim of this paper is to give an ontological account of designs that are construed as products of design activity. It is an essay on slightly revisionary metaphysics. In other words, instead of attempting to define the standard meanings of the defined notions, I will suggest a new framework within which artefacts and their designs can be considered. By approximating some fundamental intuitions on the production process, the framework will to be evaluated in the light of its consistency and explanatory power. The resulting theory will be rendered in first-order predicate language. The ontological foundation for this enterprise remains the philosophy of Roman Ingarden. Provided that this account is adequate, it will provide the grounds both for a

better understanding of what artefacts are and for the construction of computer-readable models of design specifications. The main results comprise:

- separation of different modes of artefact representations,
- distinction between the engineering and philosophical notion of design,
- conceptual framework to explicate the variable granularity of designs,
- distinction between artefact tokens and the artefact types the account suggests against a criterion of identity for artefact types.

For the sake of simplicity, the temporal aspects of artefact designs will be ignored. This constraint will force me to focus not so much on the process of designing as on the product of designing. All temporal indices in the respective definitions and theorems will therefore also be omitted. The constraint does not mean, however, that artefact designs will be construed in such a way as to exclude the behavioural or functional descriptions of artefacts.

The theory I expose here is a rightful part of a more comprehensive approach known as the four-dimensional ontology of artefacts. In this ontology, artefacts are characterised according to their purposes, designs, the background knowledge in which the designs are embedded and the instructions of use. The details of this approach may be found in (Garbacz 2004). Here I will only briefly note that my view of artefacts is convergent with the Constitution View, expressed in this issue by Lynne Baker, which states that artefacts are intention-dependent objects that are constituted by physical entities.

1. Artefacts and designs

When an engineer speaks about design, he or she usually has in mind some kind of designing activity. The standard view is that any such process consists of four stages: the clarifying of the design task, the conceptual design, the embodiment design and the detailed design (cf. Pahl and Beitz (1996)). In this paper I will not focus on such activities but rather on their *direct outcomes* which, in engineering design, are dubbed specifications.

There are a huge number of types of design specifications. There are cooking recipes, electrical schemes, elevation drawings, flow diagrams, musical scores, point paper designs, etc. Peter Simons and Charles Dement described the structure of the realm of designs by distinguishing between engineering, manufacturing and logistic bills for materials. When I speak in this paper about designs, I will be referring solely to those designs that correspond directly to engineering bills for material, or to those designs that specify the details of ready-to-use artefacts. Thus, the so-called manufacturing bills of materials and the logistic bills of materials are not designs in the sense used in this paper (cf. Simons and Dement (1996, pp. 267-268)).

The variety of engineering bills of materials is, however, still enormous. What do all these designs have in common? What sort of mental structures, if any, are eligible for artefact design? What is the relation between the design of an artefact and the physical structure of the artefact?

Artefacts are products of human action. The human production of artefacts is something that is intimately linked to beliefs, desires and intentions. The representational content of these mental

structures determines the relevant features of artefacts while co-determining certain phases of the production processes. In short, artefacts are design-based physical objects. In this paper I assume that this old Aristotelian view of artefacts is adequate also for contemporary kinds of artefacts, provided that we have a broad enough understand of the details of the picture.

Artefacts are designed and used by intentional agents. Every object capable of bringing about some state of affairs will be called an *agent*. If an agent has certain beliefs and wishes then it will be called an *intentional agent*. I assume that for every domain of artefacts, there is a community of intentional agents whose beliefs about the artefacts from that domain are reliable. I will assert that members of such a community are *rational agents*. In particular, if a rational agent believes that some artefact is designed to have certain features then her belief will be true.

What are designs? I will start with the classification problem: What is the ontological category of designs? As ordinary objects all artefacts have certain properties, take part in various processes, occur in events, etc. Thus, if we were asked to characterise them, we might come up with a set of sentences specifying the properties that the artefact has, the processes in which it takes part, etc. Every such sentence will refer to a state of affairs in which the artefact occurs. For example, specifying one feature of a resistor the sentence ‘This resistor has the resistance of 10 kΩ’ will refer to the state of affairs of the resistor having a resistance of 10 kΩ. The “fusion” of all states of affairs in which an object occurs may be called the ontic range of objects.

Artefacts are products of human activity that essentially involve mental components. Apparently, artefact designs somehow represent artefacts. If one says that x represents y , one might mean one of two things. First, one might wish to express the fact that x is an image of or mirrors y , in which case it is y that comes first and x that is presumed to imitate y . Here, it is x that is evaluated as a faithful or unfaithful representation of y . If x represents y in this sense, I will say that x *represents₁* y and call such a relation the relation of *epistemic representation*. Secondly, saying that x represents y one might wish to express the fact that y is an image of or mirrors x , in which case it is x that comes first and y that is supposed to imitate x . Here, it is y that is evaluated as a faithful or unfaithful representation of x . If x represents y in this sense, I will say that x *represents₂* y and call such a relation the relation of *poietic representation*. If x represents₁ or ₂ y , I will still claim that x *represents* y . Admittedly, in claiming that artefact designs represent artefacts, I am asserting that the former represent₂ the latter.

Usually there is nothing in the content of a representation of an entity that makes it a representation₁ or a representation₂ of that entity. It is a user of a given representation that determines that by using it either as an epistemic or poietic representation. Moreover, it is not so that any representation₁ may be viewed as a representation₂; however, any representation₂ may become a representation₁. In our case it is a community of rational agents that determines precisely which representations are poietic representations of artefacts.

My claim that artefact designs represent₂ artefacts is the weakest claim among all the claims based on the Aristotelian view of artefacts. In order to explain this metatheoretical remark, two additional qualifications of the relation of representation need to be given. I will say that x *truthfully represents* y iff every representational detail of x corresponds to some detail of y . I will also say that x *adequately represents* y iff every detail of y corresponds to some representational detail of x .

- (1.1) (i) (At least) some designs truthfully represent₂ some artefacts.
 (ii) For any artefact x , no design adequately represents₂ x .

Later on 1.1(i) will be strengthened to the “at most some” condition. 1.1(ii) accounts for the fact that human designs are products of finite minds that are unable to describe all the details of any their products. Consequently, a design of an artefact never specifies the *whole* ontic range of any one artefact.

Designs are like other representation of reality: they contain both objective and subjective components. An electrical plan is a depiction plotted in ink or chalk (the objective component of the design) by means of which a rational agent represents₂ an electrical device (the subjective component of the design). Generally speaking, designs are physical objects of various kinds (inscriptions, drawings, pictures, sounds, etc.) that are considered by rational agents to be representations₂ of other physical objects. With any artefact, the objective factor of its design may vary while the subjective factor remains constant. One might represent₂ an electrical device by means of a drawing, a complex depiction or a series of sounds, but all these physical entities represent₂ one and the same artefact. This means that the drawing, the depiction and the sounds may be perceived as representations₂ of the artefact that represent₂ it in the same way, in other words, as providing the same poietic information on that artefact. The following distinction would thus seem to be useful.

(1.2) An *engineering specification* of an artefact is a physical object considered by a rational agent to be a representation₂ of the artefact.

(1.3) A *design* of an artefact is an entity by which means a rational agent considers different engineering specifications to be representations₂ of the same artefact.

When a rational agent identifies by means of design y different engineering specifications z_1 and z_2 as representations₂ of the same artefact, I will say that z_1 and z_2 (*physically*) support y .

The foregoing remarks should make it clear that every engineering specification supports at least one design. Assuming that designing is a rational activity, I stipulate that the relation of support is a function or that no engineering specification supports more than one design.

The design of an artefact partially determines what the artefact is like, that is to say, it partially determines the states of affairs in which the artefact is involved.² This claim presupposes that artefact designs have a propositional structure, in other words, it presupposes that in principle any design might be supported by a complex inscription (i.e., a mereological sum of sentences³). Obviously, designs are not sum totals of sentences because when a designer constructs a design she does not need to formulate a sentence and a great number of actual designs are not (and presumably will not be) supported in this way. However, since it is sufficient for a sentence to exist if the language of which this sentence is a part exists, I assume that for any engineering specification, there must be a mereological sum of sentences, the sum of which supports the same design as this specification. This assumption is not as solid as it looks if we concede that drawings or charts are (or consist of) *sui generis* sentences.

What, then, are designs? It is easy to observe that designs are neither engineering specifications nor mereological sums of engineering specifications. Regarding the former claim, many designs are supported by more than one engineering specification, none of which has the privileged position of being *the* engineering specification of a given artefact. Secondly, you can multiply engineering specifications, for instance, by photocopying them, without multiplying designs.

Thirdly, you can change or even destroy an engineering specification without changing the design that is supported by the engineering specification. As for the latter claim, it should be noticed that the mereological sum of a set of objects changes when one of the objects is changed. We cannot therefore change an engineering specification without changing the sum of the engineering specifications to which this engineering specification belongs. Consequently, if designs were mereological sums of engineering specifications, then any change in any engineering specification would precipitate a change in design. For similar reasons a design cannot be identified with a distributive set of engineering specifications.

In this paper I will explore the solution to the effect that artefact designs belong to the category of intentional states of affairs introduced by the Polish philosopher Roman Ingarden. Let me briefly sketch the main ideas of his conception (cf. Ingarden (1964; 1965)).

2. States of affairs

If one slightly simplifies the issue, one might say that Ingarden distinguishes two different ontological categories of states of affairs: real and intentional. A *real state of affairs* may be defined as a part of the ontic range of an object.⁴ For instance, if John is handsome, the fact that John is handsome is part of his ontic range since the property of being handsome inheres in John. Similarly, the fact that Paul hates Barbara and Mary runs are states of affairs. Knowing about such a state of affairs is part of the ontic range of an object, I know what the object is like, in what process it is involved or when it occurs. If a state of affairs *x* is a part of the ontic range of an object *y*, I will say that *y occurs in x*.

On the matter of intentional states of affairs, Ingarden says that due to its content every representational state of mind determines an intentional state of affairs. Derivatively, due to its meaning, every declarative sentence determines an intentional state of affairs. For example, if John believes that George W. Bush is a Catholic and if he discusses with a friend the impact of this fact on Bush's international politics, then both refer to the same entity, but because John's belief is false, no real state of affairs corresponds to it. Intentional entities, including intentional states of affairs, function as intermediate entities between mind (or language) and reality. Just as real objects occur in real states of affairs, so intentional objects occur in intentional states of affairs.⁵ As a rough approximation, one might identify the category of intentional states of affairs with the category of propositions, provided that the latter are construed as mind-dependent entities. Therefore, the category of intentional states of affairs cannot be reduced to the category of sentences because two different sentences may determine the same intentional state of affairs, but one of the two may well cease to exist without the intentional state of affairs ceasing to exist. For a similar reason, the former category cannot be reduced to the category of representational states of mind.

What, then, is the relationship between the realm of real states of affairs and the realm of intentional states of affairs? Assume that an intentional agent formulates some kind of sentence. If the sentence is true, the intentional state of affairs determined by that sentence will *mirror* some actual state of affairs (i.e., a *fact*) in the ontological universe. Ingarden adds that the intentional state of affairs and its actual counterpart share the same content, but differ in their mode of existence: actual states of affairs exist autonomously and intentional states of affairs exist heteronomously. He explains that existing autonomously involves having ontic foundations in itself while existing heteronomously means having ontic origins in something else. For example, the intentional state of affairs that George W. Bush is a Catholic has its ontic roots in John's belief that George W. Bush is a Catholic.

If a representational state of mind or a declarative sentence determines an intentional state of affairs that corresponds to some real state of affairs, I will assert that the state of mind or the sentence represents (either epistemically or poietically) this real state of affairs. I will also claim that the respective intentional state of affairs represents the actual state of affairs. Similarly, I will say that intentional objects represent real objects.

I believe that given the foregoing explanation I am justified in identifying artefact designs with Ingarden's intentional states of affairs. More precisely, artefact designs are construed here as structures composed of those intentional states of affairs that represent some real states of affairs. In order to erect my account of artefact designs on firm conceptual ground, I hereby propose the following formal theory on states of affairs which encompasses some ideas taken from the philosophical legacy of Roman Ingarden.

First Ingarden's theory can be supplemented by the relation of parthood between states of affairs. As parts of ontic ranges of objects, states of affairs may be parts of one another. The state of affairs that John is handsome is also part of the state of affairs that he is handsome and shy. The state of affairs that John is an accountant is part of the state of affairs that he is an absent-minded accountant. To generalise, the state of affairs that p is part of the state of affairs that p and q . However, there are also non-conjunctive cases of parthood. For example, the state of affairs that John runs is part of the state of affairs that he is a fast runner.

The expression ' $\mathbf{Occ}(x, y)$ ' will mean that an object x occurs in a state of affairs y (or that y is a part of the ontic range of x). The expression ' $x \leq y$ ' will mean that a state of affairs x is part of a state of affairs y .

The formal theory of states of affairs is defined in axioms 2.2-2.6, and 2.8 below. Since these axioms constitute fairly weak characteristics of the respective notions, the resulting theory might be called a *minimal ontology of states of affairs*. All axioms are to hold both for real and intentional entities.

Given the relation of occurrence we may define the notions of object (2.1(i)) and of states of affairs (2.1(ii)):

- (2.1) (i) $\mathbf{Obj}(x) \equiv \exists y \mathbf{Occ}(x, y)$.
 (ii) $\mathbf{Soa}(x) \equiv \exists y \mathbf{Occ}(y, x)$.⁶

The following two axioms express the categorical constraints on objects and states of affairs.

$$(2.2) \quad \mathbf{Obj}(x) \equiv \neg \mathbf{Soa}(x).$$

$$(2.3) \quad x \leq y \rightarrow \mathbf{Soa}(x) \wedge \mathbf{Soa}(y).$$

Like Roberto Casati and Achille Varzi, I endorse the lexical principle to the effect that any relation of parthood is a partial order (Casati and Varzi 1999, p. 33). I express this principle in terms of axioms 2.4-2.6:

$$(2.4) \quad \mathbf{Soa}(x) \rightarrow x \leq x.$$

$$(2.5) \quad x \leq y \wedge y \leq x \rightarrow x = y.$$

$$(2.6) \quad x \leq y \wedge y \leq z \rightarrow x \leq z.$$

The expression ‘ $x < y$ ’ will mean that a state of affairs x is a proper part of a state of affairs y .

$$(2.7) \quad x < y \equiv x \leq y \wedge x \neq y.$$

It follows from the axioms and 2.7 that the proper parthood is irreflexive, asymmetric and transitive.

Given the metaphor of ontic range it would seem obvious that if a state of affairs x is part of a state of affairs y and an object z occurs in x (i.e., if x is part of the ontic range of z), then z will also occur in y :

$$(2.8) \quad \text{If } x \leq y, \text{ then } \forall z [\mathbf{Occ}(z, x) \rightarrow \mathbf{Occ}(z, y)].^7$$

3. Designs as states of affairs

I claim that designs are intentional states of affairs. What does this claim mean? Assume that a rational agent has designed an artefact. I argued above that:

- (i) the resulting design represents₂ part of the ontic range of the designed artefact,
- (ii) there is a mereological sum of sentences supporting that design.

Every sentence from the sum referred to in (ii) creates an intentional state of affairs. Every such state represents a part of the ontic range referred to in (i). Let X be the set of these intentional states of affairs. Because of the unity of the designer’s intentions, the design is also a unified entity. Given the above theory of states of affairs, the simplest way to represent the latter unity is to require that the design be the state of affairs that is the least upper bound of X with respect to \leq .

‘ A ’ will denote a set of artefacts. Let L_A be a language whose conceptual resources are adequate for speaking about the artefacts from A . Then S_A becomes the set of all intentional states of affairs created by the sentences from L_A .

- (3.1) For every artefact x , there is a non-empty set $X \subseteq S_A$ so that the least upper bound of X is a design of x .

How many designs does an artefact have? At first sight, the answer seems straightforward: every artefact has exactly one design. Observe, however, that an artefact may be part of another artefact and that a design of the latter may not specify all details of the former. A diode is a part of a power supply. The design of the diode that is part of the broader power supply design specifies only two parts of the diode: the anode and the cathode. Still, a more detailed design of the diode, for instance the design found in a handbook on general electronics, also mentions a semiconductor junction between the anode and cathode. Consequently, the diode has at least two designs. Consider also another case. Assume, for argument’s sake, that political organisations are

artefacts. The European Union ‘design’ lists the Republic of Poland as one of its parts and the provinces of Poland as parts of Poland. Nevertheless, the design does not mention that the many localities of Poland are parts of the EU. Still, the design of the Republic of Poland, as established in administrative Polish law, does mention these localities as parts of Poland. It therefore appears that ordinary parlance does allow us to admit that the same artefact has different designs.⁸ Strictly speaking, in their professional activities artefact designers seem to represent₂ one artefact by means of different representations₂. One, usually the most specific, representation₂ is created before the artefact is produced. Other, less specific, representations₂ are created when the artefact is used as a component of or as a tool for other artefacts. The fact that the most specific design is not deployed in the latter case is not a contingent matter but rather a consequence of the bounded rationality involved in engineering practice. In most cases, if a design x of a technical artefact were to contain the most specific designs of the components of the artefact, the design x would be an extremely complex representation. If, but only if, engineers were intentional agents with unlimited cognitive capacities, they might be able to ignore the complexity of their products. Later on I will also illustrate that the fact that one artefact has more than one design makes it possible to draw a useful distinction between artefact tokens and artefact types.

Given my identification of designs with intentional states of affairs, the most adequate approximation of the relation of specificity between designs is the relation of proper parthood between states of affairs.

(3.2) A design y_1 is *less specific than* a design y_2 iff $y_1 < y_2$.

I claim that the set of all designs associated with a given artefact has its greatest and least elements with respect to $<$. The existence of the former guarantees that every artefact is uniquely determined by its design, as far as its physical structure is concerned. The existence of the latter guarantees that there is an objective rationale for artefact token identification. Let me explain the latter claim in more detail.

Your new laptop has some dead pixels on the display. You want to reclaim it, so you visit the technical support unit. When you hear there that the laptop on the shelf is the same as your laptop, you do not start complaining that the technical support unit breaches the law of indiscernibility of identical artefacts. You presume that they mean that your laptop is a copy of the laptop on the shelf. We know that artefacts exist, so to speak, in copies. What we do not know is the conditions under which one artefact may be a copy of another.

It turns out that your second laptop is even worse than the first one. You are forced to replace one part after another. When does the original laptop on your desk cease to exist because of this replacement process? We know that in the course of time artefacts undergo various changes. What we do not know is the conditions under which artefacts preserve their identity over the course of time or cease to exist.

These and similar problems are more tractable if we reformulate them in terms of artefact tokens and artefact types. Then, instead of pursuing the question of when one artefact becomes a copy of another artefact, we pursue the question of when one artefact token is a token of the same artefact type as another artefact token. For example, the question ‘Is my laptop a copy of that laptop on the shelf?’ is superseded by the question ‘Is this token of a laptop a token of the same artefact type as that token of a laptop?’ Instead of pursuing the question of under what conditions artefacts endure in terms of time, we pursue the question of when an object is a token of a given artefact type. For example, the question ‘Does my laptop survive (i.e., preserve its identity

despite) the replacement of its battery B with a new battery B' is substituted by the question 'Will the replacement of B with B' in this artefact token lead to a change in its artefact type?' or by the question 'Is this artefact token with the old battery B of the same artefact type as the artefact token with the new battery B' '?

An artefact token is an artefact in the ordinary sense of the word. Artefact tokens are physical entities located in time and space. It is artefact tokens that we use as writing tools. They have certain properties, participate in processes and initiate events. One artefact token may be a copy of another. Within the ontological framework sketched above, one may say that artefact tokens are real objects and that their ontic ranges consist of real states of affairs. On the other hand, an artefact type is an abstract object like a universal. An artefact type gathers the features common to a group of artefact tokens. Artefact types are not located in time or space, therefore you cannot write philosophical papers with them; nonetheless, artefact types are useful conceptual devices. When I learn about the significant features of some artefact token, I may limit my investigation to the features specified by the artefact type of this token. When I notice that two artefact tokens x and y are similar in relevant respects, I may express my observation by stating that x is of the same artefact type as y . When I wish to make a general observation about a group of artefact tokens, I may express that by saying that the respective artefact type exhibits the observed regularity. In short, artefact types represent artefact tokens. Given the ontological framework sketched above, one may conclude that artefact types are intentional objects.

By now it should be obvious that the most specific design exhaustively determines the ontic range of the respective artefact type but the artefact tokens of such a type are not thereby completely determined. However, we must not compare the ontic ranges of artefact types with the most specific designs. Since we have to allow for the fact that some artefact tokens malfunction, we must not say that two artefact tokens are tokens of the same artefact type if their ontic ranges are only represented by the same most specific design. Even if an artefact token loses some of the features specified in its most specific design, the token may still "retain its artefact type". We ought to base the identity of artefact tokens on less specific designs and the most natural candidates are the respective least specific designs. If every artefact has its least specific design, then we may say that two artefact tokens are tokens of the same artefact type if their ontic ranges are represented by the same least specific design. I will thus compare the ontic ranges of artefact types with the least specific designs of artefacts.

This then places us in a position to solve our initial puzzles, the puzzle of copies and the puzzle of replacement. A physical object x is an *artefact token of an artefact type* y iff the least specific design of y represents₂ a part of the actual ontic range of x . When x is an artefact token of y , y will be termed an *artefact type* of x . A physical object x_1 is a *token of the same artefact type* as a physical object x_2 iff there is an artefact type in which both x_1 and x_2 are tokens. Observe that if the relation of being an artefact token is a function (i.e., if x is an artefact token both of artefact type y_1 and of artefact type y_2 , then $y_1=y_2$), then the relation denoted by the expression '... is of the of the same artefact type as ...' is transitive.⁹

Consequently, my laptop is a copy of that laptop on the shelf iff there is evidence of the least specific artefact design representing parts of the ontic ranges of both laptops. My laptop preserves its identity as long as its least specific design represents some part of its ontic range or if the least specific design represents some part of the ontic range of the laptop with B' rather than B , in which case the laptop survives the replacement of B with B' . On the other hand, if the design at stake represents no part of the ontic range of the laptop with B' , then the replacement process puts an end to the laptop.

Let the expression ‘**design**(y, x)’ abbreviate the expression ‘a state of affairs y is a design of an object x ’. Since designs are intentional states of affairs, the intended domain of the following formal theory of designs contains only intentional entities: intentional objects and intentional states of affairs.

I define artefacts types as “design-based intentional objects”:

$$(3.3) \quad \mathbf{Art}(x) \equiv \exists y \mathbf{design}(y, x).$$

I claim in this paper that designs are (intentional) states of affairs representing objects.

$$(3.4) \quad \mathbf{design}(y, x) \rightarrow \mathbf{Soa}(y) \wedge \mathbf{Obj}(x).$$

The most specific design of an artefact x will be called the *full design of x* and the least specific design will be called the *minimal design of x* . The fact that y is the full design of x will be denoted as **Design**(y, x). The fact that y is the minimal design of x will be denoted as **design**₀(y, x).

$$(3.5) \quad \begin{aligned} \text{(i)} \quad & \mathbf{Design}(y, x) \equiv \mathbf{design}(y, x) \wedge \forall z [\mathbf{design}(z, x) \rightarrow z \leq y], \\ \text{(ii)} \quad & \mathbf{design}_0(y, x) \equiv \mathbf{design}(y, x) \wedge \forall z [\mathbf{design}(z, x) \rightarrow y \leq z]. \end{aligned}$$

3.6 is the conclusion of the foregoing argument.

$$(3.6) \quad \mathbf{Art}(x) \rightarrow \exists y_1, y_2 [\mathbf{Design}(y_1, x) \wedge \mathbf{design}_0(y_2, x)].$$

It follows from 3.5 and 3.6 that the full design of x and the minimal design of x are unique, therefore I will denote them as **Design**(x) and **design**₀(x).

The notion of minimal design is the objective basis for artefact identification. This means that if x is an artefact type then the minimal design of x will be part of every state of affairs in which x occurs:

$$(3.7) \quad \mathbf{Occ}(x, y) \wedge \mathbf{design}_0(z, x) \rightarrow z \leq y.$$

The above axioms do not guarantee that artefact designs are not circular. There are at least two kinds of circularity at stake. The first type is more straightforward. Both artefacts and non-artefacts may occur in artefact designs but on pain of infinite regress I assume that no artefact occurs in its own design.

$$(3.8) \quad \mathbf{Design}(y, x) \rightarrow \neg \mathbf{Occ}(x, y).$$

3.8 *does not* proscribe the design supported by the sentence ‘the hammer x consists of the haft y and ...’, but it does proscribe the design supported by the sentence ‘the hammer x consists of the hammer x and ...’. Incidentally, 3.8 is related to the special sense of the expression ‘A design of an artefact determines the ontic range of the artefact’. Ultimately if y is a design of x , then y does not specify in what states of affairs x occurs but rather the states of affairs in which the objects which compose x occur.

The second type of circularity is more complex. Assume that a design of an artefact x_1 is less specific than a design of an artefact x_2 and that a design of x_2 is less specific than a design of x_1 . If we conceded that all artefacts have just one design, such a case would be excluded by the asymmetry of the relation of being less specific than (i.e., relation $<$). However, since we allow some artefacts to have more than one design, the situation depicted in Fig. 1 is possible.

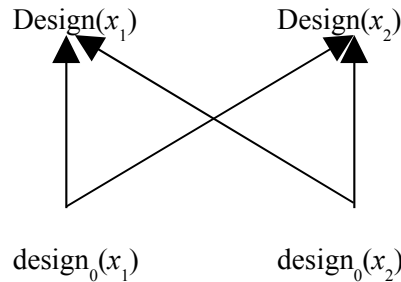


Figure 1

In this case it seems that x_1 is a proper part of x_2 (because a design of x_1 is a proper part of a design of x_2) and x_2 is a proper part of x_1 (because a design of x_2 is a proper part of a design of x_1), which is an absurd conclusion. In order to exclude such cases, I introduce axiom 3.9:

$$(3.9) \quad \mathbf{design}_0(x_1) < \mathbf{Design}(x_2) \wedge \mathbf{design}_0(x_2) < \mathbf{Design}(x_1) \rightarrow x_1 = x_2.$$

Of course 3.9 does not eliminate all cases of design circularity. For instance, 3.9 allows the following situation to occur:

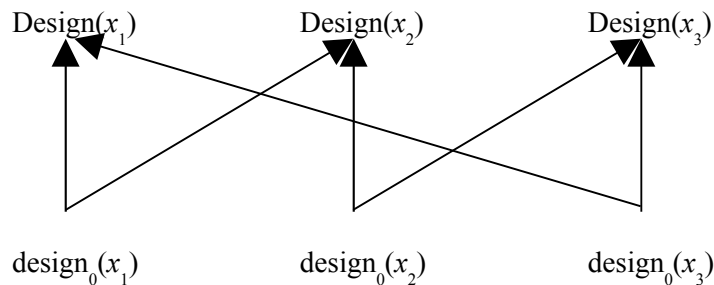


Figure 2

However, in order to exclude all such cases we either need to denumerably introduce many first-order axioms or to resort to a second-order theory.

The identifying of minimal designs with ontic ranges of artefact types means that artefacts types with the same minimal designs become identical. Needless to say, if two artefact types have the same non-minimal design, then they are identical as well. So if x_1 and x_2 are artefact types, then the following claim is valid:

$$(3.10) \quad \mathbf{design}(y, x_1) \wedge \mathbf{design}(y, x_2) \rightarrow x_1 = x_2.$$

Axioms 2.2-2.6, 2.8, 3.4, 3.6-3.10 and definitions 2.1, 2.7, 3.3, and 3.5 constitute my formal theory of artefact designs.

5. Conclusions

In this paper I have provided an ontological framework for the notion of artefact design, which is understood to be an outcome of designing activity. The minimal ontological commitments of the design discourse were revealed and formalised. I argued that given the above construal of this notion, we should account for the fact that some artefacts have more than one design. Finally, I introduced a series of constraints in order to exclude any “irrational” cases of design.

References

- Alberts, L. 1994. “YMIR: A Sharable Ontology for The Formal Representation of Engineering Design,” in: J. Gero and E. Tyugu, eds., *Formal Design Methods*, Amsterdam: IFIP Transactions, 3-32.
- Armstrong, D. 1997. *The World of States of Affairs*, Cambridge: Cambridge University Press.
- Barwise, J. and J. Perry. 1983. *Situations and Attitudes*, Cambridge, MA: MIT Press.
- Casati, R. and A. Varzi. 1999. *Parts and Places*, Cambridge, MA: MIT Press.
- Chandrasekaran, B., J. Josephson and V. Benjamins. 1999. “What are Ontologies, and Why Do We Need Them?,” *IEEE Intelligent Systems*, 14(1): 20-26.
- Garbacz, P. 2004. “The Four Dimensions of Artefacts,” in: D. Dubois et al., eds., *Principles of Knowledge Representation and Reasoning: Proceedings of Ninth International Conference (KR2004)*, Menlo Park: AAAI Press, 289-300.
- Gruber, T. and G. Olsen. 1994. “An Ontology for Engineering Mathematics,” in: J. Doyle, eds., *Fourth International Conference on Principles of Knowledge Representation and Reasoning*, Bonn: Morgan Kaufmann, 258-269.
- Hubka, V. and E. Eder. 1988. *Theory of Technical Systems: A Total Concept Theory for Engineering Design*, Berlin: Springer.
- Ingarden, R. 1964. *Der Streit um die Existenz der Welt, vol. I*, Tübingen: Max Niemeyer.
- Ingarden, R. 1965. *Der Streit um die Existenz der Welt, vol. II*, Tübingen: Max Niemeyer.
- Kitamura, Y., T. Sano, K. Namba and R. Mizoguchi. 2002. “A functional concept ontology and its applications to automatic identification of functional structures,” *Advances in Engineering Informatics*, 16: 145-163.
- Pahl, G. and W. Beitz. 1996. *Engineering Design. Systematic Approach*, Berlin: Springer.
- Simons, P. and C. Dement, 1996. “Aspects of the Mereology of Artefacts,” in: R. Poli and P. Simons, eds., *Formal Ontology*, Dordrecht: Kluwer Academic Publishers.
- Suh, N.P. 2001. *Axiomatic Design: Advances and Applications*, Oxford University Press, New York.

Endnotes

- 1 One of the numerous examples of the theory of axiomatic design, developed by N. P. Suh, which provides axioms that govern the design process, e.g., “Minimize the information content of the design”. See Suh (2001).
- 2 The precise significance of this claim is expressed by the following implication: if a design of an artefact determines that the artefact is such and such, then it is the case that the artefact is such and such. The claim *does not* entail the converse of that implication. At present I ignore the distinction between ‘is such and such’ and ‘is designed to be such and such’.
- 3 In what follows, the term ‘sentence’ will refer to sentence-tokens, i.e., inscriptions or series of phones.
- 4 Actually we should distinguish the inner ontic range of an object, i.e., its ontic structure, and the outer range of the object. The former corresponds to the notion of essence or nature; the latter to the notion of accident.
- 5 Intentional objects are in this sense different from intentional objects in the sense defined in Baker’s paper included in this issue.
- 6 Although the minimal ontology of states of affairs refers to two kinds of entities: objects and states of affairs, the ontology *is not* a many-sorted theory. Subsequently, all variables range between the same set of entities.

- 7 Since in this paper I only use my theory of states of affairs as a tool to grasp the ontological properties of artefact designs, I do not compare it with other formal accounts of states of affairs. I decided not to employ these accounts either because the complexity of their formalism is too great for the present purposes (e.g., Barwise and Perry (1983)) or because they involve more philosophical commitments than my minimal proposal permits (e.g., Armstrong (1997)).
- 8 In Hubka and Eder (1988) it is claimed that technical artefacts are represented by different engineering specifications.
- 9 Needless to say, the notion of least specific design does not solve all identity puzzles because we cannot distinguish different tokens of the same type on the basis of this notion.

Structure and Coherence of Two-Model-Descriptions of Technical Artefacts

Ulrich Krohs

Department of Philosophy, University of Hamburg,
and

Konrad Lorenz Institute for Evolution & Cognition Research, Altenberg

Abstract

A technical artefact is often described in two ways: by means of a physicalistic model of its structure and dynamics, and by a functional account of the contributions of the components of the artefact to its capacities. These models do not compete, as different models of the same phenomenon in physics usually do; they supplement each other and cohere. Coherence is shown to be the result of a mapping of role-contributions on physicalistic relations that is brought about by the concept of function. It results a sandwich-like structure of the two models, which can be reconstructed as a *two-sorted theory element*.

Keywords: Theory structure, Theory element, Function, Transistor

A technical artefact may be described in physicalistic and in functional terms. The physicalistic description accounts for structure and dynamics of the entity, while the functional description is based upon a design-and-use-centred view of the artefact. Both descriptions belong to different areas of knowledge and the question is how they cohere. Regarding the different descriptions as different models of the same entity, this question for coherence becomes a question of theory relation, i.e., of the relation between both models. To investigate this relation, I am splitting the problem of coherence into two questions. The first one is the question for consistency of the models. Though small inconsistencies would not completely corrupt coherence, they would diminish it and we should not presuppose that descriptions of technical artefacts are usually inconsistent. Consistency is reflected in theory structure, so the task is to reconstruct the structure of the “sandwich” of a physicalistic and a functional model and look whether it is consistent. Consistency alone, however, does not bring about coherence. Obviously, the junction of two isolated sets of propositions is consistent, but it would be strange to call this junction, which falls in two, coherent. Isolation has to be regarded as a feature that diminishes coherence (BonJour 1985, p. 95). Some relation must hold between the elements of physicalistic and functional descriptions that brings coherence about. Laurence BonJour regards mainly inferential relations between the elements of a class of propositions as increasing its coherence (op. cit.). According to Paul Thagard, among the candidates for such relations are explanatory and deductive relations as well as other mappings (e.g., Thagard et al. 2002). So the second task is to inquire into the kind of relation that holds between a physicalistic and a functional model and show how this relation accounts for the coherence of the description that is made up by a combination of both models.

The required concepts for the reconstruction are introduced in the first section of this paper. The second section develops the structure of consistent two-model-descriptions. The third section is dedicated to the question of coherence of physicalistic and functional models. In the fourth section, I will demonstrate the application of my reconstructive approach to the case of the

transistor. I will then discuss how this approach deals with the ambiguity of structure-function relations, and finally sum up my results.

1. Theory relations and sorts of models

The case of physics

Most investigations of theory relations have concentrated on physical theories. When two different models are given for the same process, in physics the question is usually: which is the better one? “Better”, of course, needed to be qualified and may be relative to some epistemic goal. Even if there is no ranking among the models, they are in most cases alternatives, not supplements. A description may use one model or the other, but one cannot combine both without running into inconsistency. For example, classical and relativistic mechanics cannot be mixed up. However, one of the models may be reduced to the other. This means that it is possible either to express all that can be said by the reduced model in terms of the reducing one, or at least to map all results of one model onto the results of the other. In terms of Scheibe’s synthetic concept of reduction (Scheibe 1993, p. 266), we have a *partial reduction* in the latter case.¹ Nevertheless, since such a mapping of results may hold even with inconsistent models, reducibility is not a guarantee for the consistency of two theories.

There is a case in physics in which the competition between two models is set aside and instead two seemingly incompatible models are used in parallel without a reduction relation holding between them: the case of the dualism of wave and particle descriptions of elementary particles. Neither model is sufficient for the description of all phenomena observed. Therefore both are used, as a workaround, in parallel. But in this case, both models have different ranges of application, which prevents inconsistency. So this is not a paradigm for cases like those that shall now be discussed.

The case of biology and of technology

If we look at biology, things turn out different. Here, we have two different sorts of models: physicalistic models that describe the physicochemical mechanisms of biological processes and functional models that describe the various mechanisms as realizing physiological functions. The functional account is normative since it allows for judgments about whether a component of a system behaves as it is supposed to do. So function modules are of another sort than physicalistic ones. While this characterization of the model sorts classifies models according to the involvement of functional terms, I have proposed an independent delineation criterion elsewhere (Krohs 2008).

The case of technology is similar to that of biology. Physicalistic models of technical artefacts, i.e., descriptions of the physical processes going on within an artefact, and functional models that describe these processes as being implemented to follow a certain goal, complement each other. Both descriptions are required and technical artefacts, if described only in terms of their physical properties not also from the functional perspective, would become mere physical objects and would lose their status as *technical* entities (Kroes 2002, p. 294).

Since the cases of modelling biological and technical systems both show the duality of physicalistic and functional views, many of the results that were obtained in philosophy of biology are also valid for the technological case. This holds in particular for the debate about reductionism. There was an extensive discussion of the question whether biology in general can be reduced to physics.² However, up to now all attempts failed and reduction cannot be regarded as a general means to eliminate functional talk, one of the main obstacles being the multiple

realisability of functions (Putnam 1967; Fodor 1974).³ One way to escape from unsuccessful reductionism is to regard functions as emergent properties of technical and biological systems and base the analysis of functional models on strong concepts of emergence (see, e.g., Beckermann et al. 1992). Reference to emergent properties, however, would be a way out only if functionalities were adequately conceived as properties. But they are categorically different from properties in being normative: a function may still be correctly ascribed even if the function bearer is dysfunctional and thus unable to perform the function.⁴ I therefore treat functionalities not as properties but as theoretical terms of the functional models that we use to describe biological organisms. Functional models have consequently to be conceived as being of a different sort than physicalistic models. Biological theories, then, are composites of models of both sorts. The strategy to investigate the interplay of models of different sort is to reconstruct the structure of such composite theories (Krohs 2004, chs. 7-8).

It is important to notice that a functional model of a technical artefact is not only an incomplete version of a physicalistic one. It makes use of classifications (like: being a signal or a force-generator) that cannot occur in a physicalistic model but are genuinely technical, i.e., related to design, use, or technical function of an artefact. Only the different carriers of a signal can occur in the physicalistic model. And only a functional model may be used to refer to malfunction, since this requires some kind of norm that discriminates function from malfunction – there is no such normativity in a physicalistic model. We have two non-exclusive possibilities by which to model these devices: a physicalistic model, and a functional model. Neither of them alone covers all that can be known about a technical artefact.

2. Theory structure: Consistency

To examine the coherence and consistency of mixed physicalistic/functional theories, the different models have to be described uniformly, i.e., they have to be reconstructed to some degree. For the present purpose, the structuralist version of model theory as developed by Sneed (1971) and refined by Stegmüller and co-workers (Balzer et al. 1987) seems to be the most adequate method of reconstruction. It starts from the assumption that the reconstructed models are not universally applicable but only locally, in an intended range of applications (which may be modified whenever new and adequate applications are found). This spares a lot of *caveats*, which were needed with a global approach like van Fraassen's constructive empiricism (van Fraassen 1972). So my decision to use structuralism is not based on any special "structuralist theory conception" that would claim that some theories *have* a structuralist structure (Stegmüller 1986, p. 3). I regard structuralism just as one of several tools to reconstruct theories and theory relations, and feel that it is for my purpose the most convenient of the presently available tools.

The approach

I will use a few concepts of structuralism only and will confine myself to introducing only these instead of the whole apparatus. Moreover, my version is a simplified version of structuralism.

Talking about models in the sense of structuralism, a second concept of a model has to be used that differs from the one used in science and in technology. A model in the structuralist sense is a mathematical structure, i.e., a set of objects with relations defined on it. (We can ignore for the present purpose that, strictly speaking, we would have to deal not with structures but with classes of isomorphic structures.) The structure itself, i.e., the mathematical model, is not yet the reconstruction of a technical description or of a scientific model, since it does not by itself apply to any phenomenon; it is not a model *of* anything. What is needed in addition is the specification of the intended phenomena the mathematical model has to be applied to. Therefore, the model (or

class of models) in the structuralist sense *plus* a class of intended applications makes up the smallest piece of theory or of description. This piece is called a *theory element*. A theory element, not a model in the mathematical sense, is the structuralist equivalent to a scientific model or to a technical description. For example: the mathematics of the Kepler laws defines a structure only. To turn this into a model of the dynamic of components of the solar system, we have to add that this structure is intended to describe this very dynamic, which is done in the definition of the variables. Moreover, we will not reject the Kepler laws just because they fail when applied to raindrops or to billiard balls. Instead, we simply leave these possible applications outside the class of intended applications of the laws. Only the model *plus* the intended applications allow for scientific explanation and make up a theory element.

I will refer in my argument to the following constituents of a theory element: a basis \mathbf{O} which is the set of objects the theory element deals with and a set \mathbf{R} of relations defined over \mathbf{O} . These two classes make up the core of a theory element and define the notions that are available within a particular element. \mathbf{O} specifies what can be called the ontology of the theory, and \mathbf{R} specifies the properties of the elements of \mathbf{O} . In addition, there is the set \mathbf{I} of intended applications of the theory core.⁵ So, a theory element can be characterised as (Balzer et al. 1987):

$$T = \langle \mathbf{O}, \mathbf{R}, \mathbf{I} \rangle.$$

Two-sorted theory elements

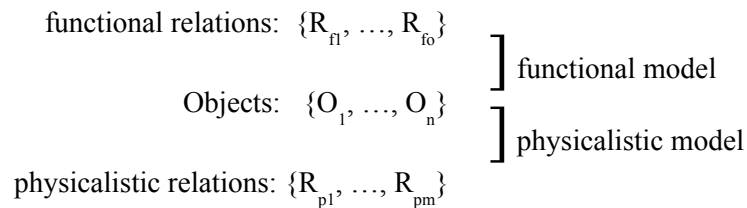
A theory element as characterised by \mathbf{I} , \mathbf{O} , and \mathbf{R} may be the reconstruction of any scientific model or technological description, be it of physicalistic or of functional sort. The sort depends on the elements of \mathbf{R} , which may be physical or functional relations. The crucial question is how to represent the interaction of a physicalistic and a functional model. I have introduced the notion of a *two-sorted theory element* to analyse this case (Krohs 2004).

A two-sorted theory element shall be conceived as a theory element that combines two “matching” models of the same phenomenon and has greater explanatory power than either of the two models standing alone. Not any arbitrary pair of models meets this requirement. The application of the concept has to be restricted to pairs of models that are (i) of different sort, that are (ii) models of the same entity or phenomenon and that (iii) describe the same aspect of this entity or phenomenon. I call two models that satisfy these requirements a pair of *corresponding* models. Satisfaction of condition (i) (sort-difference) is given in all cases where a physicalistic and a functional model are combined. Condition (ii) (identity of the described entity or phenomenon) is satisfied by any pair of sort-different models that share the same basis \mathbf{O} . However, there will usually be some differences in the bases, with a large intersection. The basis of the two-sorted theory element is the union of both bases. The overlap is significant in all cases in which it can be assumed that the modellers do regard a physicalistic and a functional model as referring to the same entity or phenomenon (e.g., the description of an integrated circuit as a set of logic gates and as a collection of semiconductors with certain electrical characteristics that bring the logical functions about; the biochemical and the information-transfer model of protein biosynthesis). Condition (iii), asking for identity of the aspect described, rules out cases in which one aspect of the entity or phenomenon is described in a physicalistic way, and another aspect in a functional way. For example, a physicalistic model of the heat production of a radio receiver and a functional description of its capacity to receive radio programs should not be regarded as being corresponding models because the described aspects are different. Corresponding models, in contrast, would be a physicalistic and a functional model of the capacity of the radio receiver to receive radio programs, or, if you like, a physicalistic and a functional model of the heat production of a receiver. In terms of structuralism: condition (iii) is met only in such cases in

which both corresponding models share the application. We can now give the definition of the correspondence relation (corr):

- (corr) Two models of different sort correspond *iff* the intersection of the bases of both models is not empty and both are applied to the same shared element of the respective classes of intended applications.

A two-sorted theory element then is a pair of corresponding models. In the simplest case, it is made up of the shared basis O_{T_2} , plus the classes of physicalistic and of functional relations, R_p and R_f . This results in a “sandwich-structure” of a two-sorted theory element: Two different slices of bread (the two sets of relations) share the butter (the set of objects):



In the general case, as already mentioned, the classes of objects of both models differ but intersect. In addition, the class I_{T_2} of intended applications has to be specified. I_p and I_f may be identical. In this case, the class I_{T_2} of intended applications of the two-sorted theory element will be the same as for each of the separate models. The general case is that the two models have slightly different but overlapping intended applications. So, in general, I_{T_2} is the intersection of the classes I_p and I_f of the classes of intended applications of the two separate models. According to the requirement of (corr), this intersection is never the empty set.

So, the general structure of a two-sorted theory element is:

$$T_2 = \langle O_p \cup O_f, R_p, R_f, I_p \cap I_f \rangle,$$

with indices “p” and “f” for the physicalistic and functional model, respectively.

Let me point out again that a two-sorted theory element requires that its models differ with respect to their sort. Otherwise they would represent either incompatible or redundant conceptions of the entity in question. In particular, the relations that belong to the different models are of different sort and are not shared among the models. This accounts for the consistency of the theory element, given that the models are themselves consistent: as long as two models share only objects and intended applications, but not relations, no contradictions can occur between them; their combination is consistent. The set of both models will of course contain any inconsistency that might occur within each of the isolated models, but the combination will not give rise to any further inconsistencies.

3. Coherence

My account of correspondence reconstructs how a physicalistic and a functional description of a technical artefact may be used in parallel without competition. In a two-sorted theory element there is, so to say, room for sets of physicalistic relations as well as of functional relations. This shows the consistency of two-model descriptions of technical artefacts. However, this is not yet a satisfying answer to the question of coherence of physicalistic and functional descriptions. It does not spell out the way in which the two models interact.

In the introduction, I have pointed out that this interaction has to be reconstructed in terms of formal relations that hold between the elements of a body of knowledge. The body of knowledge that is in question here is the technical description as reconstructed in terms of a two-sorted theory element. What we are interested in now are the *relations* that hold between corresponding functional and physicalistic models. These may be of different kind.

First, there might be classical inferential relations, as favoured by Bonjour as mediators of coherence (1985). Due to the multiple realisability of functions and the heterogeneity of the functions that a function bearer may perform (see below, section 5), standard accounts of inference do not seem to be able to explain the coherence of two-sorted theory elements. The same holds true for the most important aspect of the approach of Thagard and co-workers, namely of explanatory and deductive relations (Thagard et al. 2002). Establishing relations of such kind would result in establishing a reduction relation between both sorts of models, which was shown not to hold in the cases considered here (cf. section I). But Thagard exposes another kind of relation that accounts for what he calls coherence by analogy, namely a mapping relation that holds between subsets of a body of knowledge (ibid.). I will concentrate on this mapping relation, which turns out to be a central mediator of coherence in the present case, without ruling out that other relations may hold that further increase coherence of the theory element.

As I have mentioned in my exposition of the notions of structuralism, the core structure of a theory element defines the notions and relations that are available. But not every (if any) scientific theory or technical description is independent of any other theory. Concepts that are defined in other theories can be incorporated by so called *intertheoretical links* that hold between the two theories (Balzer et al. 1987). Being theoretical terms of other theories, such concepts can be applied in a theory element in question according to the rules that are specified within the theory they originate in. The links are additional constituents of the core structure of a theory element. My claim is that the coherence between a physicalistic and a functional description of a technical artefact is brought about by a concept that is defined in a more general theory and provided by a link: the concept of function.

The link makes the notion of a function available as provided by a theory of function. Different such theories are discussed in the literature, resulting in different accounts of the concept of function. It does not matter for my present purpose which one is preferable, as long as it provides criteria for the functionality of processes or structures as described in a physicalistic way. In principle, this might be an etiological account like Millikan's (1984), based on the causal history of an entity, or one based on use plans like that of Houkes and Vermaas (2009). I have proposed an alternative explication that avoids shortcomings of other theories. It is based on the role that reference to design plays in accounts of the ontogeny or construction of an entity (Krohs 2004, chs. 4-5; Krohs 2009). "Design" includes intentional and natural design, the latter being the outcome of evolutionary processes. My approach therefore integrates biological and technical functions. However, I will not enter the discussion about the adequacy of different theories of

function here. For the reconstruction of two-sorted theory elements we merely need a workable concept of function that ascribes functions to components of technical (and biological) systems. Any such concept secures the interaction of physicalistic and functional descriptions.

The role of the theory of function is the following: It provides the notion of function that can be applied to the objects O_i and the relations R_j of the physicalistic model. Only if elements of R_p can be shown to be functional according to the theory of function used, can it be regarded as justified to supplement the physicalistic model with a functional one: the concept of function becomes part of the physicalistic model, but it allows for the addition of a functional one. The theory of function that is applied maps the functional relations R_f onto the physicalistic relations R_p . This, obviously, need not be a one-to-one mapping. The structures of both models may differ greatly, as long as any relation out of R_f can be mapped onto R_p as a whole set (Fig. 1).

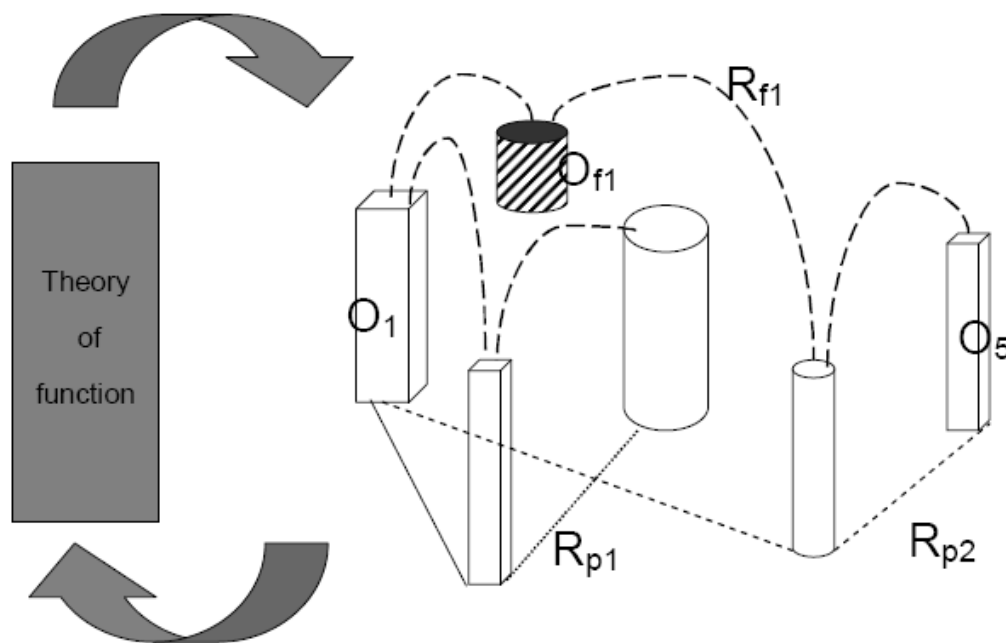


Fig. 1: Visualisation of a two-sorted theory element. Right side of the figure: The physicalistic model consists of a set of objects, O_1 - O_5 (only two of the five objects are labelled in the figure), and a set of two relations, R_{p1} (solid lines) and R_{p2} (dotted line). The functional model of the same aspect of the same entity consists, in this example, of the same objects plus an object O_{f1} that does not occur in the physicalistic model, and of a single relation, R_f (dashed line), that requires the additional object. The theory of function that is applied maps this relation onto the relations of the physicalistic model, as indicated in the left part of the picture. As is the case in the example, this need not be a 1:1 mapping.

To sum up, coherence of models in a two-sorted theory element (given the identity of the intended application) is brought about by (i) an – at least partly – shared class of objects and (ii) by the status of some of the physicalistic relations that hold between the objects as being functional according to a link to the theory of function, therefore by a mapping of the functions to physicalistic relations. Such a mapping increases the coherence of a body of knowledge by elucidating, in Thagard's terminology, an analogy between its different subsets.⁶ However, since the mapping imposes relations on the components of the two-sorted theory element, it could in principle introduce not only an account of analogy, but also inconsistencies. This possible source

of inconsistencies is not covered by the argument about the structural independence of the pair of corresponding models (section 2), so it is crucial that it is observed by the theory of function. Not giving rise to inconsistencies in the mapping of functions to physicalistic relations might be regarded as a criterion that has to be met by any sound theory of function.

4. Application

I now want to demonstrate the adequacy of my reconstruction using descriptions of the transistor as an example. A transistor is a semiconductor element that was developed as a device for amplification of electric current. From the literature, we can reconstruct a physicalistic model that describes the structure and electric characteristics of the transistor, and a functional model that describes how it works as an amplifier. In the descriptions given in the original U.S. Patents, both models are more or less blurred (Shockley 1950; Bardeen and Brattain 1950). For reconstruction, I therefore refer to the account of the transistor that Darlington gives on the basis of these first patents. In his patent of a transistor pair (Darlington 1953), he gives the following description of the structure of the transistor:

Transistors comprise, in general, a body of semiconductive material and three connections, termed emitter, base and collector, to the body. ... Transistors ... may be classified further as to conductivity type; in an N-type junction transistor, the intermediate zone, i.e., the one with which the emitter and collector zones form junctions, is of N-conductivity type (Darlington 1953, p. 1).

His description characterises the class of objects the model deals with: Elements of \mathbf{O} are a piece of semiconductive material (usually silicon or germanium), atomic impurities, and three connections (made of metal). Free electrons may be added that did not need to be mentioned in the description because they are present in the used material anyway. Next, Darlington describes the physical relations that hold with respect to these objects:

In the utilization of transistors, a number of what may be considered as circuit parameters or aspects are of prime moment. Among and illustrative of such parameters are the current multiplication factor, commonly designated alpha, and the emitter and collector resistances (ibid.).

According to this description, elements of \mathbf{R} are the mathematical functions stating the current multiplication factor, and the emitter and collector resistances. In the structuralist reconstruction, the mentioned parameters belong to the class \mathbf{R}_p of relations of the physicalistic model. In addition, so-called “holes” or defect electrons as defined by the model of semiconductivity (Wilson 1931) will be elements of this class.⁷ Darlington further characterises the physicalistic model by describing how the parameters just mentioned are influenced by structural variations:

[The current multiplication factor] is dependent upon a number of controllable factors, such as the width of the intermediate zone Similarly, the emitter and collector resistances are amenable to control by design (ibid.).

So, if the required data were provided, the physicalistic model could be stated in a completely quantified way.

What about the functional model? Darlington, like Shockley and Bardeen & Brattain, speaks about amplification of a signal. “Amplification” and “being a signal”, in contrast to the already

mentioned physicalistic concepts of a current multiplication factor and of a current, are functional classifications of the transistor's electric characteristics and of the change or time course of an electric current, respectively. This view is supported by the fact that there is not one single correlate to the signal. It occurs in various realisations, as it may in general be the case with functions. A voltage change of different currents may be the very same signal. Since signal processes are functional, we are dealing with an element of the class \mathbf{R}_f of functional relations that are part of the functional model.

Since the structure of the transistor as referred to in the functional model does not differ from the description given in the physicalistic model, domain \mathbf{O} is identical for both models. (However, only two connections instead of three may be required in some other functional descriptions.) The "sandwich-structure" of the two-sorted theory element is therefore made up of two classes of relations, \mathbf{R}_p and \mathbf{R}_f , and a common domain \mathbf{O} . Both models, the physicalistic and the functional one, share the intended application I_i , which is the behaviour of the transistor in an electric circuit of a certain type. Therefore, all three requirements for correspondence as posed in section 2 are met: sort difference, identity of the described entity, and identity of the modelled aspect. With respect to the two models, (corr) holds, and the technical description given in the patents has the structure of a two-sorted theory element.

To account for the coherence of a two-sorted theory element, we need to have a look at the mapping of the relations of both models. Above, I have introduced this mapping as a matter of the application of the concept of function. We have seen that the functionalities of amplification and of being a signal are mapped on the physicalistic roles of a current multiplication factor and of fractions of different currents. We do not rely on any particular explication of the concept of function here, so no restriction applies in this regard. It is clear from Darlington's description that such a mapping holds, mediated by a concept of function that is not further specified. So, coherence by analogy can in fact be found in the transistor example.

5. Accounting for functional ambiguity

We have not yet addressed the fact that the twinning of functional with physicalistic models may be ambiguous. Most technical artefacts can serve different functions, and in general, functions can be realised differently. This is known as heterogeneity of functionality and as multiplicity of realisation, respectively (Carrier 2000). It has to be asked on the one hand, whether the reconstruction presented can account for both, and on the other hand, whether an account of ambiguity brings in any unwanted arbitrariness. Sticking to the example of the transistor, I will first have a look at the heterogeneity of its possible functions. Above, I was concerned with the functional model of the transistor as used as an amplifier. However, it may be used in different ways. Some of these are mentioned in the quoted patents, others not.⁸ Most prominent might be the function of the transistor as a switch. The switch model refers to off- and on-states of the transistor. They are characterised by saturation of collector current in the first state, and by zero collector-emitter-voltage in the second state. The characteristic of the transistor in the range in between, which is most important for the amplifier-model of the transistor, almost does not matter for the switch model. Nevertheless, the physicalistic model that corresponds to the functional switch-model is the same one that corresponds to the functional amplification model. Each pair of corresponding models has a common intended application (transistor in electric circuit of a certain type). So we have two different two-sorted theory elements describing the transistor. Functional heterogeneity of technical artefacts (see the two left columns of table 1) can easily be accounted for.

Table 1: Multiplicity and heterogeneity.

Function	Artefact		
Amplification	Transistor	Tube	<i>Multiple realisability of amplification</i>
Switching	Transistor	Relay	<i>Multiple realisability of switching</i>

Functional heterogeneity of transistor use

Multiple realisability of a function can be accounted for in a similar way. Let us regard the function of signal amplification. This function can be realised by a transistor and described by the two-sorted theory element already discussed. Alternatively, it may be realised by a vacuum tube, which was functionally the predecessor of the transistor. Similarly, there are alternative realisations of electrical switches. The switching function may be realised, e.g., by a relay instead of a transistor (see right column of table 1). Let us regard the two realisations of an amplifying element only, a transistor and a tube. Physically, they are completely different. So the physicalistic models are different as well. However, the functional model will differ only with respect to the elements of the domain that refer to the physical entity: to the structure of the vacuum tube and of the transistor, respectively. But the two functional models will define almost the same functional relations over these domains. So again, we have two different two-sorted theory elements, in this case describing two different technical artefacts in similar functional contexts. This accounts for multiple realisability of functions.

At first sight, my approach seems to be fairly liberal with respect to the functions that might be assigned to an artefact. This might be judged as being problematic. Therefore, I want to add some considerations about this issue. I was dealing with *descriptions* of technical artefacts. Doesn't this mean that I have to accept any function as relevant for a reconstruction in a two-sorted theory element that anybody likes to describe, no matter whether it has any relation to the actual or intended use of the artefact, the goals that anybody might pursue by using it, or the norms that are set on its construction? First, it seems quite clear that functions of technical artefacts are in fact context dependent in some way. Herbert Simon conceives an artefact as the interface of an internal structure and a (physical) environment. In his view, exactly this interface position seems to bring about the association of a goal with the artefact (Simon 1996, pp. 6-7). Peter Kroes has adjusted this conception and made clear that the relevant environment is not the physical environment but human action (Kroes 2002). This results in the view of the dual nature of a technical artefact *as* embedded in a context of human action. My account seems to go another step further and change the environment of human action into an environment of human description. This, I agree, would be a step too far. However, the impression of complete dependency on the description is false. There are two reasons for my claim. (i) I admit that my conception of correspondence allows for a wide range of combinations of physicalistic and functional models. Nevertheless, there is an additional demand for coherence: the theory of function itself has to be applicable in the framework of the physicalistic model. This guarantees that not any effect or relation that can be ascribed to a component of a complex entity may be considered in the functional model, but only functional relations that can be isolated from physicalistic models. So the theory of function has to be specific enough to prevent arbitrary function ascriptions. Reconstruction of theory structure only has to show *that* the theory of function is crucial in this respect; it does not have to develop this theory. (ii) The validity of the functional part of a two sorted theory element has to be proven by standards as strong as those that are used for the physicalistic part. A functional model that is not in accordance with observations of the structure, dynamics, and possible use of an artefact is not a good model and

must not be regarded as a valid functional description. (Such an evaluation would mostly fall within the scope of the special sciences.) But discrimination of good and bad models cannot be a matter of reconstruction. Reconstruction deals with theories only and has no grip on the modelled phenomena. Moreover, it has to be possible to reconstruct even implausible combinations of physicalistic and functional descriptions for the sake of analysis. My claim is only that the reconstruction method has a grip on “mixed” models, not that it, purely as a method, sorts out the best model.

6. Conclusion

Descriptions of technical artefacts consist of physicalistic and functional elements, which can be ascribed to different models. The description as a whole was reconstructed as a two-sorted theory element, which, in the simplest case, combines a pair of corresponding physicalistic and functional models in a sandwich-like structure. This accounts for the consistency of a mixed physicalistic-functional description of a technical artefact. The same theory structure can be found in mixed physicalistic-functional descriptions in biology. It was shown how the concept of function, as provided by an adequate theory of function, maps the relations of the functional model onto relations of the physicalistic one. The mapping relation brings the coherence of a two-sorted theory element about, in terms of what Thagard has dubbed coherence by analogy.

References

- Agazzi, E., ed. 1991. *The Problem of Reductionism in Science*. Dordrecht: Kluwer.
- Balzer, W., C.U. Moulines and J.D. Sneed. 1987. *An Architectonic for Science: The Structuralist Program*. Dordrecht: Reidel.
- Bardeen, J. and W.H. Brattain. 1950. *Three-electrode Circuit Element Utilizing Semiconductive Materials*. U.S. Patent No. 2,524,035, Oct. 3, 1950.
- Beckermann, A., H. Flohr and J. Kim, eds. 1992. *Emergence or Reduction? Essays on the Prospects of Nonreductive Physicalism*. Berlin: De Gruyter.
- BonJour, L. 1985. *The Structure of Empirical Knowledge*. Cambridge, MA: Harvard University Press.
- Carrier, M. 2000. “Multiplicity and Heterogeneity: On the Relations between Functions and their Realizations,” *Studies in the History and Philosophy of Biology and Biomedical Science*, 31: 179-191.
- Darlington, S. 1953. *Semiconductor Signal Translating Device*. U.S. Patent No. 2,663,806, Dec. 22, 1953.
- Fodor, J. 1974. “Special Sciences (Or: The Disunity of Science as a Working Hypothesis),” *Synthese*, 28: 97-115.
- Houkes, W. and P.E. Vermaas. 2009. “Produced to Use: Combining Two Key Intuitions on the Nature of Artefacts,” *Techné*, 13 (2): 123-136.
- Kroes, P. 2002. “Design Methodology and the Nature of Technical Artefacts,” *Design Studies*, 32: 287-302.
- Kroes, P. 2006. “Coherence of Structural and Functional Descriptions of Technical Artefacts”. *Studies in History and Philosophy of Science*, 37: 137-151.
- Krohs, U. 2004. *Eine Theorie biologischer Theorien: Status und Gehalt von Funktionsaussagen und informationstheoretischen Modellen*. Berlin: Springer.
- Krohs, U. 2008. “Conservative and Non-Conservative Models in Molecular Biology and the Concept of Genetic Information,” presented at the 21st Biennial Meeting of the Philosophy of Science Association, Pittsburgh.
- Krohs, U. 2009. “Functions As Based on a Concept of General Design,” *Synthese*, 166: 69-89.

- Krohs, U. (forthcoming). "Dys- mal- et non: l'autre côté de la fonctionnalité," in: J. Gayon & M. Mossio, eds., *Épistémologie de la catégorie de fonction: des sciences de la vie à la technologie*. Paris: Presses Universitaires de France.
- Millikan, R.G. 1984. *Language, Thought and Other Biological Categories: New Foundations for Realism*. Cambridge, MA.: MIT Press.
- Putnam, H. 1964. "Psychological predicates," in: W. Capitan and D. Merrill, eds., *Art, Mind, and Religion*. Pittsburgh: University of Pittsburgh Press: 37-48.
- Scheibe, E. 1993. "A new theory of reduction in physics," in: J. Earman, A.I. Janis, G.J. Massey and N. Rescher, eds., *Philosophical Problems of the Internal and External Worlds: Essays on the Philosophy of Adolf Grünbaum*. Konstanz: Universitätsverlag Konstanz / Pittsburgh: University of Pittsburgh Press: 249-271.
- Shockley, W. 1950. *Semiconductor amplifier*. U.S. Patent No. 2,502,488, Apr. 4, 1950.
- Simon, H.A. 1996. *The Sciences of the Artificial*. 3rd edition, Cambridge, MA.: MIT Press (1st ed. 1969).
- Sneed, J.D. 1971. *The Logical Structure of Mathematical Physics*. Dordrecht: Reidel. 2nd ed. 1979.
- Sober, E. 1999. "The Multiple Realizability Argument against Reductionism," *Philosophy of Science*, 66: 542-564.
- Stegmüller, W. 1986. *Theorie und Erfahrung, dritter Teilband: Die Entwicklung des neuen Strukturalismus seit 1973*. Berlin: Springer.
- Thagard, P., C. Eliasmith, P. Rusnock and C.P. Shelley. 2002. "Knowledge and Coherence," in: R. Elio, ed., *Common Sense, Reasoning, and Rationality*, Vol. 11, Oxford: Oxford University Press: 104-131.
- van Fraassen, B. 1972. "A Formal Approach to the Philosophy of Science," in: R.G. Colodny, ed., *Paradigms and Paradoxes. The Philosophical Challenge of the Quantum Domain*, Pittsburgh: University of Pittsburgh Press: 303-366.
- Van Regenmortel, M.H. and D.L. Hull, eds. 2002. *Promises and Limits of Reductionism in the Biomedical Sciences*. Chichester: Wiley.
- Wilson, A.H. 1931. "Theory of Electronic Semi-Conductors," *Proceedings of the Royal Society London*, A 133: 458-491.

Endnotes

- 1 The case of partial reduction is all but an exception. Even for the mentioned standard example for reduction we cannot obtain more than this (Scheibe, op. cit.).
- 2 For a discussion of reductionism, see, e.g., the contributions in Agazzi (1991), Beckermann et al. (1992), and Van Regenmortel and Hull (2002).
- 3 For a dissenting view, see Sober (1999).
- 4 The concept of function has even more normative aspects than the generally acknowledged case of dysfunctionality (Krohs, forthcoming).
- 5 Specification of some further classes would be required to fully characterise the theory core; only one of them needs to be introduced below.
- 6 In addition, practical inferences may play an important role, as Peter Kroes has recently pointed out (Kroes 2006).
- 7 The concepts of Wilson's model are again provided *via* a link between his theory and the physicalistic model of the transistor that makes use of it.
- 8 Bardeen and Brattain (1950, p. 1) generously write about "a novel method of and means for translating electrical variations for such purposes as amplification, wave generation, and the like."

Book Reviews

***Participatory Democracy, Science and Technology* by Karl Rogers (Palgrave MacMillan, 2008). 256 pp. ISBN: 978-0230522060.**

This book makes good on many of the promises made by Rogers' previous work, *Modern Science and the Capriciousness of Nature*. How exactly ought people in a well functioning democracy interact with the strongest forces that shape their lives (science and technology)? Do we threaten the objectivity of science when we democratize it? What role do technical experts have in a democratic society? What does "democracy" mean? Rogers ambitiously attempts to answer all of these questions while simultaneously building a convincing case that the democratization of science and technology isn't simply a good thing for democratic societies vis-à-vis the realization of democratic ideals, but is in fact a necessary component of "good" science and technology.

After a careful opening critique of technological determinism found in the substantivist theories of technology, (Heidegger, Marx, Marcuse, Ellul, Heilbron's soft determinism, etc.) the third chapter pulls heavily from Feenberg to suggest a "dialectical" theory of technology. This, at its core, is an attempt to make sense of the dialectical nature of the relationship between technology and society (i.e., how technologies are shaped by human choices and how human choices are shaped by technology). This departs from (or perhaps supplements) Feenberg's account by emphasizing an irresolvable ambiguity between what Feenberg calls the primary and secondary instrumentation of technology. This ambiguity arises out of the dialectical nature of technology, allowing Rogers to sweep away the last vestiges of determinism from the substantive theories of technology on which he is building.

Chapter four, on participatory democracy, is largely meant to explain how Feenberg's call for "deep democratization" is supposed to play out. The author worries that without a full account of how "deep democratization" is understood, there is no clear path to move from a technocratic authoritarianism to a democratic technological society without remaining open to traditional technocratic arguments or claims of impracticality or perhaps even the undesirability of overthrowing the current technocratic regime. Participatory Democracy is meant to be that path. Here, the author relies heavily on Benjamin Barber's notions of thick and thin democracy. Through thickening (increasing participation in) traditional (thin) liberal constitutionalist structures, a democratic citizenry may emerge out of technocratic authoritarianism without the need for violent revolution or over-reliance on an impractically motivated citizenry. This is because through thickening already existent thin democratic political structures it is possible to rely on liberal constitutionalist ideals during the transition, eventually replacing them with a fully functioning participatory democratic society. In this way, liberal constitutionalist values serve as a propaedeutic, as a ladder that, once used to climb up, may be thrown away. It is through participating in participatory democratic institutions that a citizenry capable of the kind of self reliance necessary for a well functioning participatory democratic society emerges.

Next, the author turns toward the question of scientific and technical expertise in a participatory democracy. It is here that the author makes good on the promissory note in the first chapter to defend the idea of democratic participation as a practically valuable thing in a technological society (rather than democracy being a moral good that trumps technical considerations). Via an appeal to Polanyi's understanding of the nature of science, the author argues for a relation

between the form of technical work and the form of the wider society within which this work emerges; since scientists' intuitions at having correctly established a connection with an independent reality through theoretically understood material are themselves a manifestation of a more comprehensive societal gambit involved in constructing a technological society, the democratization of scientific research and technological development requires the democratization of the society at large. This means that science and technology will, in so far as they are a part of society, emerge as more democratic institutions that communicate bi-laterally with non-scientist members of the public. This, in conjunction with a brief rehearsal of Wynne's classic piece on sheep farmers in Cumbria amounts to the argument for the practicality of democratizing science and technology. This is an important way of making the technological infrastructure of regions and nations more sustainable, diverse, flexible and robust because they are better integrated into local social circumstances in which they are located. Excluding the public from technical or scientific decisions "simply leads to badly implemented and developed science." Thus, democratic structuration has practical value for the rational development and implementation of scientific research and technological innovation because it "increases the social capacity to intelligently and creatively adapt and respond to events in our messy, complicated, and capricious world."

The final chapter fleshes out this relationship between democracy and the "rational society." Drawing conspicuously on Habermas' notion of the ideal speech situation, the author notes that, because of the dialectical relationship between science and technology and society, at any given moment in history exactly what counts as "rational," what epistemological or moral standards, and what constitutes free and open deliberations can be in dispute. This creates a situation in which rationality itself is something that needs to be contested in a democratic forum. In the absence of a universal agreement on what constitutes good reasons for action the most "rational" thing to do is to place technical decisions into the democratic sphere so as to scrutinize them from as many perspectives as possible. A rational society will be one that uses science and technology in a way that conforms as much as possible to the desired structure of the lifeworld (i.e., a society in which people are empowered to shape how science and technology shape their material conditions). In this way, opening decisions, typically decided via instrumental rationality, to participatory democratic fora will ensure decisions adhere more closely to something that approximates a rational decision. In this way participatory democracy becomes "an ontological condition for the maximization of the societal capacity for survival, creativity, experimentation, and freedom."

The book is a most welcomed addition to the growing number of works in STS devoted to the intersections of democracy theory and science and technology. While this reviewer would have liked to see a more selective focus on particular forms of participatory democracy (i.e., deliberative democracy [the word deliberation is used no less than 39 times in the final chapter]), the book successfully does a lot of the heavy lifting of demonstrating the fertile areas in which science and technology may constructively interface with democratic theory while making both the science and the society better for it.

Brandiff R. Caron

Virginia Polytechnic Institute and State University

***New Waves In Philosophy Of Technology*, edited by Jan Kyrre Berg Olsen, Evan Selinger, and Soren Riis (Palgrave Macmillan, 2009). 384 pp. ISBN: 978-0230220003.**

Don Ihde, in his foreword to this volume, classifies the authors collected here as representing a fourth wave – the new wave – among philosophers focusing on technology, technologies, and technological culture. Ihde includes himself and myself among the third-wave philosophers, and it may not have been wise on the part of the editors of *Techne* to have invited me to comment on our next-wave successors. Presumably each wave has added, or claims to have added, something both new and different – and, one would hope, better – to the contributions of its predecessors. In doing such comparative measurements, I once, here in the pages of the predecessor version of SPT's *Techne* (whole volume 4), summarized the types of measures typically used. Some contributions to the literature are said to be “quantitatively” better: like advances in scientific fields, what is said to be better is supposed to build explicitly on prior knowledge. Other alleged improvements are merely “qualitative”: they offer allegedly better *value* judgments, better *syntheses* of prior work, or, finally, only more *originality*. What, in these terms, can be said about the work of the “new wave” philosophers collected in this volume? I run through them one by one.

Keekok Lee, “*Homo faber*: The Unity of the History and Philosophy of Technology.” Lee has actually been around long enough to have joined those of us in Ihde's third wave, though it is true that she was not prominent among philosophers of technology in our era. In this essay, she argues that – although so many changes have taken place in the history of western philosophy since the Greeks that it seems unlikely that technology, in all its forms from primitive to contemporary, could be understood within a single philosophical framework – there is a common thread in the notion of *Homo faber*. Unfortunately, to my eyes, her survey is so sweeping that it might well have been written during Ihde's first wave.

Jan Kyrre Berg Olsen, “Becoming through Technology.” This is actually an essay on science, not technology, though it does pay some attention to the technologies of time measurement. To me, it reads like a reworking of a running conflict between Milic Capek and Adolf Grunbaum as far back as the 1960s. Berg Olsen puts a novel twist on the argument. But I can't help remembering how, when Capek retired to our philosophy department at the University of Delaware and, kind person that he was as a colleague, he wondered why I would have turned from philosophy of science (good) to philosophy of technology (at best questionable). I'm sure he would have the same doubts about Berg Olsen, at least in this essay.

Robert Rosenberger, “Quick-Freezing Philosophy: An Analysis of Imaging Technologies in Neuroscience.” This is an interesting – while difficult for anyone not familiar with the neuroscience literature discussed – application of Ihde's “postphenomenology” type of analysis to a case study in a specialized subfield of neuroscience, the nature of synaptic vesicles in neurotransmission. A good technoscience case study, building on a third-wave predecessor in philosophy of technology.

David M. Kaplan, “How to Read Technology Critically.” Kaplan has also been around for awhile, but this is a genuinely novel approach, though it relies principally on the thought of Paul Ricoeur, and Kaplan admits that Ricoeur has actually contributed little to the philosophy of technology. Ricoeur as here interpreted by Kaplan should contribute a great deal to the “fourth wave.”

Graham Harman, “The McLuhans and Metaphysics.” This is an original replay of Marshall and Eric McLuhan's use of the tetrad (defined as a “fourfold”) as an analytical structure in all fields, with special reference to the elder McLuhan's “understanding media.” It is largely based on *Laws of Media* (1988), in which the younger McLuhan tried to breathe new life into his father's ideas,

then in something of a decline not only among the third-wave philosophers of technology but generally. The essay, in my opinion, is decidedly original, as well as refreshingly comprehensive.

Soren Riis, “The Question Concerning Thinking”; and Iain Thomson, “Understanding Technology Ontotheologically, or: The Danger and the Promise of Heidegger, an American Perspective.” Ihde refers to these essays as dealing with the specter or ghost of Heidegger that is still found wandering through the fourth wave. I would leave it to Robert Scharff, the leading Heideggerian of the third wave (leaving aside Ihde’s post-heideggerianism), to say whether or not there is even anything really original in these two essays.

Nick Bostrom, “The Future of Humanity,” and Philip Brey, “Human Enhancement and Personal Identity.” This paired set of essays, oddly inverted in order, reflect Bostrom’s posthumanism and Brey’s critical assessment of it. Brey actually goes out of his way to be fair to Bostrom (and his fellow posthumanists) in a long essay, saving his devastating “ethical considerations” for just the last couple of pages. There Brey argues that, “Even if new inequalities could somehow be prevented, which seems unlikely, the question would remain whether human enhancement would really improve human lives”(p. 182). Incidentally, Bostrom has been around a good while, and even contributed to an SPT meeting in 1997.

Benjamin Hale, “Technology, the Environment and the Moral Considerability of Artefacts.” In this complex and difficult essay, Hale begins by recognizing three versions of a “pragmatic turn” in environmental ethics: to Peirce, James, Dewey and the American Pragmatists; to the Frankfurt school of neo-Marxists, including Marcuse and Adorno; and to “discourse theorists,” where he lists Apel and Habermas, as well as himself. A good third of the essay is then devoted to Habermasian theorizing, before Hale turns to his curiously abstract argument (in an essay supposedly devoted to a “pragmatic turn”) about the *lack* of “moral considerability” of artifacts in relation to environmental philosophy.

I pause here to make a point about Habermas and Ihde’s “third wave” in his foreword. Habermas, for some reason, always held back from any relationship to the Society for Philosophy and Technology (the home of this journal); and in all of this “new wave” book there are precious few references to anyone in the “third wave” except Ihde himself. Even Andrew Feenberg, an offshoot of the Frankfurt school (like Habermas himself) is rarely mentioned; and the same is true for recent proponents of Dewey as a philosopher of technology, or “technical” philosophers of technology such as Kristin Shrader-Frechette or Joe Pitt, or even Mario Bunge, who has a wide following among some European philosophers of technology. Whatever shortcomings these authors find in the “third wave” (I will get to an explicit claim, by Evan Selinger, in a moment), they seem to be shortcomings of Ihde himself and other phenomenological philosophers of technology. (I don’t mean to say that Selinger’s critique is not valid – as I will show when I get to him.)

Peter-Paul Verbeek, “Cultivating Humanity: Towards a Non-Humanist Ethics of Technology.” Verbeek’s approach is explicitly “postphenomenological,” consciously building on Ihde’s approach. But “non-humanism” in the essay also owes a good deal to the Bruno Latour of *We Have Never Been Modern* (1993). Verbeek’s very cautious conclusion is this: “Only by approaching the human as more-than-human does it become possible to adequately give shape to the respect for humanity the humanist tradition has rightly been defending for so long” (final sentence). Along the way, Verbeek uses the technology of antenatal ultrasound as the basis of his argument, rejects both Heidegger and anti-Heideggerians, and falls back on pre-modern virtue ethics as better than “modernism’s” favored duo of deontology and consequentialism. (I should

add that Verbeek's book, *What Things Do*, 2005, does constitute an advance over Ihde, a "new wave" in that sense.) Finally we come to the two essays in the volume that, in my opinion, best deserve the "new wave" label:

Evan Selinger, "Technology Transfer and Globalization: A New Wave for Philosophy of Technology?" Selinger begins with what are to me non-controvertible historical facts, that the Society for Philosophy and Technology was tardy in facing the globalization issue (the theme of its biennial conference only as late as 2007) and that such treatments of globalization as there have been among philosophers, not all of them self-consciously philosophers of technology, have been woefully abstract and have reflected a Western bias. To counter this, Selinger focuses, in this multiply nuanced essay, on a concrete case, "village phones," a "gift" of Grameen Banks primarily to women in rural Bangladesh. The result is an admirable case study, in the tradition of Science and Technology Studies, that both tries to eliminate Western bias *and* critiques non-Western critiques, on the ground in Bangladesh, of this technological development. It's about time, I would say, for such a melding of the STS case study approach with philosophy of technology. More traditional philosophers of technology of Ihde's "third wave" have done case studies, but not with Selinger's attention to concrete practice in a non-Western setting. This essay alone is almost worth the price of the book, and an equally good one follows.

Casper Bruun Jensen and Christopher Gad, "Philosophy of Technology as Empirical Philosophy: Comparing Technological Scales in Practice." What these Danish authors mean by "empirical philosophy" is a use by philosophers of anthropological-style approaches (they give due credit to Harold Garfinkel, *Studies in Ethnomethodology*, 1967; and Barney Glaser and Anselm Strauss, *Discovery of Grounded Theory: Strategies of Qualitative Research*, also 1967) in order to deal with concrete case studies. Their examples are the introduction of "bush pumps" in Zimbabwe (Annemarie Mol and Marianne de Laet) and "fishery inspection" on the vessel *Vestkysten* (*West Coast*; one of the two authors, Gad, did fieldwork on the ship in 2006 and 2007), though they also refer at length to Marilyn Strathern's "Enabling Identity? Biology, Choice and the New Reproductive Technologies" (1996) as well as to other concrete STS-type studies. Their conclusion, which makes empirical philosophy reflect the approach of the Social Construction of Technology (see Wiebe Bijker and John Law, eds., *Shaping Technology/Building Society*, 1992), is this: "Empirical philosophy assumes that we are often faced with technological situations of ambivalence, danger and possibility, in which indigenous and academic forms of action, value and conceptualization are associated and often at stake." And their last word is this: "In such cases we believe that this analytical mode offers a viable and interesting point of entry for a nuanced engagement with pressing technological matters of concern." To which I say Amen.

In short, though there is some originality here with respect to the reworkings of old material – some going all the way back to Ihde's first and second waves, but predominantly the third – it is doubtful that they offer much more than mere reworkings. Whether that – together with the five or six genuinely original essays – constitutes a new fourth wave or not, I would leave to readers of the book. The editors of the *New Waves in Philosophy* series clearly think so, but this member of Ihde's third wave has his doubts.

Paul T Durbin
University of Delaware

References

- Bijker, Wiebe, and John Law, eds. 1992. *Shaping Technology/Building Society*. Cambridge, MA: MIT Press.
Garfinkel, Harold. 1967. *Studies in Ethnomethodology*. Englewood Cliffs, NJ: Prentice Hall.

- Glaser, Barney, and Anselm Strauss. 1967. *Discovery of Grounded Theory: Strategies of Qualitative Research*. Chicago, IL: Aldine.
- Latour, Bruno. 1993. *We Have Never Been Modern*. Cambridge, MA: Harvard University Press.
- McLuhan, Marshall, and Eric McLuhan. 1988. *Laws of Media: The New Science*. Toronto: University of Toronto Press.
- Strathern, Marilyn. 1996. "Enabling Identity? Biology, Choice and the New Reproductive Technologies." In S. Hall and P. du Gay, eds., *Questions of Cultural Identity*. London: Sage. pp. 35-52.
- Verbeek, Peter-Paul. 2005. *What Things Do: Philosophical Reflections on Technology, Agency, and Design*. University Park, PA: Pennsylvania University Press.

***Engines of the Imagination: Renaissance Culture and the Rise of the Machine*, by Jonathan Sawday (Routledge, 2007). 424 pp. ISBN: 978-0415350611 (hbk), 978-04153562x (pbk), 978-0203696158 (ebk).**

This book is a rich source of information on an insufficiently researched topic. However, the philosopher of technology will have to use Sawday's material to make his or her own generalizations and interpretations.

The author surveys a vast amount of renaissance literature touching on technology. Research on technology and the humanities has concentrated primarily on the late seventeenth century and after. Second in coverage is medieval technology and culture, on which figures such as Lynn White and Jean Gimpel and their followers have written. However, the renaissance has been a relatively neglected area in the field of technology and literary culture. Sawday's work is extremely valuable for covering this area. However a great many of the sources he uses are from the mid-seventeenth century, on the border between the Northern Renaissance and the Age of Reason. Although the author seems, from his earlier works, to be most familiar with this period, it would have been desirable, given the theme and title of the work, to focus more on the Italian Renaissance of the fifteenth century and the Northern Renaissance of the sixteenth and very early seventeenth century.

Sawday has surveyed a huge amount of literature and culled references to technology. Some of the topics he covers are Montaigne's positive references to technology in his diary, which contrast with the negative attitude toward new technology in his essays; the religious and political symbolism of the moving of the immense obelisk on the Pope's orders in Rome; the role of the clock and of printing (covered briefly here, and among the few topics in renaissance technology and general culture that have been extensively covered by others); the writings of Agricola and Ramelli; Women and spinning as it appears in art of the renaissance; Bacon and Hooke on machines; Milton and the engine; and Andrew Marvell and Shakespeare on the natural and the mechanical. One of the intriguing aspects of this literature, one that the author emphasizes, is the fact that much of the technology that the literary figures discuss was seen as surprising and totally novel in a way that is hard for us to apprehend.

Sawday supplies numerous quotations from original sources in literature, and discusses, to a lesser extent, painters such as Valasquez and Bruegel. The author also shows familiarity with quite a bit of secondary literature, including theorists such as Lewis Mumford and Marshall McLuhan, more recent writers on philosophy and history of technology such as Langdon Winner, David Landes, David Noble, and Thomas Hughes, and historians of science such as Paula Findlen, Charles Webster, and Jan Golinski. Despite the broad framework of interpretive sources used, the work is seemingly lacking in strong overarching themes or sharply formulated conclusions.

Another limitation, aside from the extensive treatment of Montaigne and references to Leonardo, Agrippa, Alberti, and others, is that the work culls its material primarily from English sources. There is, of course, a vast literature on Leonardo's inventions and speculations, and an extensive one on Alberti and Italian renaissance architecture. Thus Sawday's work on Montaigne and Milton is original and welcome. However, as a survey of technology in the renaissance in general, the work, extensive as it is, is incomplete.

At times the author draws comparisons to later issues and themes such as Leo Marx's classical investigation of themes of machinery and nature in American literature. There are references to recent and contemporary events, such as the attempt to erase the past and return to an agricultural utopia in Pol Pot's Cambodia, Silvio Berlusconi's speech against Islam, and Paul Roberts on the "end of oil." However, I did not find these references to later issues and problems to be sufficiently systematically, or continuously developed and linked to the theses of the work.

There are several interpreters of renaissance technology in relation to the general culture whom Sawday does not utilize, despite the immense range of reference he deploys. One is Roger D. Masters' *Fortune is a River*, which discusses evidence for a collaboration between Machiavelli and Leonardo in canal building. Sawday mentions the collaboration in a footnote, but cites only a source that briefly casts doubt on its reality. Masters' work is very suggestive on the relations between notions of physical and political power in renaissance thought. Another is Paolo Rossi's *Technology and the Arts in the Early Modern Era*. The works of Edgar Zilsel, from which I suspect a major theme of Rossi's is taken, concern the social crisis that threw together partially literate technologists and experts in the crafts with literary and learned humanists lacking in practical technical knowledge. Two other figures that have interpreted the relation between the broader renaissance society and economy and the mechanical approach to nature are Franz Borkenau and Henryk Grossman. Zilsel, Borkenau and Grossman are all non-orthodox Marxists. Sawday does make a number of references to Marx, but not to twentieth century Marxists who discussed renaissance technology. The three just mentioned also wrote primarily in German, and Sawday's primary focus in terms of secondary sources is English. Borkenau elaborates on Marx's claim that Descartes saw the world through the eyes of manufacture. Grossman criticizes Borkenau (and is used by writers such as E. J. Dijksterhuis and to criticize Marxism in general), but develops a much more nuanced and historically accurate (in terms of time sequences of developments) case about the relation of the capitalist economy to the mechanical view of the world.

The conclusion of the book characterizes the book's theme as one of the natural and the artificial. However, I found the book to be more of a collection of separate essays than a continuously developed argument. This is not a major criticism, given the valuable material that the author surveys and presents. Nevertheless, the philosopher of technology will need to mine the book for examples to be interpreted, not for a major connecting theme or striking thesis.

I recommend this book as a treasure trove of fascinating quotations from English writers concerning technology. However, the issue of in what respects renaissance writings concerning technology differed from those in the Middle Ages and Enlightenment still needs to be developed.

Val Dusek

References

- Borkenau, F. 1987 "The Sociology of the Mechanistic World-Picture" *Science in Context* 1 (1): 109-127.
- Freudenthal, G. and P. McLaughlin, eds., *The Social and Economic Roots of the Scientific Revolution*, New York: Springer Publishers, 2009.
- Masters, R. 1998 *Fortune is a River: Leonardo Da Vinci and Niccolo Machiavelli's Magnificent Dream to Change the Course of Florentine History*, New York: Penguin Putnam Inc.
- Rossi, P. 1970. *Philosophy, Technology, and the Arts in the Early Modern Era*, New York: Harper and Row.
- Zilsel, E. 2000. *The Social Origins of Modern Science*, ed. by D. Raven, W. Krohn, and R. S. Cohen, Boston: Kluwer Academic Publishers.

***Philosophy and Design From Engineering to Architecture* edited by Pieter E. Vermaas, Peter Kroes, Andrew Light, and Steven A. Moore (Springer, 2009). 362 pp. ISBN: 978-9048127337.**

Philosophy and Design is an anthology of very interesting papers focused loosely on philosophy and technological design. The strength and weakness of the volume are the same, the broad range of contributions. The papers are diverse and generally of high quality, though the volume does not have a strong coherence. According to the introduction, the volume was developed with the insight that the growing complexity of engineering design reduces the distinction between engineering and architecture and, hence, architecture's long tradition of consciously influencing human interaction and social organization through design could be used to illuminate engineering design. For the most part, however, the cross fertilization of fields is confined to the Introduction and Part III where it is often implicit. I confess that the parallels between technological design and architectural design were not what drew me to the book. I read the book to get a sense of where the field of philosophy and technology is now, and to find out what well-established and new scholars in the field are thinking. In this regard, the volume did not disappoint.

After the Introduction, *Philosophy and Design* consists of three parts: Engineering Design; Emerging Engineering Design; and Architectural Design. In order to do justice to the range of papers, I will take up each part separately.

Part I Engineering Design

Each of the papers in Part I addresses a question or concern loosely centered around the design of technology. The set begins with Franssen taking on one of the deepest philosophical questions – the metaphysical status of artifacts. He argues that what an artifact is or is for, is indeterminate, and he draws out the implications of this for role of designers. Houkes argues for a use-plan analysis ("design involves the construction and communication of a use-plan"). Ihde uses the intentional fallacy as a parallel for understanding the many interpretations of technology design. Although these papers are quite distinctive in approach, all three wrestle with the mismatch between the intentions of artifact designers and the way artifacts may be understood and used. Brey takes on evolutionary accounts of technological innovation and change. Van Gorp and van de Poel, and Verbeek focus on ethical issues in design; Van Gorp and van de Poel give an account of ethical issues in engineering design and Verbeek gives an account of values in technological design. Both of these papers are crisp and insightful statements of ideas the authors have elaborated before. Feng and Feenberg give an account of design that synthesizes social theory, science and technology studies, and critical theory. Their argument is that technologies are underdetermined and, therefore, there are opportunities for alternative perspectives in the design process. Part I concludes with a paper by Naoe on culture being inscribed in technology and Thompson's account of how alienability, rivalry and exclusion costs are values that get expressed in institutional design.

Part II Emerging Engineering Design

The papers in Part II are focused on the future – the kind of technologies likely to be developed and the engineering and design practices that could or should be used in the future. The papers grapple in quite different ways with how to think about and what to do about the technological future that is coming. Sullins’ struggles with the role and nature of robots and how they will be integrated into human activity, in particular given the potential for affective robotics. Rieder and Schafer argue that open source software shows that technological design will in the future not just be a product of engineers or the institutions of engineering; there will be “fluctuations in how technical artifacts are created”. Nordmann argues that we may be witnessing a regression in the way technology is understood; he suggests that technology is being naturalized and might be thought of as being “as enchanted and perhaps frightening as nature used to be when humanity started the technological process of disenchantment and rationalization”. These papers are followed by a cluster on redesigning humans (human enhancement). Cerqui and Warwick provide a techno-enthusiastic account the redesign of humans since they describe what might happen to humankind and how “bright the future might be.” In stark contrast, Melo-Martin provides a more sophisticated, meta-analysis of genetic technology, pointing to the misunderstanding about knowledge creation in the current discussion and arguing for a better understanding that would help with the decisions humans will have to make about genetic enhancement. The cluster on human enhancement ends with Schmidt making some points about concepts that shape the question “should we redesign humans?” [One cannot help but wonder whether it was oversight or a political statement that led to the use of ‘Man’ in the title and within the paper.]

Part II ends with two papers on design methodology and two papers on responsibility. Miettinen uses design to examine systems methodology and systems engineering. In a paper that harks back to the discussion in Part I about the relationship between the artifact designers and artifacts, Krohs argues that artifact designs influence the design of a society but because societies are self-organizing, artifacts influence only to a minor degree. Neeley and Luegenbiehl focus on design as a discourse framework to be distinguished from the framework of technological development. They argue that the technological development framework brings with it the perception of inevitability while the design framework “enhances perceptions of choice and consequently, of individual responsibility.” Cook, also concerned about responsible design, argues that we must distinguish between natural, artifactual and human systems, though the three are interdependent. At a time when we are critically dependent on artifactual systems, Cook argues, a blurring of the distinctions between these three types of systems undermines “effective and responsible design”.

Part III Architectural Design

The six papers in Part III focus, in one way or another, on architecture. Three of the papers are by architects and three by philosophers. Davis sets the scene with an historical account of architecture emphasizing the split that took place in the 19th Century between designing and building. The paper concludes with a perspective on the current state of architectural practice. Moore and Webber focus on architects as experts who provide the public with a representation of reality – a representation that hides as well as reveals. The paper seeks to understand the consequences of architects making normative judgments that limit the range of choices others have. Cavanagh uses house construction as a case study. He seems to push against the purpose of the volume (i.e., to draw parallels between design contexts). Pitt elaborates earlier work on the criteria for successful design. Focusing on the Michael Graves complex in The Hague, he argues for what he characterizes as a common sense approach to design, an approach that draws on the insights of William James. Hanks responds to what he calls the crisis of cities in the U.S. He examines and evaluates two possible responses: New Urbanism and Civic Environmentalism. The book ends with a paper by Parsons in which he considers “the relationship between the

aesthetic appreciation of the built environment and the aesthetic appreciation of the natural environment". Parsons argues that these two aesthetics should not be thought of as opposed; he argues for producing built environments that mirror nature but he has a complex notion of nature.

I was glad that I read the book and was heartened to see philosophers of technology engaged in such rich and substantive discourse, discourse that has real-world implications. I was left with the impression that the field is far from coalescing around any solid girders of understanding but perhaps that doesn't matter.

Deborah G. Johnson
University of Virginia

***Pragmatism as Post-postmodernism: Lessons from John Dewey* by Larry A. Hickman (Fordham University Press, 2007). 284 pp. ISBN: 978-0823228423.**

As a specialist in John Dewey studies, Larry A. Hickman has made and continues to make contributions to the development of Dewey's philosophy. This new collection of papers from more than three decades of work is his latest approach defending and extending Dewey's classical pragmatism. As Hickman writes, his aim is to "extend the reach of John Dewey's insights into areas where they have so far had little or no recognition" (p. vii). In the pragmatist tradition, Hickman's concern is especially to help people everywhere promote their intelligent resources and practical capacities to solve social problems. Pragmatism here is the classical philosophical program that derives from Charles Pierce, John Dewey, and William James. For Hickman, productivity is central to this pragmatism, and thus the book works to "produce" creative artifacts for communities and not simply to think as an end in itself. In light of such an understanding, Hickman denominates Dewey's position as *productive pragmatism*. From the perspective of productive pragmatism technology is understood as "a natural activity of human beings, a part of their attempt to secure transitory goods and improve the conditions of their lives, both as individuals and groups" (p. 84). Like all productive human activities, it uses what Dewey calls the method of inquiry to seek and secure goods. The discussions in this book - on broadly overlapping topics such as postmodernism, neomodernism, globalization, and environment - all provide further examples of this method. The conscious reader will thus find the book's structure to be philosophical and even pragmatic. The technical route is to start from "theories" looked at as "practices" in relation to technology and its context. Afterward, using these reflections, argumentation moves back to theories again in order to advance classical pragmatism. Early in this book, Hickman locates productive pragmatism in the contemporary history of philosophy by comparisons with postmodernism and neopragmatism. Before the terminology of postmodernism was invented, classical pragmatism had already taken an antifoundationalist and deflationary attitude toward traditional metaphysics. However, Hickman's approach here is not simply to "postmodernize" pragmatism, but to make a distinctive philosophical argument that unlike highbrow postmodernism, Dewey and classical pragmatism provide "a theory of experimental inquiry that takes its point of departure from real, felt existential affairs" (p.29), in opposition to postmodernist cognitive relativism emphasizing difference, discontinuity and incommensurability. This is why Dewey's pragmatism can be called "post-postmodernism". Classical pragmatism, Hickman argues, also offers significant advantages over some currently popular versions of neopragmatism. For instance, Richard Rorty's neopragmatism blurs the distinctions between arts and technosciences and attempts to displace classical pragmatism's thick program of active experimental reconstruction with thinner projects that present hoping and coping as the best available paths to progress. By contrast, Dewey's classical

pragmatism honors the distinctive roles of the arts and technosciences and emphasizes their objective results over subjective attitudes we might take toward them. Dewey is thus more able to mobilize the pragmatic enthusiasm for engaging and solving social problems, especially those characteristic of technological culture. Here, Hickman is in agreement with other interpreters such as Junichi Murata, an active Japanese pragmatist who maintains that the Deweyan contribution to the ethics of technology is to solve sociotechnical problems by means of creative long-term technology assessment. After presenting his vision of pragmatism as post-postmodernism, Hickman thus turns to consider the specific advantages of Dewey's viewpoint for intractable issues of technology and environment.

Hickman is one of the earliest pragmatists to reconsider Dewey as primarily a philosopher of technology. Especially in *John Dewey's Pragmatic Technology* (1990) and *Philosophical Tools for Technological Culture* (2001), he has used a Deweyan approach to create theoretical and practical resources for disciplines such as the philosophy of technology and technology studies. Among three other social-critical philosophers of technology with whom he has entered into dialogue in these previous and the current book — Jürgen Habermas, Andrew Feenberg, and Albert Borgmann — Hickman argues that Feenberg's social-critical theory of technology is closest to Dewey. Hickman commends Feenberg for moving away from his teacher, Herbert Marcuse, toward the critique of technology advanced by Dewey. From a pragmatist perspective, Habermas and Borgmann are more deficient. Habermas places too much weight on the noninstrumental side of the unstable dualism of strategic action versus communicative action, and lacks an adjustive historicist perspective on human situations. Borgmann's device paradigm is too broad and seems to deprive humans of creative uses of technology.

Hickman argues that where technology fails, the problem is ourselves. It is our lack of ability to invent new tools and to criticize our own highly cherished values. Hence, Dewey's critique of technology in Hickman's narrative calls for "naturalizing" technology, locating it in a realm that is neither supernatural nor extranatural and in which the only telic elements are the natural ends of objects, individuals, and events, all of which in turn may become means to further ends. This leads directly to Hickman's treatment of environmentalism as a related practical theme amenable to a Deweyan perspective. In this section, Hickman compares Dewey's environmental naturalism with that of Aldo Leopold and some other green pragmatists. Dewey would accept much of their work in environmental philosophy, but his naturalism would not accept the idealized, nonhuman nature, or mystic ideals sometimes encountered in Leopold and others. Evolutionary naturalism is Dewey's main theoretical framework in all his reflections on the human world. In the last part of book, Hickman tries to encapsulate the central concepts in Dewey's classical pragmatism. These ideas include the theory of inquiry (what Dewey called "epistemology industry"), warranted assertibility, habits as artifacts and productive pragmatism (Hickman's key term). Instrumentalism and experimentalism are two highlighted methodologies. But this part also tries to think through classical pragmatism from a higher level, elaborating on earlier descriptions of classical pragmatism as a post-postmodernism. Although contextualist, productive pragmatism also promotes the creative invention of new "tools" to solve problems in different situations. Its experimentalist inquiry produces new artifacts, including new habits, making it more active than either postmodernism or neopragmatism. For scholars in philosophy of technology and other technology studies disciplines, this book offers two main contributions: First, compared with other books on pragmatist philosophy of technology, it presents a more theoretical and systematic account of Dewey's pragmatism. Second, the volume is an intelligent resource for philosophy and technology studies. More specifically, in problematical sociotechnical culture, it actually helps produce creative artifacts in the forms of tools to address social problems. In sum, Hickman's

most prominent achievement is to present classical pragmatism as a creative philosophy of production.

Qin Zhu

Dalian University of Technology

Colorado School of Mines

Dorsality: Thinking Back through Technology and Politics (Posthumanities Series) by David Wills (University of Minnesota Press, 2008). 280 pp. ISBN: 978-0816653454.

David Wills' *Dorsality: Thinking Back through Technology and Politics* begins with a neologism of sorts, a noun derived from an adjective referring to the backside of a body; under Wills' novel spin, it also refers to the back of our thought. Regarding the former, the notion of dorsality serves to describe the constitution of a human body and, as such, humanity; regarding the latter, it is a standard philosophical practice of looking back or beyond and into the customary conditions of possibility of philosophy. *Dorsality* is not a book about the latest technological developments in metallurgy or biotechnology; rather, it is a philosophical treatise concerning the conceptual framework that governs our understanding of technology.

As spelled out by Plato and Aristotle, and interpreted by Heidegger, Derrida, and now Wills, *techne* means both art and craft – that is, both artistic creation and technological production. To fully grasp the meaning of technology, one must inquire into the nature of both. Moreover, following the thesis of his *Prosthesis*, which according to Wills, is to be seen as a “back-ground” for *Dorsality*, there is no pure, natural, non-prosthetic origin; instead, everything is always already infused by the artificial (245). The same applies to humans: there is no pristine, simple human that later creates technology; instead, moving the timeline of evolutionary biology by following the anthropologist Leroi-Gourhan, Wills asserts that technology is literally embedded in our upright stance which in turn frees our thought-creating brains and tool-making hands. Technology as production/creation by humans of something other than human, as a differentiating force, is, after all, not something other than human.

Dorsality is made up of a series of critical readings of sources ranging from Exodus and Homer to Rimbaud, Sade, Heidegger, and Derrida. Given his background in literary theory and practice in deconstruction, Wills mainly focuses on the Western literary and philosophical tradition. Wills' method and style are decidedly deconstructive. Unlike his *Prosthesis*, *Dorsality* does not employ an elaborate personal autobiographical conceit and, as such, is more akin to his *Matchbook*. Thematically, it explores the areas of ethics, politics and sexuality. Wills references the standard bearers of continental philosophy and literary theory such as Blanchot, Barthes, Lyotard, Deleuze, Derrida, Nancy, and Lacou-Labarthe, as well as the more recent, rising stars such as Giorgio Agamben and Bernard Stiegler. On a more personal note, one of the concepts developed – namely, that of “leaving” as “the originary moment of thinking (and desiring)” – is said to be owed to Branka Arsić to whom, it may be inferred from a reference to the first name in the dedication, the book is dedicated (251).

Dorsality is “framed” by black and white reproductions of art works that precede epigraphs. The works range from Salvador Dalí and Frida Kahlo to Bill Viola, and even include a photograph of Emily Dickinson's tombstone by Wills himself. The common theme among most of these seem to be women and (fragmented, disintegrating) bodies.

“The Dorsal Turn” serves as an introduction of both the notion of dorsality and the rest of the book. In the “Facades of the Other: Heidegger, Althusser, Levinas,” in addition to discussing Althusser’s analysis, via the notion of interpellation, of the constitution of the political subject in “Ideology and Ideological State Apparatuses” and Levinas’ notion of ethical relation, Wills offers a reading of a number of Heidegger’s texts with a focus on “The Question concerning Technology.” Through the analysis of Heidegger’s shifts, turns and step-backs, Wills attempts to recover Heidegger’s rejection of technology.

“No One Home: Homer, Joyce, Broch” describes the odyssey of “polytropic” and “polytechnic Odysseus” by developing the concept of “originary exile” or “technotopological departure.” Departing from one end of Western literary history, the chapter arrives at the other by the examination of Joyce’s *Ulysses* and Broch’s *Death of Virgil*. Along the way, Wills asks: “What if, ‘before’ any act of creation or procreation, before any domestication via the womb or the earth, before any Earth Mother or Uranus, any Rangi or Papa, any Zeus or Hera, there were only the fiction of the same? What if the origin could only ever be conceived (of) in the form of such a construction, if the originary home were a possibility of a concept, a technotopological hypo-prosthesis that is the opening to inventing, to thinking and to fiction?” (82). In a similar fashion, “A Line Drawn in the Ocean: Exodus, Freud, Rimbaud” looks further into the formation of national identity, this time by means of, literally speaking, oceanographic exploration – that is, by describing the rhetorical force of the ocean in Exodus, Freud and Rimbaud.

“Friendship in Torsion: Schmitt, Derrida” examines the possibility of unnatural (technological, prosthetic) friendship as developed by Wills’ “sorely missed” friend Derrida in his analysis of Schmitt in *The Politics of Friendship*. “Revolutions in the Darkroom: Balász, Benjamin, Sade” is an essay in aesthetics that takes a penetrating look at dorsal sexuality via a series of reversals (theory/practice, aesthetic/political, nature/technological artifice) cinematically developing in Sade’s darkrooms. Moving from perversion to controversion, the final chapter, “The Controversy of Dissidence: Nietzsche,” examines Nietzsche’s deicide and concludes that: “Henceforth, whichever way we walk, we are all on Nietzsche’s path” (243).

Effectively demonstrating Wills’ dexterity and the breadth and scope of his interest, *Dorsality* is an excellent book. It is an essential reading for those practicing continental philosophy, aesthetics or literary theory. It could be an interesting read for those interested in philosophy in general or those engaged in broad, definitional aspects of technology studies.

Vladimir D. Thomas
Herzing University New Orleans