Erkenn (2014) 79:1367–1389
DOI 10.1007/s10670-014-9605-1

ORIGINAL ARTICLE

Function Ascription and Explanation: Elaborating an Explanatory Utility Desideratum for Ascriptions of Technical Functions

Dingmar van Eck · Erik Weber

Received: 24 June 2013 / Accepted: 4 February 2014 / Published online: 19 February 2014

Abstract Current philosophical theorizing about technical functions is mainly focused on specifying conditions under which agents are justified in ascribing functions to technical artifacts. Yet, assessing the precise explanatory relevance of such function ascriptions is, by and large, a neglected topic in the philosophy of technical artifacts and technical functions. We assess the explanatory utility of ascriptions of technical functions in the following three explanation-seeking contexts: (i) why was artifact x produced?, (ii) why does artifact x not have the expected capacity to ϕ?, (iii) how does artifact x realize its capacity to ϕ? We argue that while function ascriptions serve a mere heuristic role in the first context, they have substantial explanatory leverage in the second and third context. In addition, we assess the relevance of function ascriptions in the context of engineering redesign. Here, function ascriptions also play a relevant role: (iv) they enable normative statements of the sort that component *b* functions better than component *a*. We unpack these claims by considering philosophical theories of technical functions, in particular the ICE theory, and engineering work on function ascription and explanation. We close the paper by relating our analysis to current debates on the explanatory power of mechanistic vis-à-vis functional explanations.

1 Introduction

The philosophy of technical artifacts is gaining momentum. Whereas it was initially assumed that analysis of technical artifacts and technical functions was a rather

D. van Eck (✉) · E. Weber
Department of Philosophy and Moral Science, Centre for Logic and Philosophy of Science, Ghent University, Blandijnberg 2, 9000 Ghent, Belgium
e-mail: Dingmar.vanEck@Ugent.be

E. Weber
e-mail: Erik.Weber@Ugent.be

trivial task for philosophy, and technical functions could easily—and in passing—be accounted for by theories of biological functions, analyses developed in the past 10–15 years have proven otherwise (e.g., Preston 1998; Vermaas and Houkes 2003). The phenomenology of technical artifacts and technical functions presents intricacies that are not well accounted for by theories of biological functions. At present, a number of separate analyses focusing on the technical domain have been advanced, addressing issues such as theories of technical functions (Vermaas 2006; Houkes and Vermaas 2010), mechanistic artifact explanation (de Ridder 2006; de Winter 2011), the epistemology (Houkes 2006) and ontology of technical artifacts (Houkes and Meijers 2006), and comparisons of the dual–intentional and structural—'nature' framework of technical artifacts viz-a-viz collectivist frameworks (Houkes et al. 2011).

While we consider these developments valuable, we argue that something vitally important is missing in current theorizing about technical artifacts and technical functions, to wit: careful reflection on the explanatory relevance of technical function ascriptions. When and why are function ascriptions explanatorily relevant? Basically, current philosophical theories of technical functions are mainly concerned with specifying conditions under which agents are *justified* in ascribing functions to technical artifacts (and their components and processes). Yet, assessing the precise *explanatory relevance* of such function ascriptions is, by and large, a neglected topic in the philosophy of technical artifacts and technical functions.[1] The focus lies on developing normative accounts for *justifiable function ascription*, rather than on utility assessments of function ascriptions. In this paper we address this lacuna.

We assess the explanatory utility of ascriptions of technical functions by considering three explanation-seeking contexts that often figure in the philosophical literature on functions (and explanations). Applied to the technical domain, these are:

i why was artifact x produced?
ii why does artifact x not have the expected capacity to ϕ?
iii How does artifact x realize its capacity to ϕ?

In addition, we analyze the utility of function ascriptions in the context of engineering redesign, i.e., a context in which explanation is intertwined with design. We there focus on the utility of function ascriptions for making claims of the sort that:

iv component *b* functions better than component *a*

In addressing the first question we use the "ICE" theory of technical functions, in which elements from Intentionalist, Causal role, and Evolutionist theories of function are incorporated, as an instrument to assess the relevance of functions ascriptions. We argue that on the basis of the ICE theory, two parallel explanations can be constructed for the first explanation-seeking question, a functional one that incorporates function ascriptions and a teleological one that does not. We argue that,

[1] Cf. (Preston 1998; Kroes 2003; Vermaas and Houkes 2003; Krohs 2009; Houkes and Vermaas 2010).

in this explanatory context, the teleological explanation is superior to the functional explanation. The functional explanation black-boxes relevant difference making properties with respect to occurrence of the phenomenon to be explained that are included in the teleological one. We also indicate how the alternative function theories of Preston (1998) and Krohs (2009) fare with respect to this explanation seeking context. We argue that, for similar reasons as in the case of the ICE theory, on these alternative theories, function ascriptions turn out explanatorily irrelevant for the first explanation-seeking question. Hence, we conclude that, in this context, function ascription is, at best, merely heuristically useful in guiding the construction of a satisfactory explanation which does not invoke function ascriptions.

Our analysis of the second explanation-seeking context of explaining artifact malfunction, does reveal substantial utility of function ascriptions. By considering an engineering methodology for the analysis of artifact malfunction, developed by Price (1998) and Bell et al. (2007), we show that functional explanations are explanatorily superior to full—blown mechanistic explanations. Functional explanations, here, black-box irrelevant causal details and thereby focus on relevant difference making properties with respect to explaining artifact malfunction. Hence, function ascription is required to construct satisfactory explanations.

In the third explanatory context of capacity explanation, we argue that function ascription and mechanistic explanation work in tandem. By considering an engineering methodology for reverse engineering, developed by Otto and Wood (1998, 2001), we show that (systemic) capacities are best explained by providing mechanistic explanations. These explanations provide the most complete story on relevant difference making properties. Yet, we also argue that in order for mechanistic explanations to convey in-depth understanding to agents with respect to the phenomenon to be explained, the functional contributions of the components' behaviors need to be known. Hence, function ascription is also substantially relevant in this context. This utility of function ascriptions carries over to the final context we consider, namely artifact redesign. There, we argue, function ascriptions are relevant for comparing the functional performance of components' behaviors.

In arguing these points we employ a key desideratum from several accounts of explanation (Woodward 2003; Strevens 2004; Couch 2011; cf. Weisberg 2007) according to which those, and only those, factors that make a difference to the occurrence of a phenomenon to be explained should be referred to in an explanation.

In considering the second, third, and final context we closely engage functional modeling practices in engineering science. Function is a key concept in engineering, yet not a unitary one (Erden et al. 2008). Rather, different meanings of the term 'function' are systematically used side-by-side in engineering functional modeling practice. This engagement with engineering functional modeling also enables us to clarify the utility of specific notions of function in the contexts we consider.

We close the paper by discussing a ramification of our analysis for the philosophy of (mechanistic) explanation in general, to wit: assessments of the explanatory force of mechanistic vis-à-vis functional explanations. Against the notion that 'complete' mechanistic explanations have more explanatory power than less elaborate functional ones *tout court* (cf. Machamer et al. 2000; Craver 2007; Piccinini and Craver 2011), we argue that it depends on the request for explanation whether functional or mechanistic explanations are to be preferred. We argue that in the context of malfunction explanation, functional explanations fare better than mechanistic ones. In this discussion we also briefly consider explanation in the biological domain.

We proceed in the next section with a brief description of the core tenets of the ICE theory after which we address the first explanation-seeking context. We then discuss engineering interpretations of technical function in section three. In the fourth section we consider artifact malfunction, and in the fifth section we analyze capacity explanation and present our engineering redesign case. In section six we briefly expand our analysis to current discussions on the explanatory force of mechanistic explanations. We present our conclusions in section seven.

2 Functional Versus Teleological Explanation: Why was Artifact x Produced?

Here we employ the ICE theory of technical functions (Houkes and Vermaas 2010) as a conceptual instrument to assess the explanatory utility of function ascriptions with respect to the explanation-seeking question:

i why was artifact x produced?

We choose to focus in-depth on the ICE theory in our analysis since it, in our view, provides the most sophisticated theory on technical functions, and provides the richest conceptual apparatus to address this question. It invokes more relevant difference-making factors when compared with alternative function theories. Yet, the results we present in this section are not conditional on use of the ICE theory. After our assessment in terms of the ICE theory, we indicate how the alternative function theories of Preston (1998) and Krohs (2009) fare with respect to the above explanation seeking question. As in the case of the ICE theory, also on these alternative theories, function ascriptions turn out heuristic.

2.1 The ICE Theory of Technical Functions

The most elaborate statement of the ICE theory of technical functions can be found in *Technical functions: on the use and design of artefacts* (Houkes and Vermaas 2010). The authors chose a normative rather than descriptive approach:

> This choice means that we approach both artefacts and the actions in which they play a role largely from a *normative* rather than a descriptive perspective. We do not offer a theory about how people actually use or design artefacts, or how they in fact describe them in functional terms; instead we seek to provide a framework for evaluating some aspects of these activities, and we theorise about rational and proper artefact use, and about justifiable function ascriptions. (p. 4)

The authors review three function theories for technical artifacts: the intentional (I) theory (Neander 1991; Bigelow and Pargetter 1987; McLaughlin 2001; Searle 1995), the causal-role (C) theory (Cummins 1975) and the evolutionist (E) theory

(Millikan 1989).[2] After exposing the weaknesses of these theories, Houkes and Vermaas propose their own theory, which is called the ICE-theory, because it combines insights from the three basic theories.[3] Function ascriptions to artifacts are analyzed against the background of artifact use and design. The use of an artifact is viewed as the carrying out of a use plan for the artifact. Design is seen as primarily the development of new use plans for artifacts. Another important feature is that the theory is agent-oriented rather than property-oriented: the ICE theory takes the form of a theory of *justifiable function ascriptions* by human agents rather than a theory that identifies functions as properties of artifacts.

The core of the theory comprises two definitions of justifiable functions ascriptions (one for designers or justifiers, one for passive users; see 2010, pp. 88–89). These definitions can be merged into a single definition. At the EPSA 2011 symposium in which the book was discussed, Houkes and Vermaas proposed the following general definition, which does not distinguish between the two types of agents:

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

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

We will use this definition in our analysis. As can be seen, the ICE theory is a *normative* theory about *justifiable function ascription*: it concerns when function ascriptions are justified and how they have to be justified.

Although the question *why and under which conditions function ascriptions are explanatorily useful* is—as in other theories of technical function—not explicitly addressed, the ICE theory can be invoked to address this issue. We do so here with respect to the following question:

i why was artifact *x* produced?

2.2 Heuristics of Technical Function Ascriptions

We argue that by applying the ICE theory to answer the question why an artifact *x* was produced, two parallel explanations can be constructed, a functional one and a, what we may call, teleological one. Whereas the former, by definition, contains function ascriptions, the latter does not. The question, now, is, which explanation is to be preferred? We address this question in terms of the notion, emphasized in several accounts toward explanation, that those, and only those, factors that make a

[2] Neander's (1991) theory counts as an evolutionist one in the context of biology. Applied to technology, it becomes an intentionalist one (Houkes and Vermaas 2010).

[3] Since these weaknesses have no bearing on the argumentation presented in this paper, we do not elaborate them here.

difference to whether or not a phenomenon to be explained occurs should be specified in an explanation (Woodward 2003; Strevens 2004; Couch 2011; cf. Weisberg 2007).[4] Applying this constraint or desideratum has substantive implications: in the explanation-seeking context under consideration, function ascription and functional explanation have a mere heuristic role and, we argue, teleological explanation is to be preferred.

2.2.1 Case 1: Backward Looking Explanation

The first type of cases we consider are questions of the following form:

1. Why was artifact *x* produced?

   With respect to the ICE theory, functional explanations that we give to answer such questions have the following format:

2. Artifact *x* was produced because there was a designer *d* who justifiably ascribed the physicochemical capacity ϕ as a function to x.[5,6]

   Let us consider an example:

[4] Note that this *desideratum* is different from the theory or model constraint of 'simplicity'. When endorsing 'simplicity' a theorist or modeler may intentionally exclude reference to factors that make a difference to whether or not a phenomenon occurs. The constraint which we endorse here, requires that an agent should strive for describing all the factors that make a difference to whether or not a phenomenon occurs. Whether an agent succeeds in doing so is, of course, a different matter. Weisberg (2007) labels this constraint an "1-causal" representational ideal, and distinguishes it from the representational ideals of "simplicity" and "completeness". The latter requires that an explanation should specify both difference making properties with respect to whether or not a phenomenon occurs, as well as the "higher order causal factors" that affect the precise manner in which the phenomenon occurs (cf. Weisberg 2007, p. 651).

[5] An anonymous referee pointed out that (justified) function ascription could have played no role in answering the first explanation-seeking question since there was no physical artifact yet to which a designer could have ascribed a function to. Agreed, yet our answer is in keeping with the ICE theory: "The historical perspective required to ascribe ICE functions may be limited to the design process; it need not extend to earlier generations of artefacts. *An artefact can therefore straightaway be ascribed the capacity for which designers selected it, even if the artefact is a completely novel one* (the case of the first nuclear plant)" (Houkes and Vermaas 2010, p. 93) (our italics). In other words, the answer accords with the ICE theory. To be sure, we here take function ascriptions as answers to the explanation-seeking question under consideration to be 'proper' function ascriptions. Proper function ascriptions are discussed by Houkes and Vermaas (2010) against the backdrop of what they call 'proper use plans'.

[6] An anonymous referee pointed out that regarding production, belief initially is sufficient and justified belief only becomes relevant in continuation of the production process. Again, agreed. However, justified belief is central to the ICE theory, both in the ascriptions of functions to technical artifacts, and in accommodating central desiderata put forward in the function literature, such as the proper-accidental function distinction, function ascription in innovative contexts, and the handling of malfunction statements. The underlying reason is that the ICE theory is a "normative rather than a descriptive perspective" on "justifiable function ascriptions" (Houkes and Vermaas 2010, p. 4). Given this perspective, the requirement of justified belief for explaining the production of an artifact is either a bullet one has to bite when adopting the ICE theory, or the ICE theory should be extended to also encompass a descriptive perspective in which 'mere belief' suffices for explaining the production of an artifact. Hence, our use of the term 'justified'.

3. Why was the computer mouse produced?

   A possible answer is:

4. The computer mouse was produced because there was a designer *d* who justifiably ascribed the capacity to indicate X–Y positions on computer screens as a function to the computer mouse.
   Another possible answer that can be constructed in terms of the key concepts invoked in the ICE theory, is the following non-functional one:

5. The computer mouse was produced because there was a designer *d* who had a use plan *up* for it and an account A. *d* believed (i) that the computer mouse has the capacity to indicate X–Y positions on the computer screens, (ii) that *up* leads to its goals due to, in part, this capacity. *d* could on the basis of *A* justify these beliefs. *d* communicated *up* and testified these beliefs to other agents.

So we here have two explanatory formats: a functional explanation (2, exemplified in 4) and a teleological explanation (5, with some details filled in). Now, the latter more elaborate explanatory format naturally leads to several follow-up questions: who was the designer *d*? What was the use plan s/he had in mind? What was the goal? To whom were the beliefs communicated? For instance, the goal may have been to facilitate computer use by feeding commands into the CPU without touching the keyboard. And the people to which the beliefs were communicated may include production managers, financial and marketing managers and the general manager of the enterprise in which the designer is working.

Given the constraint that an explanation should specify those factors that make a difference to whether or not a phenomenon occurs – here the production of artifact x—, a satisfactory explanation of the fact that the computer mouse was produced should include the details referred to in these additional questions. Information on the designer(s), goal(s), use plan(s), and agents involved in the communication chain(s), is crucial to understand how a given artifact x came to be: a design for a computer mouse without an accompanying use plan for it, nor a specified goal for which it can be employed, and neither a financial and marketeering strategy to put the product in the market, simply will not go into production.[7]

Now, the information about the designer can be included without giving up functional talk:

6. The computer mouse was produced because Douglas Engelbart justifiably ascribed the capacity to indicate X–Y positions on computer screens as a function to the computer mouse.

However, the rest of the required information cannot be communicated by means of function talk: from explanation (6) we cannot derive what Engelbart's use plan was,

[7] We focus on those difference making factors that are part of the conceptual framework of the ICE theory, and do not consider other potential difference making factors, such as, say, the choice of materials for the computer mouse. Therefore, our labelling of the notion that explanations should specify difference-making factors as a *desideratum* (cf. note 4). That there are, in the explanatory context under consideration, other difference making factors does not affect the outcome of our comparison of the explanatory superiority of functional vis-à-vis teleological explanations.

what his account was, to whom he talked, etc. So this explanation has a *heuristic* role: it is a first step towards a more satisfactory explanation. And, importantly, this satisfactory explanation does not employ function talk: function ascription is removed in order to fill in other, more detailed, information: his use plan, goals, communication partners, etc.

The point generalizes: explanations that fit in scheme (2) are only a first step, even if we include the name of the designer(s) and the capacity, as we did in (6). The satisfactory explanation requires an implementation of the following scheme:

7. Artifact x was produced because there was a designer *d* who had a use plan *up* for it and an account A. *d* believed (i) that *x* has the capacity to ϕ, (ii) that *up* leads to its goals due to, in part, this capacity. *d* could, on the basis of *A*, justify these beliefs. *d* communicated *up* and testified these beliefs to other agents.

In this teleological scheme, the word 'function' does not occur. So, in the explanations in which the factors are specified that make a difference with respect to the phenomena to be explained there are no function ascriptions.[8] In other words, in this context, functional explanations black-box relevant difference making properties with respect to the occurrence of the phenomenon to be explained, which are included in the teleological explanation.

This result is not conditional on use of the ICE theory. Also when applying Preston's (1998) pluralist theory of (biological and) technical function, function ascription turns out irrelevant with respect to the explanation-seeking question "why was artifact x produced". Preston invokes both the concepts of 'system (or causal role) function' and 'proper function' in the ascription of technical functions to capacities of artifacts. She argues that intended capacities for which artifacts are constructed by designers or inventors initially only have or can be ascribed system/causal role functions (p. 243, pp. 249–250). It is only when artifacts continue to be reproduced, that proper functions can be ascribed to those capacities for which the artifacts were reproduced, and this continued production is contingent on successful performance as determined by users, not designers or inventors (pp. 244–245).

Applying Preston's account, a possible answer to the explanation-seeking question "why was artifact x produced" has the following format:

> Artifact x was produced because a designer or inventor intended artifact x to perform a certain capacity, to which s/he ascribed a system function.

Now, the last clause 'to which s/he ascribed a system function' adds no explanatorily surplus to the explanation and thus should be removed from it. The fact that a designer or inventor constructed an artifact to perform a certain capacity that s/he desired, suffices. Designers/inventors and desired capacities are the difference making factors here, not the ascription of system functions.

Applying Krohs' (2009) theory leads to the same conclusion that function ascriptions have no added explanatory value in this explanation-seeking context. On Krohs' (2009)

[8] Note that the argumentation presented here is not to be confused with conceptual explication of the term 'technical function'. On the ICE account, 'technical function' refers to a physical–chemical capacity. We here invoke the ICE function ascription machinery to construct two parallel explanations.

account of (biological and) technical function, function is explicated in terms of the causal role concept of function and the notion of 'general design'. General design is defined as the 'type-fixation' of, in the case of technology, components of designed artifacts, i.e., the process by which a configuration/organization of components is brought about. Such processes include construction and assembly plans (pp. 74–75). On this account: "function is the contribution of a type-fixed component to a capacity of a system that is the realization of a design" (p. 79). In the context of artifact designing, a function is 'intended' if a component should make a certain contribution/perform a certain role in order to achieve the goal(s) of a designer (p. 85).

Applying Krohs' account, in the case of components, a possible answer to the explanation-seeking question "why was artifact x produced" has the following format:

> Artifact x was produced because a designer intended artifact x to make a certain contribution to a capacity of a system in order to achieve his/her goals.

A possible answer in the case of a system composed of a configuration of components has the following extended format:

> Artifact x was produced because a designer intended the components making up the artifact to make certain contributions. The system, in turn, is constructed via type-fixation processes, such as construction and assembly planning.

Again, in both scenarios, the ascription of a function here is irrelevant for explaining artifact production. Designers, goals, construction and assembly plans, and contributions are the difference making factors here. Function ascription adds nothing.

In considering our next explanation-seeking context of malfunction explanation, the situation is very different: there, we argue, black-boxing is precisely what gives functional explanations their explanatory leverage vis-à-vis mechanistic ones. Before presenting this argumentation, we briefly consider different meanings of the term 'function' as used in engineering. We then also clarify the explanatory utility of specific notions of function in the contexts we consider in the remainder of this paper.

3 Imposing Precision on Technical Functions: Engineering Meanings of Function and Functional Decomposition

Function is a key term in engineering (e.g., Chandrasekaran and Josephson 2000; Stone and Chakrabarti 2005; Kitamura et al. 2005), yet has no uniform meaning: different functional modeling approaches advance different concepts (Erden et al. 2008), and some researchers use the term with more than one meaning simultaneously (Chandrasekaran and Josephson 2000; Chakrabarti 1998; Deng 2002; Srinivasan and Chakrabarti 2009; Srinivasan et al. 2012). Vermaas (2009, 2011) has regimented three '*archetypical*' engineering concepts of function:

- *Behavior function*: function as the desired behavior of a technical artifact
- *Effect function*: function as the desired effect of behavior of a technical artifact

- *Purpose function*: function as the purpose for which a technical artifact is designed

The concept of behavior function is advanced in several engineering design and reverse engineering methods (Pahl and Beitz 1988; Stone and Wood 2000; Chakrabarti and Bligh 2001; Otto and Wood 2001). In these methods, a function is described as a conversion of flows of materials, energy and signals, where input flows and output flows in the conversion (are assumed to) match in terms of physical conservation laws (cf. Otto and Wood 2001). The function "convert electricity to torque" of an electric screwdriver's motor, for instance, is represented as a conversion of an input flow of "electricity" into corresponding output flows of "torque", "heat", and "noise" (cf. Stone and Wood 2000, p. 364). Since these descriptions of functions are specified such that input and output flows match in terms of physical conservation laws, they are taken to refer to specific physical behaviors of technical artifacts (cf. Otto and Wood 2001; Vermaas 2009; van Eck 2011a, b).

Effect function descriptions are also used in design methods (Lind 1994; Deng 2002), as well as in knowledge management methods (Kitamura et al. 2005) and diagnostic reasoning approaches (Bell et al. 2007). There, functional descriptions refer to only the technologically relevant *effects* of the physical behaviors of technical artifacts: the requirements are dropped that descriptions of these effects meet conservation laws and that matching input and output flows are specified (Vermaas 2009; van Eck 2011a). The function of an electric screwdriver's motor is then described simply as, say, "produce torque", leaving it unmentioned what the physical antecedents are of this effect.[9]

Purpose function descriptions are also employed in engineering design (Gero 1990; Chakrabarti 1998; Deng 2002). Such descriptions refer to by designers intended states of affairs in the world, which are to be created by the physical behaviors and effects of the technical artifact concerned (Vermaas 2009; van Eck 2011a, b).[10] The function of an electric screwdriver's motor is then described as, say, "having a rotational force".

Engineering descriptions and explanations of the workings of technical artifacts and artifacts-to-be-designed often are constructed by breaking down/functionally decomposing functions into a number of other (sub) functions. The relationships between functions and sets of their sub functions are often graphically represented in functional decomposition models. Like the concept of function, such models come in a variety of flavors. Elsewhere, one of us regimented this diversity in terms of the three archetypical engineering concepts of function (van Eck 2011a):

- *Behavior functional decomposition*: a model of an organized set of behavior functions;
- *Effect functional decomposition*: a model of an organized set of effect functions;

[9] Another example illustrating the distinction, given by Vermaas (2009), are the functions of a sound barrier: its behavior function can be described as 'converting acoustic energy to thermal energy' and its effect function as 'absorbing sound'.

[10] In methodologies that advance effect and/or purpose functions the concept of behavior is typically introduced as well, and through descriptions of the behavior of technical artifacts the physical conservation laws are taken into account.

- *Purpose functional decomposition*: a model of an organized set of purpose functions.

The use of functional decomposition is ubiquitous in engineering science. Stone and Wood (2000) use behavior functional decompositions in, for instance, the conceptual phase of engineering design to analyze the desired functions of some artifact-to-be, and in the reverse engineering of existing artifacts for archiving functional descriptions of these artifacts and their components. Otto and Wood (2001) also use behavior functional decompositions in reverse engineering tasks to determine the organized components and sub functions (behaviors) of artifacts—their mechanisms –, by which artifacts produce their overall (behavior) functions. Bell et al. (2007) use effect functional decompositions for explaining malfunctions of artifacts. Finally, Deng (2002) uses purpose functional decompositions in the conceptual phase of engineering design.

Against this backdrop of diverse engineering meanings of function and functional decomposition, we return to our main objective of assessing the explanatory relevance of ascriptions of technical functions, and in doing so comment on the utility of specific engineering interpretations of the term.[11]

4 Malfunction Explanation

In the situation described so far functional explanations are not optimal for explaining why an artifact x was produced: there is a non-functional/teleological alternative that is better. We now move on to our second explanatory context: diagnostic reasoning. There, we argue that function ascriptions and functional explanations provide the most satisfactory explanations. We make our case by discussing an engineering methodology for malfunction analysis.

A widely adopted desideratum in the literature on technical functions is that function theories should advance a notion of *proper* function that allows malfunctioning. In different accounts, this is done in different ways. According to the ICE theory, agents that ascribe functions to capacities of artifacts should be able to justify their beliefs that those artifacts have these capacities on the basis of either experience, testimony, or scientific or technological knowledge (the account *A*). Nevertheless, this measure of support, in principle, leaves open the possibility that an artifact malfunctions, despite the agent's (erroneous) belief that the artifact does have the capacity. Hence, malfunction is accommodated within the ICE theory. Krohs (2009) proceeds in different fashion. Rather than justified yet erroneous belief as in the ICE theory, in Krohs' theory, the notion of type fixation determines standards for the contributions of components which they can fail to achieve. Similarly, in the account of Preston (1998) successful performance as measured by users provides a yardstick to accommodate malfunction. Yet, of course, the

[11] In van Eck (2011a, b) the relevance of ascriptions of functions is analyzed in the context of routine designing, and the conversion of functional descriptions across routine design frameworks. In this paper we analyze the explanatory utility of function ascriptions in other and more varied contexts of engineering design, i.e., artifact production, failure analysis, reverse engineering, and redesign.

accommodation of malfunctioning artifacts within schemes for the ascription of functions to technical artifacts, is completely different from *explaining* the occurrence of malfunctioning artifacts. Notions like ‘justified yet erroneous belief’ (ICE theory), ‘unsuccessful performance as measured by users’ (Preston), and ‘not meeting standards for components’ contributions’ (Krohs) are not difference making factors that explain the occurrence of specific malfunctions. Malfunction explanation requires (contrastive) explanation that isolates the specific fault(s) that cause malfunction(s).

Therefore, we here focus on engineering diagnostic reasoning methods invoked to explain occurrences of malfunctions in technical artifacts, and clarify the structure of the explanatory formats that these methods advance, to wit: *contrastive* functional explanations.

When an artifact does not serve a function which we expect it to do, explanation-seeking questions of the following format arise:

Why does artifact x not serve the expected function to ϕ ?
For instance: why does this electric screwdriver fail to drive screws?

Such questions are *contrastive*: they contrast the actual situation with an ideal and expected one (cf. Lipton 1993). Now, malfunction explanations that answer contrastive questions have a different format than reverse engineering—mechanistic—explanations which answer questions about plain (non-contrastive) facts, such as explanations of why an artifact has a certain capacity (e.g., an electric screwdriver’s capacity to drive screws). Contrastive or malfunction explanations, as developed in engineering by, for instance, Price (1998), Hawkins and Woollons (1998), and Bell et al. (2007), pick out only those causal factors that make a difference to the occurrence of a specific malfunction, rather than also specify those factors that both normal functioning and malfunctioning technical artifacts have in common and which do not affect the occurrence of the malfunction under consideration.[12] Judged by the format of these explanations, a full mechanistic story on the organized components and behaviors of an artifact is overkill for malfunction explanation. Most information about other components, their behavior, and the manner in which the (malfunctioning) component is organized with them is left out.

We take it that this is focus on only the (expected) properties that make a difference with respect to the occurrence of a specific artifact malfunction is done for sound reasons. For instance, the mechanism of an electric screwdriver by which it fulfills its function of driving screws likely has a substantial number of elements in common with mechanisms underlying a malfunction of this type of screwdriver, say, the complete failure to drive screws or the driving of screws without sufficient torque. It might be the case that, say, the conversion of electricity into torque is sub-optimal in the malfunctioning screwdriver, whereas most other operations are similar in both normally functioning screwdrivers and in the dysfunctional one. In both cases, say, components generate electricity, transport electricity, and insulate

[12] Some of these factors that normally functioning and malfunctioning artifacts have in common might affect the precise manner in which a malfunction occurs, yet do not affect the occurrence itself. Weisberg labels factors that affect the precise manner in which a phenomenon occurs, “higher order causal factors” (cf. Weisberg 2007, p. 651).

heat and noise etc. Only the sub-optimal electricity-into-torque conversion then marks a relevant *contrast* between functioning and malfunctioning artifacts and, hence, should be specified in a contrastive malfunction explanation. It is, in this example, this factor that makes a difference to whether or not the specific malfunction phenomenon will occur. These other operations might affect the precise manner in which the malfunction phenomenon occurs but are explanatorily irrelevant for explaining the occurrence of a specific malfunction.

Hence, rather than an elaborate description of organized component parts and behaviors, a mechanistic explanation (Machamer et al. 2000; Glennan 2005; Bechtel and Abrahamson 2005; Craver 2007), an explanatory format is required that pinpoints what has gone wrong. This can be done by a function ascription, or more precisely, by a contrastive functional explanation describing a component *malfunction*. Such explanations provide the most insightful way to explain what went wrong. Most information about the other entities, their behavior, and the manner in which the component is organized with these entities and behaviors is better left out, since these details are explanatorily irrelevant and only obscure the difference making factors that matter.

Consider, by way of example, a methodology for malfunction analysis and explanation, called Functional Interpretation Language (FIL), developed by Bell et al. (2007).[13] The FIL methodology was developed and is used in industry for a variety of diagnostic reasoning tasks, in particular Failure Mode and Effect Analysis (FMEA). In short, in FMEA analyses, the effects of a malfunctioning component on the overall behavior of an artifact are analyzed, by comparing the overall behavior of artifacts working correctly with the overall behavior of ones that do not, due to a component failure/malfunction (e.g., Price 1998; Hawkins and Woollons 1998; Bell et al. 2007).

In FIL, the representation of a technical function consists of three elements: the *trigger* of a function, its associated and expected *effect*, and the *purpose* that the function is to fulfill. Triggers and effects describe (appropriate) behavioral states of components of a technical artifact, and purposes describe desired states of affairs in the world that obtain when a trigger results in an expected effect (Bell et al. 2007, p. 400). For instance, with FIL, the function of a stop light of a car is described in terms of the trigger "depress_brake_pedal", the effect "red_stop_lamps_lit", and the purpose "warn_following_driver" (p. 400). This functional description conveys the idea, that car drivers in the direct vicinity are being informed that a car is slowing down when its (right and left) red stop lamps are lit, as a result of the car's brake pedal being depressed (cf. Bell et al. 2007, p. 400).

Now, as Bell et al. (2007) stated, trigger and effect representations serve two explanatory ends in malfunction analyses: firstly, they highlight relevant behavioral states/properties, and, simultaneously, provide the means to *ignore* less relevant

[13] The FIL is one of the most visible methodologies in engineering fault analysis; work on the FIL dates back to the late '90 s (cf. Price 1998), and continues to be further elaborated to this day. The approach is well-entrenched in the broader literature on 'function' in engineering, building upon classics in the field (like Chandrasekaran and Josephson 2000) and, in addition, is not only successful in the academic engineering literature, but also successfully employed in industry (Bell et al. 2007). The FIL is developed both for automated failure analyses as well as intended as a method for fault analysis done by human engineers.

behavioral aspects of a given artifact; secondly, they provide the means to assess which components are malfunctioning (pp. 400–401).

For instance, the trigger representation “depress_brake_pedal” highlights the behavioral state of a pedal being depressed, yet ignores the behaviors and components of, say, the pedal lever mechanism(s) itself. And the effect representation “red_stop_lamps_lit” likewise ignores, say, the behaviors and components comprising the electric circuitry of the stop lamps. It represents only a desired and expected effect. Thus, trigger-effect descriptions of functions represent factors that make a difference with respect to the occurrence of malfunction, and black-box factors that are explanatorily irrelevant.

Functional descriptions in terms of triggers and effects support the analysis of the actual states of triggers and effects, i.e., assessment of whether the expected behavioral states in fact obtain, and support assessing which and how components are malfunctioning (Bell et al. 2007). A normally functioning artifact, say the car’s stop lights, has both a trigger and an effect occurring; the brake pedal is depressed and the stop lights are lit. Trigger-effect descriptions support analysis of two varieties of malfunction. First, a trigger may occur, yet fail to result in the intended effect. Say, the brake pedal is depressed, yet the stoplights are not on. Second, a trigger may not be occurring, yet the effect is nevertheless present. Say, the brake pedal is not depressed, yet the stoplights are on (see Bell et al. 2007). Such analysis of the actual states of triggers and effects allows one to focus on the most likely causes of failure (Bell et al. 2007). Say, if the pedal is depressed and the lights fail to ignite, first likely causes to investigate may be whether the electrical circuit in the lights are broken or the ‘on/off’ connection between the brake and electrical circuitry (connected to the lamp) is damaged. On the other hand, if the pedal is not depressed and the lights are lit, a first likely cause to investigate may be whether the ‘on/off’ connection between the brake and the electrical circuitry is damaged.

In such assessments of which malfunctioning component(s) resulted in the absence of a given capacity (the capacity of the stoplights to emit light), functional descriptions pick out only the difference making factors with respect to the occurrence of component(s) malfunction.[14] Full, mechanistic, specification of, say, all the details of the electric circuitry and/or the pedal lever mechanism(s) is unwanted.

Thus here we have a case in which function ascriptions and malfunction claims have clear explanatory relevance. They are useful labels to highlight relevant capacities or behaviors, and ignore less relevant and irrelevant ones: function ascriptions black-box most information on the behaviors, components and organization of artifacts, which allows analysts to focus only on those features that make a difference—malfunctioning components—with respect to explaining why an artifact does not manifest an expected capacity. Functional explanations are here explanatorily superior to full—blown mechanistic explanations, since they

[14] To support more detailed malfunction analyses, functions are often decomposed into sub functions in FIL. We here focus on the simple case. It suffices to illustrate our case without introducing unnecessary complexity.

enable isolating relevant difference making properties and suppress reference to irrelevant mechanistic details that only obscure the factors that matter.

Note that this example contrasts with our first case where a functional explanation couched in terms of the ICE theory leaves out information that is relevant (see explanation scheme 6), and additional details should be included to arrive at a satisfactory explanation (see the complete explanation scheme 7, which does not include function ascriptions).

Trigger-effect descriptions refer to desired effects of behaviors (cf. van Eck 2011a). In the above brake system example, the effect of the ignition of the lamps is described yet not the behaviors, say, electricity conversions into light and heat, underlying this effect. In light of the above analysis, we can also understand why in FIL descriptions of functions refer to desired effects of behaviors, rather than purposes or behaviors, and why this is the best choice. Function descriptions are used to black-box or suppress reference to unwanted behavioral and structural details, and to highlight the relevant difference making properties with respect to malfunctioning artifacts. Given this explanatory objective, more elaborate behavior function descriptions are ill-suited since these include irrelevant details such as, say, the thermal energy generated when lamps are lit and/or the input electricity required to make the lamps ignite. Purpose function descriptions refer to desired states of affairs in the world. Such descriptions provide a useful yardstick to assess whether such states of affairs indeed obtain. If this is not the case, say, the purpose assigned to a car's stop light of informing fellow car drivers in the vicinity that the car is slowing down is not achieved, this indicates that an artifact malfunctions. Effect function descriptions are then invoked to explain such malfunctions: these highlight the relevant cause(s) due to which a given purpose fails to be achieved, say, a broken electrical circuit in the lights. Put differently, purpose function descriptions are useful to specify the phenomenon to be explained—a malfunctioning artifact—whereas effect function descriptions are useful to explain that phenomenon/malfunction.

5 Reverse Engineering Explanation and Redesign

So far, we have focused on the relevance of function ascriptions in explanatory contexts. In this section, we consider our third and fourth case of reverse engineering and redesign, respectively. Reverse engineering explanation and redesign are intertwined. Often, redesign phases are preceded by a reverse engineering phase (Otto and Wood 1998, 2001; Stone and Wood 2000) in which the following explanation-seeking question is addressed:

iii How does artifact x realize its capacity to ϕ?

We address the relevance of 'function' with respect to this question first and then assess the utility of function ascription in redesigning.

Reverse engineering explanation is 'prototypical' mechanistic explanation: determining the organized components and sub functions (behaviors) of artifacts by which artifacts produce their overall (behavior) functions. In other words, providing an answer to the question: how does artifact x realize its capacity to ϕ?

In engineering science, reverse engineering and engineering design go hand in glove (e.g., Otto and Wood 1998, 2001; Stone and Wood 2000). In Otto and Wood's (1998, 2001) method, a reverse engineering phase in which reverse engineering explanations are developed for existing artifacts, precedes and drives a subsequent redesign phase of those artifacts.

In the reverse engineering phase, an artifact is first broken down component-by-component, and hypotheses are formulated concerning the functions of those components. In this method, functions are behavior functions and represented by conversions of flows of materials, energy, and signals. After this analysis, a different reverse engineering analysis commences in which components are removed, one at a time, and the effects are assessed of removing single components on the overall functioning of the artifact. Such single component removals are used to detail the functions of the (removed) components further. The idea behind this latter analysis is to compare the results from the first and second reverse engineering analysis in order to gain potentially more nuanced understanding of the functions of the components of the (reverse engineered) artifact. Using these two reverse engineering analyses, a behavior functional decomposition of the artifact is then constructed in which the behavior functions of the components are specified and interconnected by their input and output flows of materials, energy, and signals (Otto and Wood 1998, 2001). An example of a behavior functional decomposition of a reverse engineered electric screwdriver is given in Fig. 1.

After the reverse engineering of a technical artifact, aimed at providing detailed understanding of the mechanism(s) by which it operates, the redesign phase starts by identifying components that *function sub-optimally*, and, thereby, cause artifacts to manifest their overall functions in sub-optimal fashion. Redesign efforts are subsequently directed towards designs with improved functionality of these components (Otto and Wood 1998, 2001). Otto and wood (1998) discuss an example of redesigning an electric wok. The (reverse engineered) artifact's desired capacity to "deliver a uniform temperature distribution across the bowl" failed to be achieved due to the fact that the electric heating elements of the wok, such as a bimetallic temperature controller, were housed in too narrow a circular channel (Otto and Wood 1998, p. 235). Redesign efforts were subsequently directed towards a design with improved functionality of the heating elements, inter alia resulting in a design with a thicker bowl and different shape than in the reverse engineered electric wok.[15]

What is the utility of function ascription in the first reverse engineering phase, and in the subsequent redesign phase? We address these issues in turn, starting with reverse engineering.

5.1 Functions in Reverse Engineering

With respect to answering the question how a complex system-mechanism, here a technical artifact x, manifest its capacity to ϕ, the literature seems to converge on

[15] This redesign step involves a lot of mathematical modeling, use of physical and technological principles, and/or prototype building (Otto and Wood 1998, 2001). These details need not concern us here.

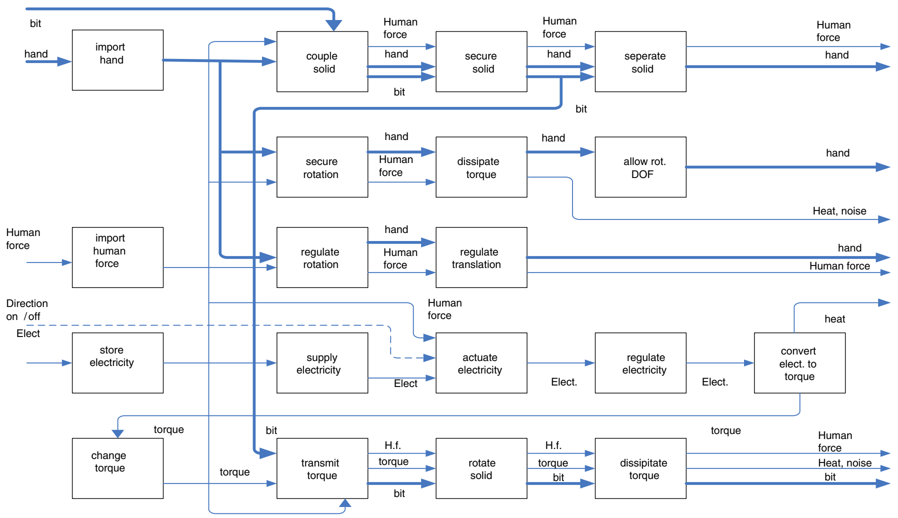

Fig. 1 Behavior functional decomposition of an electric power screwdriver (adapted from Stone and Wood 2000)

the perspective that mechanistic explanations are the optimal choice (Machamer et al. 2000; Glennan 2005; Bechtel and Abrahamson 2005; Craver 2007). These explanations provide the most insightful story on relevant difference making properties with respect to the phenomenon to be explained. Like in malfunction explanation, the desideratum is that all the explanatorily relevant factors with respect to whether or not the phenomenon occurs should be described (Couch 2011). Yet, in the context of this request for explanation, the factors that make a difference are many more than the ones in malfunction explanation. Here, also those factors, organized behaviors and components, that both normal functioning and malfunctioning technical artifacts have in common are to be described in order to understand how a technical artifact x realizes its capacity to ϕ.

Now, descriptions of mechanisms do not include function ascriptions but, rather, are descriptions of organized behaviors and components (Machamer et al. 2000; Glennan 2005; Bechtel and Abrahamson 2005; Craver 2007).[16] Yet, function ascription is emphasized in the construction of mechanistic explanations (Machamer et al. 2000; Craver 2001; Bechtel and Abrahamson 2005; McKay Illari and Williamson 2010) for *individuating* the organized behaviors of components that contribute to the phenomenon to be explained. Those components' behaviors that are discovered to make a functional *contribution* to an overall phenomenon, relative to the mechanistic organization in which they are situated, are described in a

[16] The precise lingo for describing mechanisms differs. Some, for instance, prefer activity and entity talk (e.g., Machamer et al. 2000), others operation and working part terminology (e.g., Bechtel and Abrahamson 2005), or behavior and part parlance (Glennan 2005). These differences have no bearing on the argumentation presented in this paper.

mechanistic explanation.[17] Thus, although function ascriptions do not figure in mechanistic explanations themselves, they are explanatorily relevant with respect to discovering how and which components' behaviors *contribute* to the phenomenon to be explained. Hence, and concurring with this perspective, function ascription is also relevant in the context of mechanistic explanation and reverse engineering explanation. For instance, knowing that and how the behavior "convert electricity to radiation" (Otto and Wood 1998, p. 230) of an electric wok's heating coil contributes to temperature distribution across the wok bowl, deepens our understanding of the mechanism(s) underlying temperature distribution.[18]

5.2 Functions in Redesign

The utility of function ascriptions is also evident in the final context we consider: artifact redesign. Here, we argue, function ascriptions are relevant to *compare* reverse engineered artifacts with novel designs with respect to functional performance. Somewhat more formally, functional parlance is useful in redesign for making the claim that: a novel component *b functions better* than component *a*, or: if component *a* is replaced by component *b*, the function which *a* has will be fulfilled in a better way by *b*. In the wok example, the thicker and differently shaped bowl fulfills its function better than did the old bowl. Or a halogen heat lamp fulfills the function of "converting electricity to radiation" in a better way than the wok's heating coil (cf. Otto and Wood 1998, p. 236).

Since such claims are *normative* ones about improved performance of a given component in comparison to other ones, behavior talk in itself is not enough. Both the old heating coil and the new halogen heat lamp display the behavior of converting electricity to radiation, and both the old and new bowl conduct heat. To compare and distinguish these components in terms of performance, terminology is needed to highlight differences in performance; statements like component *b functions better* than *a* do express such facts. Hence, functional parlance is relevant in this context.

Let us, finally, briefly comment on the relevance of the notion of behavior function that is employed by Otto and Wood (1998, 2001) in their reverse engineering and redesign method. Given that redesign starts with a reverse engineering phase, the choice to employ behavior functions is the optimal one. Specifying organized components and behaviors gives the most detailed information on the mechanisms by which artifacts operate. Effect function descriptions omit relevant details and purpose function descriptions are irrelevant for deciphering the

[17] Functional contribution is understood in terms of the notion of "mechanistic role function" (Craver 2001), being an offshoot of Cummins' (1975) notion of function. On the mechanistic account of role functions, function ascription is intimately tied to the manner in which the behavior of a component is organized within a mechanism. In Cummins' account, organization is treated more loosely and not restricted to mechanisms (cf. Craver 2001).

[18] In Otto and Wood's (1998, 2001) reverse engineering method, the terms 'function' and 'functional decomposition', referring to (sets of organized) behavior(s), are used to describe and explain the workings of artifacts. In the mechanist literature, the term 'function' is reserved for individuating mechanisms, whereas descriptions of mechanisms are given in term of organized behaviors and components.

internal workings of artifacts, since they describe state of affairs in the world to be realized by the behaviors of artifacts. Also for the subsequent comparative analyses, behavior function descriptions are the most useful for at least two reasons. Firstly, these are the most detailed and hence provide the most information to assess the performance of components and make comparisons between components. Say, the novel halogen heat lamp fulfills the function of ‘converting electricity to radiation’ in a better way than the wok’s heating coil since the halogen lamp produces less heat or noise, or both. Secondly, in replacing components one needs to take the structural configuration of the reverse engineered artifact into account, i.e., how the to-be replaced component is organized with other components, in order to ensure that the novel component can indeed be placed in this configuration. Descriptions of behavior functions, and sequences thereof as specified in reverse engineering models in which the behavior functions of the components are specified and interconnected by their input and output flows of materials, energy, and signals, provide the most elaborate information on structural configurations.

6 Explanatory Force of Functional and Mechanistic Explanations

So far, we have focused our analysis on the technical domain. We close the paper by discussing a ramification of our analysis for the philosophy of mechanistic explanations in general, i.e., assessments of the explanatory force of mechanistic vis-à-vis functional explanations, broadening our focus to the biological domain.

Assessments of the explanatory power of mechanistic explanations vis-à-vis functional ones basically hinge on the distinction between, one the one hand, sketches of mechanisms or functional explanations and complete models of mechanisms, on the other. Whereas mechanism sketches/functional explanations specify some, yet not all, entities and/or activities and/or organizational features relevant for explaining explananda phenomena, complete models specify all the features of mechanisms considered relevant for explaining explananda phenomena (Craver 2007; Kaplan and Craver 2011; Piccinini and Craver 2011; Gervais and Weber 2013).

Craver (2007) stipulates one core requirement or “central criterion of adequacy” (p. 139) that mechanistic explanations ought to meet: mechanistic models should fully account for their explananda phenomena. To ‘fully account for’ or ‘render intelligible’ is equated with providing (ideally) complete descriptions of the mechanisms underlying phenomena. The more complete, the more adequate an explanation is judged to be (cf. Machamer et al. 2000; Kaplan and Craver 2011; Piccinini and Craver 2011). Completeness is achieved when models specify all the entities, activities, and organizational features that are relevant to explananda phenomena. Mechanists who adopt this ‘completeness’ perspective, always come down on the side of ‘complete’ mechanistic explanations, and always take functional explanations to have less explanatory power. Our analysis of contrastive malfunction explanations significantly corrects this perspective.

The ‘completeness’ perspective, indeed, seems exemplified by reverse engineering explanations: elaborate behavior functional decompositions are constructed to describe and explain the mechanisms of artifacts. But what about contrastive malfunction explanations? In that case, as we saw in section four, functional explanations or mechanisms sketches, which leave out a lot of mechanistic details, are explanatorily superior to full-blown mechanistic ones. By black-boxing most information on the behaviors, components and organization of artifacts, functional explanations allow analysts to focus only on those features that make a difference—malfunctioning components—with respect to explaining why an artifact does not manifest an expected capacity. In the case of malfunction explanation, less is more when it comes to adequate explanations. Functional explanations are here thus explanatorily superior to full—blown mechanistic explanations, since they isolate relevant difference making properties and suppress reference to irrelevant mechanistic details that only obscure the factors that matter.

The point generalizes. Also in the case of biological malfunction, explanations that omit reference to mechanistic details are more adequate than ones in which increasing amounts of mechanistic details are specified. Consider, for instance, impaired blood circulation in the circulatory system. Malfunction explanations here better single out only those steps—entities engaging in activities—in the circulatory system's mechanism(s) that cause the circulation of blood to be impaired, i.e., make a difference to whether or not impaired blood circulation occurs. Of all the activities of entities that one can enumerate in the circulatory system's mechanism for blood circulation (e.g., spontaneous electric variation in specific myocardium, coordinated contraction of the cardiac muscle, ejection of blood from the ventricles into the aorta and the arterial system, etc.) only some are relevant to explain *impaired* blood circulation. Say, blood transport by (normally undamaged) vessels.[19] In the case of impaired blood distribution, the cause may be that blood transport is disrupted in particular vessels as a result of thrombosis in those vessels. If so, an explanation for impaired blood distribution by the circulatory system is best given in terms of the claim that some vessels fail to transport blood. Like in the case of malfunction explanation in engineering, the functional explanation or mechanism sketch should only specify those difference making factors that underlie the occurrence of a specific malfunction.

Our analysis thus shows that assessments of the explanatory power of mechanistic versus functional explanations that solely hinge on the distinction between complete models and mechanism sketches are too simple. In some contexts, in casu malfunction explanation, providing the best explanations implies providing functional rather than mechanistic ones.

In sum, what is missing in current discussions that wage the explanatory force of mechanistic versus functional explanations solely in terms of 'completeness', is consideration of the explanation-seeking questions for which explanations are sought: it depends on the explanatory context whether mechanistic or functional explanations are to be preferred.

[19] We adapt this example from (Nervi 2010).

7 Conclusions

Too often in the philosophical literature on functions it has been assumed that function ascriptions in themselves are explanatory and, hence, relevant (e.g., Wright 1973; Millikan 1989; Neander 1991). Wimsatt (1972) and Wouters (2003), however, caution against the idea that function ascriptions, by definition, provide explanations. Whether or not function ascriptions have explanatory leverage is an issue that requires careful analysis. In this paper we assessed the explanatory relevance of ascriptions of technical functions, an issue by and large neglected in the literature on technical artifacts and technical functions. We analyzed the relevance of technical function ascriptions in three explanatory contexts and in a specific engineering redesign context. We argued that whereas function ascriptions serve a mere heuristic role in the context of explaining why artifacts are produced, they play a substantial role in explaining artifact malfunction and capacity explanation, as well as in comparative analyses of functional performance in engineering redesign. These assessments were developed by comparing the explanatory merits of teleological, functional, and mechanistic explanations in these contexts. We closed the paper by discussing a ramification of our analysis for current thinking about the explanatory force of mechanistic vis-à-vis functional explanations, and in doing so broadened the focus to the biological domain.

Acknowledgments We thank two anonymous referees, Conor Dolan, Raoul Gervais, and Merel Lefevere for very useful comments.

References

Bechtel, W., & Abrahamson, A. (2005). Explanation: A mechanist alternative. *Studies in History and Philosophy of Biological and Biomedical Sciences, 36*, 421–441.

Bell, J., Snooke, N., & Price, C. (2007). A language for functional interpretation of model based simulation. *Advanced Engineering Informatics, 21*, 398–409.

Bigelow, J., & Pargetter, R. (1987). Functions. *Journal of Philosophy, 84*, 181–196.

Chakrabarti, A. (1998). Supporting two views of function in mechanical designs. In *Proceedings 15th national conference on artificial intelligence, AAAI'98*, July 26–30, 1998, Madison, WI, USA.

Chakrabarti, A., & Bligh, T. P. (2001). A scheme for functional reasoning in conceptual design. *Design Studies, 22*, 493–517.

Chandrasekaran, B., & Josephson, J. R. (2000). Function in device representation. *Engineering with Computers, 16*, 162–177.

Couch, M. (2011). Mechanisms and constitutive relevance. *Synthese, 183*, 375–388.

Craver, C. F. (2001). Role functions, mechanisms, and hierarchy. *Philosophy of Science, 68*, 53–74.

Craver, C. F. (2007). *Explaining the brain: Mechanisms and the mosaic unity of neuroscience*. New York: Oxford University Press.

Cummins, R. (1975). Functional analysis. *Journal of Philosophy, 72*, 741–765.

De Ridder, J. (2006). Mechanistic artefact explanation. *Studies in History and Philosophy of Science, 37*, 81–96.

De winter, J. (2011). A pragmatic account of mechanistic artifact explanation. *Studies in History and Philosophy of Science, 42*(4), 602–609.

Deng, Y. M. (2002). Function and behavior representation in conceptual mechanical design. *Artificial Intelligence for Engineering Design, Analysis and Manufacturing, 16*, 343–362.

Erden, M. S., Komoto, H., Van Beek, T. J., D'Amelio, V., Echavarria, E., & Tomiyama, T. (2008). A review of function modeling: Approaches and applications. *Artificial Intelligence for Engineering Design, Analysis and Manufacturing, 22*, 147–169.

Gero, J. S. (1990). Design prototypes: A knowledge representation schema for design. *AI Magazine, 11*(4), 26–36.
Gervais, R., & Weber, E. (2013). Plausibility versus richness in mechanistic models. *Philosophical Psychology, 26*(1), 139–152.
Glennan, S. (2005). Modeling mechanisms. *Studies in the History and Philosophy of the Biological and Biomedical Sciences, 36*(2), 375–388.
Hawkins, P. G., & Woollons, D. J. (1998). Failure modes and effects analysis of complex engineering systems using functional models. *Artificial Intelligence in Engineering, 12*(4), 375–397.
Houkes, W. (2006). Knowledge of artifact functions. *Studies in History and Philosophy of Science, 37*, 102–113.
Houkes, W., & Meijers, A. (2006). The ontology of artefacts: The hard problem. *Studies in History and Philosophy of Science, 37*, 118–131.
Houkes, W., & Vermaas, P. E. (2010). *Technical functions: On the use and design of artefacts*. Dordrecht: Springer.
Houkes, W., Kroes, P., Meijers, A., & Vermaas, P. E. (2011). Dual-nature and collectivist frameworks for technical artefacts: A constructive comparison. *Studies in History and Philosophy of Science, 42*, 198–2005.
Kaplan, D., & Craver, C. (2011). The explanatory force of dynamical and mathematical models in neuroscience: A mechanistic perspective. *Philosophy of Science, 78*, 601–627.
Kitamura, Y., Koji, Y., & Mizoguchi, R. (2005). An ontological model of device function: Industrial deployment and lessons learned. *Applied Ontology, 1*, 237–262.
Kroes, P. (2003). Screwdriver philosophy; Searle's analysis of technical functions. *Techné, 6*(3), 22–35.
Krohs, U. (2009). Functions as based on a concept of general design. *Synthese, 166*, 69–89.
Lind, M. (1994). Modeling goals and functions of complex industrial plants. *Applied Artificial Intelligence, 8*, 259–283.
Lipton, P. (1993). Making a difference. *Philosophica, 51*, 39–54.
Machamer, P. K., Darden, L., & Craver, C. F. (2000). Thinking about mechanisms. *Philosophy of Science, 57*, 1–25.
McKay Illari, P., & Williamson, J. (2010). Function and organization: Comparing the mechanisms of protein synthesis and natural selection. *Studies in History and Philosophy of Biological and Biomedical Sciences, 41*, 279–291.
McLaughlin, P. (2001). *What functions explain*. Cambridge: Cambridge University Press.
Millikan, R. (1989). In defense of proper functions. *Philosophy of Science, 56*, 288–302.
Neander, K. (1991). The teleological notion of "function". *Australasian Journal of Philosophy, 69*, 454–468.
Nervi, M. (2010). Mechanism, malfunctions and explanation in medicine. *Biology and Philosophy, 25*, 215–228.
Otto, K. N., & Wood, K. L. (1998). Product evolution: A reverse engineering and redesign methodology. *Research in Engineering Design, 10*, 226–243.
Otto, K. N., & Wood, K. L. (2001). *Product design: Techniques in reverse engineering and new product development*. Upper Saddle River NJ: Prentice Hall.
Pahl, G., & Beitz, W. (1988). *Engineering design: A systematic approach*. Berlin: Springer.
Piccinini, G., & Craver, C. F. (2011). Integrating psychology and neuroscience: Functional analyses as mechanism sketches. *Synthese*, *183*, 283–311.
Preston, B. (1998). Why is a wing like a spoon? A pluralist theory of functions. *Journal of Philosophy, 95*, 215–254.
Price, C. J. (1998). Function-directed electrical design analysis. *Artificial Intelligence in Engineering, 12*(4), 445–456.
Searle, J. (1995). *The construction of social reality*. New Haven: Free Press.
Srinivasan, V., & Chakrabarti, A. (2009) SAPPhIRE: An approach to analysis and synthesis. In *Proceedings of the 17th international conference on engineering design,* Stanford, CA, USA, August 24–27, 2009. Design Society, pp. 2.417–2.428.
Srinivasan, V., Chakrabarti, A., & Lindemann, U. (2012) A framework for describing functions in design. In *Proceedings international design conference—design 2012*, Dubrovnik, Croatia, May 21–24, pp. 1111–1121.
Stone, R. B., & Chakrabarti, A. (2005) Guest editorial. Special Issue: Engineering applications of representations of function, Part 2. *Artificial Intelligence for Engineering Design, Analysis and Manufacturing, 19*(3), 137.

Stone, R. B., & Wood, K. L. (2000). Development of a functional basis for design. *Journal of Mechanical Design, 122*, 359–370.
Strevens, M. (2004). The causal and unification approaches to explanation unified—causally. *Noûs, 38*(1), 154–176.
van Eck, D. (2011a). Supporting design knowledge exchange by converting models of functional decomposition. *Journal of Engineering Design, 22*(11–12), 839–858.
van Eck, D. (2011b). Incommensurability and rationality in engineering design: The case of functional decomposition. *Techné: Research in Philosophy and Technology, 15*(2), 118–136.
Vermaas, P. E. (2006). The physical connection: Engineering function ascriptions to technical artefacts and their components. *Studies in History and Philosophy of Science, 37*, 62–75.
Vermaas, P. E. (2009). The flexible meaning of function in engineering. In *Proceedings of the 17th international conference on engineering design (ICED 09)*, pp. 2.113–2.124.
Vermaas, P. E., & Houkes, W. (2003). Ascribing functions to technical artefacts: A challenge to etiological accounts of functions. *British Journal for the Philosophy of Science, 54*, 261–289.
Weisberg, M. (2007). Three kinds of idealization. *The journal of Philosophy, 104*(12), 639–659.
Wimsatt, W. C. (1972). Teleology and the logical structure of function statements. *Studies in History and Philosophy of Science, 3*, 1–80.
Woodward, J. (2003). *Making things happen*. Oxford: Oxford University Press.
Wouters, A. G. (2003). Four notions of biological function. *Studies in History and Philosophy of Biology and Biomedical Science, 34*, 633–668.
Wright, L. (1973). Functions. *Philosophical Review, 82*, 139–168.