Material Culture and Technological Determinism

Hector MacIntyre

Thesis submitted to the
Faculty of Graduate and Postdoctoral Studies
in partial fulfillment of the requirements
for the Doctor of Philosophy in Philosophy

Department of Philosophy
Faculty of Arts
University of Ottawa

Abstract

This dissertation has two results. First, I argue that each of the two basic components of technological determinism (TD)—what I call the *inexorability thesis* and the *autonomy thesis*—are plausible claims on a naturalistic stance. Second, I argue that a normative model for the design of cognitive systems can guide the practice of cognitive engineering, e.g. the task of building cognitive aids and enhancements.

TD conjoins two logically independent but empirically related claims. The inexorability thesis is the claim that technology change is an evolutionary process. I defend this claim against considerations raised by Lewens, most notably the lack of a robust account of artifact reproduction that would underwrite genuine transmission. I consider (but reject) the solution of memeticists to this problem. I find that theorists of cultural evolution, e.g. Boyd and Richerson (among others), do present a plausible response. Technologies can be said to evolve via the cumulative selective process of cultural retention.

The autonomy thesis is the claim that features of human cognitive agency arise from material culture. I argue for this thesis through a consideration of the merits of Preston's theory of material culture. Her sociogeneric approach attributes human cognitive agency to a material cultural genesis, and this approach is backed by strong anthropological evidence. Preston would not accept the thesis but she does not manage to exclude it, despite an admirable attempt to develop an account of innovation. I also consider the design of technologies in the practice of cognitive engineering and propose adopting a normative theory of factitious intellectual virtue as a model to guide design in this arena.

Table of Contents

Acknowledgements

Let me first thank my supervisor Andrew Sneddon for all his feedback, guidance, and support during the research and writing, as well as on career matters. Thanks also to the members of my examination committee for their diligence and expertise. Philosophical research is something of an indulgence, and I have my family to thank for indulging me their support and encouragement. This is for my parents Joan and Norman, my Grandma Doris, brother Tyler, sister Teaghan, the Olivers, and most importantly our late Grandpa Gordon Oliver. And Tuffy.

I have many friends and colleagues to thank as well. I would not have come to Ottawa if not for Stephen Latta. The first year rooming with Jennifer Whitson was a time of warmth and fun. I have fond memories of life at the Booth Street pad with Andrew Clark and Shawn Moi, with frequent and stimulating visits from Aaron O'Brien, Matthew Anderson, Dave Cannon, and Sean Anderson. Special thanks to Michelle Hawkins and Lucas Jurkovic for many fun conversations over the years, to Marie-Hélène Villeneuve for hosting me when I visited. Early on I was an upstart padawan waiting to be set straight by Leah Armontrout-Spencer. I gained an inspiring colleague in Michelle Ciurria, and I also have and a great friend and colleague in Mary Butterfield. Additional thanks to Jennifer Chandler and Ian Kerr for formative discussions and seminars in the early stages and, last but not least, to the staff of the Department of Philosophy, Jocelyne Lacasse and Catherine Bernard, without whom I would not have finished. You are all very wonderful.

Introduction

A footpath has a peculiar effect on people. It invites them to walk along a certain route, e.g. through some brush or down an embankment, largely owing to the convenience of there being a route that is smooth, unobstructed, and visible. It's an attractive and reliably efficient way to move between two places. It has the benefit of being quite durable so long as it is frequently used (though this can depend on the material in which it's trodden). It requires no construction to maintain apart from regular use. Often there is no design or plan that produces a footpath. It simply forms where people walk.

The peculiar thing about a footpath, though, is that once one appears, it tends to stay around because people tend to use it. That is, once made, it seems to get maintained just by virtue of having been there in the first place. Now of course we might want to say that strictly speaking, this is not quite so. Pedestrians are the ones who maintain a footpath by walking over the same ground again and again. But as a causal explanation of how the footpath gets maintained, this is purely mechanical. Saying that people are the ones who maintain the footpath suggests that this is an intentional action of theirs, but this does not hold up to scrutiny. People who use a footpath almost never intend by using it also at the same time to maintain the path—all they intend is to go from one spot to another. The footpath is largely a byproduct of their intention to make that trip.

What explains this tendency of pedestrians to keep using the footpath? We might presume that it is simply the fact that people want to move easily between destinations, and a footpath (once produced) satisfies that desire. It first comes about as a byproduct of people just forging ahead between two places and, once trodden, it tends to encourage further use in a peculiar sort of feedback effect. Walkers don't intend to maintain the path in using it, they just intend to move between two places, yet maintaining the footpath is itself another byproduct, just as the initial series of troddings was. However, the byproduct is not just a one-off side effect. It has a peculiar self-reinforcing tendency. Once it gets started it seems to self-promote.

The contention of this dissertation is that this odd feedback dynamic, which is not guided by the intentions of human agents yet still piggybacks on their activity, is a feature of many technologies. Almost all modern examples follow this pattern. Telecommunication technologies are perhaps the most prominent example. A century ago, the prevailing methods of communication were personal face-to-face conversation and post. Compared to cellular telephones, electronic mail transmission, and video communication, the old technologies are grossly inefficient. The increased efficiency of telecommunications has many secondary effects that have transformed social relations, but the key one is that some of *these* effects tend to self-reinforce. People who adopt smart phones encourage others to utilize them as well, which makes it more likely that the initial adopters will continue using theirs. It's not helpful to view this sort of feedback effect as a side effect because of the transformative scope of its self-reinforcing dynamic.

At Cambridge, Huw Price has recently founded a project to study the "existential risk" our own technologies pose to human society, while Nick Bostrom directs the Future of Humanity Institute at Oxford on a similar path of research. The "dual nature" of technology is an old saw whereby any tool can be used for good or evil, depending on the intention of the agent who wields it. But a surprising number of intellectuals seem to suspect that modern technologies in particular seriously jeopardize long-term human safety even as it is argued that these technologies are what provide the measure of safety we currently enjoy. As Perlman (2004) notes in an excellent survey, despite all the attention that artifacts and their functions have received in the philosophy of biology and naturalistic literature in recent decades, hardly any effort has been made to connect these discussions with those on technology from social theory and other traditions.[1] Perlman grants that there exists a wide gulf between these philosophical styles. Much of the philosophy of technology is obscure and speculative, while naturalistic studies of artifacts stemming from function theory and action theory tend to be quite narrow.

The past two decades, however, have seen many works by philosophers of biology, of action, of anthropology, and of cognitive psychology that make it possible to piece together some naturalistically coherent insights about technologies and their impacts on individuals, on groups, and on social organization. Tackling topics that bridge the divide between Amglo-American and European traditions has never been an easy task, and in this effort I have been greatly helped by recent exemplars of empirically minded approaches such as Margolis and Laurence's anthology *Creations of the Mind* (2007) and Beth Preston's excellent (2013) monograph, which each take broad views informed by the best of all available viewpoints. In the rest of this introduction I want to spell out the kind of view I'm aiming to assess—technological determinism—as well as some of the new work on which I draw to assess it. Then I'll sketch the argument.

§1.1 Technological Determinism

The view that technology change is a leading cause of social change is known as technological determinism (TD). TD has a long sordid history in social and political the-

[1] Perlman (2004), 45.

ory, and few defenders. I say TD has a sordid background because many theorists have at different times and in different senses been regarded as proponents of TD. Marx's account of social change is deterministic in the sense that it posits an inexorable sequence of phases and a material basis, and was for this reason seen as an example of TD for much of the last century. Jacques Ellul is another author whose work has been cited as an example of TD. However, I suspect that once the commitments of the view are made plain, it will be doubtful whether anyone has ever actually held it in a pure form.

According to Bimber (1990; 1994), there are two central components of TD and three candidate readings from which to draw its most viable interpretation. He proposes the rather easy stance that any statement of TD must be (1) deterministic and (2) technologically so. Getting clear on the doctrinal content of TD is a task primarily concerned with the scope of these two component claims. How should one relate the first claim to e.g. traditional hard determinism, given that its scope appears to be squarely within the domain of the social sciences with their different approaches to causality? What do we understand by "technology" given the nebulous definitions available for this term?

Take claim (1): The philosophical problem of hard determinism is almost certainly an illusion. Not only do the fundamental sciences now tell us that causality is probabilistic, but even if they told us that every causal relation was realized in a fully deterministic manner, it still would not follow that the sort of agency required for e.g. moral responsibility was incompatible with this fact. This is because the conception of agency at work has never been based on a causally unhinged freedom of the will. This has long been affirmed by the strong tradition of compatibilism in philosophy and by a variety of writers from Strawson to Frankfurt and Dennett. I take it that TD raises concerns about agency and not about philosophical determinism; to address the staple question of determinism in this project would lead us far astray from the philosophically untrodden ground that makes TD an interesting topic. Questions about the nature and possibility of agency are largely independent from questions about causal freedom. Furthermore, much has yet to be settled by philosophers of social science concerning how to study the effects of their phenomena of interest. This project proceeds against an overtly compatibilist backdrop with an aim to make more plain how individuals participate in their material culture, and how it reciprocally affects their lives.

Component (2) of TD is that technology "should play a necessary part in the way that preceding events or states of the world determine the future…the laws of nature determining human history do so through technology. Technology is the medium through which general laws, some of which are learned through the sciences, shape the course of human events."[2] Any determinism worthy of the qualifier "technological" ought to rely on specific features of technologies to explain these determining conditions or laws.[3]

Bimber argues that there are no hard or soft (strong or weak) versions of TD to be profitably distinguished because the logical standard for using the view to adjudicate explanations of social and historical change requires that it have a firm meaning. Other authors contend the opposite along the following lines: a strong version has it that a set of technologies perpetuates a corresponding set of social relations that in turn maintain it, while the weak version says that a set of technologies is compatible with a corresponding set of social relations even though it does not tend to perpetuate them.[4] Langdon Winner's famous example of the stronger version is an analysis of nuclear power indicating

[2] Bimber (1994), 87.
[3] Bimber (1990), 340.
[4] Kroes and Verbeek (2013), 17.

that these systems both require and propagate authoritarian structures of political power. I tend to agree with Bimber that the weaker version is virtually meaningless as a tool for explaining (let alone predicting) social and historical change. It suggests no criteria for ascertaining compatibility between technologies and social organization, nor does it say why this compatibility would be significant if it could be ascertained.

Importantly, Bimber claims that "technology" cannot be defined too broadly. He argues that the definition ought to be restricted to artifacts or machines. "Including factors broader than artifacts or machines in the definition of technology, and hence in the explanation of the determining factors for human history, requires the conclusion that social change is dependent in part upon social factors."[5] This cannot meet the standard of TD because it attributes (even partial) agency to society. "For these reasons I assume that, to make sense, technological determinism must define technology as physical artifact or machine and the associated material elements by which these are produced."[6] There are many clear examples of artifacts that have political effects and instantiate power dynamics between groups and individuals. The accessibility of walkways and buildings is one example, like concrete versions of financial impediments that prevent the poor from attending ivy league colleges. Among Winner's favourite examples of politically active technologies are industrial tomato harvesting machines which not only produced crops of fruit that differed from hand-harvested tomatoes in both taste and plumpness, but threatened to put traditional family tomato farms out of business altogether.[7]

Let the following stand as a provisional definition of TD:

---

[5] Bimber (1994), 88.
[6] Bimber (1994), 88.
[7] Winner (1986).

> [TD] can be seen as the view that, in light of the past (and current) state of technological development and the laws of nature, there is only one possible future course of social change. This might mean that various technological processes, once begun, require forms of organization or commitments of political resources, regardless of their social desirability or of previous social practices. It could mean that an enterprise (for example, the railroad) necessitates subsequent technologies (such as the telegraph, or large-scale hierarchically organized steel-production facilities) and requires a pool of labour, the availability of capital, an insurance and banking industry, and so on, so that a fixed and predictable course of economic, social, and cultural change follows inevitably from the adoption of the railroad.[8]

Bimber calls this formulation the nomological reading of TD. It is one of three basic formulations he examines, and it is the one that, according to him, comes closest to a pure statement that can be given to specify the substance of TD.

The version of TD I defend conjoins a weaker version of TD's deterministic component, in the form of a claim about technology change as an evolutionary process which I call the *inexorability thesis*, with an *autonomy thesis* that develops the "technological" component of TD in terms of a dynamic feedback effect found in material culture. The inexorability thesis is philosophically defensible along basically the same lines as Bimber would hope. But the autonomy thesis could be seen as verging on a normative account. I think it need not be seen this way once more basic facts about agency are made plain.

§1.1.1 Three Readings of Technological Determinism

Besides the nomological account, Bimber also considers "normative" accounts and "unintended consequences" accounts of TD. Only the nomological reading is found to pass muster as both deterministic and technological in the required senses. The other two readings each fail to satisfy the basic criteria.

---

[8] Bimber (1994), 83-84.

Normative accounts of TD fail because they “do not impute special agency to the characteristics of technology as artefact, nor do they demonstrate how the disjunction of norms is to be permanent, predetermined, or beyond human control.”[9] Instead these accounts rely on cultural mechanisms ultimately controlled by human actions. The ways in which social change derives from technology change on the normative account are determined by intentional actions. It is the adoption of or adherence to particular norms emphasizing maximal efficiency, or mechanistic logic, that generates such close affinity between social change and technology change. But the norms are social or behavioural in origin, not inherently technological. A society could in principle abandon the norms if everyone agreed to do so. Hence normative accounts fail to meet the second criterion Bimber gave. For this reason, such common social science tropes as self-fulfilling prophecies, etc., would also fail to meet the “technology” criterion of a proper TD theory, as would the infamous argument of Jacques Ellul (1967) concerning technique.

Unintended consequences accounts also fail, but by the first criterion. These accounts argue that it is the unanticipated effects of design—the side effects of technology—which compound to fix human social change. Such accounts fail to meet the deterministic criterion in that they simply result from the human failure to foresee the impacts of intentionally undertaken human actions. “Unintended consequences are basic facets of social action, rather than the special products of technology. Unintended Consequences Accounts do not justify attributing the unpredictability of social outcomes to features of technology.”[10] Hence these accounts too are rejected.

[9] Bimber (1990), 341.
[10] Bimber (1990), 341.

This leaves only nomological accounts. Bimber finds that these accounts do in fact meet the criteria for TD. They do so by "holding that society evolves along a fixed and predetermined path, regardless of human intervention. That path is itself given by the incremental logic of technology."[11] Technology drives social change. At the heart of this picture of TD are two distinct claims. Bimber himself characterizes them in slightly different ways between his two papers. In 1990 he describes the two basic claims of nomological TD in this way: "human history develops over a course which corresponds to developments in technology; and developments in technology themselves proceed strictly according to an internal sequence and logic deriving from physical laws."[12] And in 1994 he describes them in this way: "technological developments occur according to some naturally given logic, which is not culturally or socially determined, and…these developments force social adaptation and changes."[13] (Winner's notion of reverse adaptation, despite being a normative account, articulates a similar idea.[14]) Mitcham characterizes the nomological idea of TD as "path dependency": once undertaken, a particular path of innovation must be followed, as in the progression from steam engines to combustion engines, computers to smartphones, or from electricity to toasters.[15]

However, Bimber's formulation of the nomological reading is so strong that it appears to cast doubt on his claim that the standard for TD "does not admit of 'hard' and 'soft' versions of determinism."[16] Can it really be that TD says "there is only one possible future course of social change" set by innovation and that it occurs "inevitably" once par-

[11] Bimber (1990), 341.
[12] Bimber (1990), 342.
[13] Bimber (1994), 84.
[14] Winner (1977), 226.
[15] Mitcham (2013), 18.
[16] Bimber (1994), 87.

ticular technologies arise? Surely this must be a straw view. But even if no actual writer has ever held such a strong position, it may still be useful to articulate and use it to clarify what claims weaker versions are actually staking. Let us call the industrial-strength version of nomological TD the "inevitability version" of the view. It is implausibly strong, and so if Bimber insists that it is the only viably pure form of TD on offer then there is little hope that TD can be defended. However, my strategy will be to interpret the nomological reading of TD according to two component theses: an *inexorability thesis* and an *autonomy thesis*. I argue that an "inexorability version" of the nomological view is defensible even if it does not meet the burden of Bimber's version. I also argue that a version of the autonomy component which focuses on the phenomenon of agency—and hence ventures into normative territory—is also defensible.

These two theses correspond roughly to the two major domains studied by philosophers writing about technology, artifacts, and material culture, namely function theory and action theory. I will argue for each thesis on its own. Though they are not unrelated, an argument for TD can (by basic rules of inference) simply assert a conjunction of the two claims once each has been defended on its own. The rest of this introduction briefly explains the content of these two theses, their relations to TD and to each other, and indicates the phenomena of interest.

§1.1.2 The Inexorability Thesis

Let the following statement stand as TD's *inexorability thesis*: the incremental sequence by which technology change occurs is an evolutionary process. This is a weaker claim than what Bimber insists is required. Recall that he characterized the nomological account in terms of an "inevitable" progression of technology change. But Bimber does not say why TD requires such a strong criterion of inevitability, nor does he suggest a mechanism for this progression. Rather than saying that technological development takes only one possible course, while neglecting to supply a mechanism, the inexorability thesis does give a mechanism—natural selection—and says that technological development may take different courses (e.g., in different cultures and societies) that are not guided by human intention or direction. Technology change occurs in a way that human beings do not determine, and if the right conditions are present it will continue to do so inexorably. The choice of label here—"inexorable" as opposed to "inevitable"—reflects an attempt to capture the basic idea of technology moving along in a way that is beyond the control of the human agents who nonetheless manufacture it.

There are two benefits to formulating in this way the aspect of TD which Bimber and others try to capture by such terms as "logic" and "inevitable development". First, it sharpens the somewhat rough idea about how technological development proceeds in relation to agent intention that seems to drive much of the empirically misbegotten social-theoretic writing on this topic. Certainly the idea that technology evolves is neither novel nor particularly insightful, but it is surprising how often it is asserted on the basis of nothing more than casual investigation. It also provides a good test case for revealing what sorts of limitations can be exposed in the philosophy of evolutionary theory.

Second, a formulation like this helps to situate the study of technology in relation to other social scientific investigations. The theory of evolution by natural selection is extraordinarily powerful and many attempts have been made to invoke it in domains other than biology, with limited success. It remains an attractive option because the social sciences are more like biology than they are like chemistry, physics, or cosmology. The more phenomena we can scrutinize in the light of this theory, the better. It is sometimes thought that the study of technology by its very orientation is closer to fundamental science because technology is an output of engineering, which is rooted in the application of physical and mathematical design. But the application is merely a symptom, and not all technology creation is rooted in engineering. Other species also create technology yet they have no engineering prowess and arguably no genuine intentions at all. To the extent that we can investigate the impact of technologies on society in terms of processes that are well understood in other domains, namely evolutionary biology, we can begin to grapple with very complex questions in a more tractable way.

Here I am expanding on Bimber's (1994) definition of determinism as the claim that "history is determined by laws or by physical and *biological* conditions rather than by human will" (my emphasis). Putting this in terms of inexorability is an attempt to understand his claim that technology follows an inevitable course of progression as a claim that it evolves. Understood my way, any progression is merely figurative. The forces at work in evolution are both stupid and blind. But an evolutionary account of technology change would still offer a sound basis on which to show that such change exhibits an inexorable sequence that, while perhaps instantiated in and realized by the guidance of human designers, is not reliant on this guidance for the specific route it takes.

On my suggestion, the inexorability of technology change is characterized as an algorithmic process. It is inexorable because it follows a sequential, stepwise path not unlike an effective procedure. Natural selection is often said to work this way: elements in an environment combine to form organic compounds such as proteins and acids, which eventually produced replicating systems, which in turn acquired increasing complexity as mutations and replications occurred (or not) in relation to new environments. Given the presence of three basic conditions—variation, reproduction and heredity—a process of natural selection will only fail to occur in the event of environmental calamity. This is not a progression, but the gradual accumulation of features which increase in both the complexity and functional usefulness of the traits that are retained, can resemble one.[17] It makes sense to use evolutionary theory to defend an inexorability reading of TD.

I presume that prima facie, the idea that technology evolves by a process of natural selection is actually quite plausible. Many writers have thought that even if the presence of intentional selection (e.g. design in engineering and manufacturing) makes it difficult for the three basic conditions of selection to obtain, since reproduction is occurring only at the whim of intelligent agents, it nevertheless occurs when foresight fails. Virtually the entire computing industry, for instance, has expressed belief in "Moore's Law" as a reliable forecast of technological progress. This is significant given that many of those who endorse this forecast are trained scientists and engineers. But the very reliability of the development of computer technology should indicate that it is too stable to be an evolutionary process, even if no person, group, or arrangement thereof is driving the observed pattern. Hofstadter explains the point with his typical clarity: "Moore's three-dimensional law is a remarkable 'epiphenomenon' of our global culture, by which I mean that it is a statistical regularity that emerges from a swarm of unknown, mutually independent activities…the unpredictable microevents are the hundreds of small scientific and engineering breakthroughs scattered over time and around the planet."[18]

[17] Cf. Dennett (1995), chapter two, §4 and §5.
[18] Hofstadter (2004), 166.

Nevertheless, the theory of evolution is powerful and successful enough that one must carefully consider the possibility that selection is indeed at work in any system where the three basic conditions are found. If the conditions are present, there is more reason to presume that selection would occur than that it would not. And it does seem that, computational linearity aside, the vast and rather unpredictable accumulation of innovation in large-scale technological development in the past century alone points to an uncoordinated process of selective retention. The mutually independent activities and unpredictable microevents keep generating innovations, and these keep generating subsequent innovations, and each daughter generation of technology retains more and more useful traits while shedding less useful ones. Nobody seems to be directing this process and, what's worse, nobody seems able to alter its course in any impactful way.

Of course, under closer scrutiny with the aid of evolutionary biology, this sense begins to evaporate. While there is clearly a great deal of variation in technologies, the major problem is giving a clear picture as to how technology can reproduce apart from the direct actions of intelligent human agents. Even automated technologies the sole function of which is to produce other machines, as in e.g. automated factories and assembly lines, or sophisticated software systems, are ultimately set in motion according to the directives and designs of human engineers and their clients. In the entire domain of technology there appears to be only production and never any genuine reproduction. Hence there can be no truly automatic copying of the sort required for selection to act, since all that occurs is a series of productions each of which is initiated by intentional selection, and hence no room for blind heritability and retention to take hold.

This tack is the one Lewens (2004) adopts to criticize the idea of evolution in the domain of technology. Even though he thinks that technology change can be construed as an evolutionary process, he argues that it is intractably difficult to discern reproduction of the sort that would support a process of natural selection. But it is not entirely clear that things are quite so simple. In the next chapter I frame this question in terms of how technologies—i.e., artifacts and material culture—as well as the ultimately biological intentions which generate and sustain them, get created. The dominant account that we have from philosophy of biology is that these entities are generated and propagated in virtue of their proper functions. I expound this approach to the technological domain through the lens of Ruth Millikan's historical theory of function and the criticisms that have been raised against it. Lewens is also concerned to deflate this theory, and I consider several avenues of reply against his criticisms in chapter four.

One such avenue of reply is the suggestion that technologies are memes that get copied without a direct mechanism of reproductive transmission. Defenders of memes often invoke the image of a parasite to illustrate the way in which technologies might go about getting themselves blindly reproduced while still utilizing the intentional selection of intelligent agents. In the same way that parasitoid worms might infect an insect or bird and co-opt its nervous system to ensure their own successful reproduction, technologies affect the organizational patterns and hence the patterns of human activities. The idea is not that any one particular technology has co-opted the behaviours of human agents, but perhaps more charitably that a critical mass of technological influences has impacted human behaviours and organizational tendencies to the point where the sorts of activities we pursue just tend to be the kind that are geared toward innovations. Even a footpath is like a parasite in that it channels human ambulatory activity along a trajectory which results in not only its being sustained, but in its expansion into new trajectories and routes.

This I freely admit is far-fetched, but the image is sound. The question of what exact processes constitute biological reproduction is not entirely settled. There are a number of reproductive processes in plants which seem to defy the basic idea of a copying process, yet are still counted as reproductive. Cuttings, for example, are used to keep a variety of plant species alive that would otherwise have died out. Sperber (2007) gives an interesting picture of the sorts of issues that make the notion of reproduction less clear cut than it might appear from biological theory. Some examples of parasites can be counted among these problem cases. If the notions of reproduction and propagation turn out to be constructed as resemblance notions (rather than e.g. natural kind processes, etc.), then we must confront the possibility that technologies can count as reproductive entities along a spectrum of reproductive activities that is not defined by necessary/sufficient criteria.

The proponent of technology change as an evolutionary process, and hence the proponent of TD, does face a steep uphill battle. But even the most basic issues are far from settled. My inexorability thesis is defensible given the rich range of recent work in theorizing about cultural evolution. Once we begin to frame the issues facing the study of phenomena in the domain of technology in terms of material culture (following e.g. Preston's account), we start to see that many of the very lucid concerns of Lewens and others miss the mark. The case is by no means closed, but there is enough reason to think the inexorability thesis is defensible as a component of nomological TD.

§1.1.3 The Autonomy Thesis

Let the following statement stand as TD's *autonomy thesis*: the basic competencies of human cognitive agency are a technogenic development. The term "technogenic" is intended to capture the idea that technologies have played an important role in generating features of cognitive agency. This gets at Bimber's notion of technology as a medium through which "the course of human events" is shaped. If our agency has been shaped over time by the artifacts that become available in our local culture, then the actions of those agents (individuals and groups) who make up our culture would be products of that pattern of influence. The term is also intended to recognize that an attempt to draw a straightforward causal connection here would be both futile and facile because agents and artifacts are intricately linked as far back as it is possible to trace, and their causal connection is largely what is at issue in many of these studies. The autonomy thesis proposes that technology plays a critical role in ratcheting up the particular feedback dynamic that drives the production of features of cognitive agency in human cultures.

How might one argue for this thesis? Does this formulation of the idea stray too close to what Bimber calls the normative reading, given the focus on culture? A number of methodological hurdles must be cleared in order to assess this idea. For one, it is difficult to make sense of "technology" broadly and monolithically construed. In §1.2 of this introduction I give an overview of the sheer variety of phenomena which philosophers have classed as artifacts, and it might seem foolish to think that criteria can be cited for such classification at all because none could capture all of them. However, these phenomena are the focus of rigorous empirical social scientific study, especially in anthropology. Because such investigations are scientific, we are under no obligation to give firm a priori criteria by which to classify the entities of interest. I make an effort to speak of technologies (plural) and of material culture, but monolithic characterizations are not out of order, so long as their provisional utility is kept well in mind.

Referring to the thesis as a claim about autonomy is an attempt to preserve the origin of the idea in social and political theory. There is a long, fraught tradition of discussion about the allegedly social nature of technology, about its impact on individuals and society, and about its moral status. The philosophical study of technology is, by contrast, a quite recent enterprise, and it has tackled these same questions in a rather more brusque and provocative manner. The notion of autonomous technology that has been taken up by philosophers (from its origin in social theory) was formulated largely in isolation from the philosophical study of the concept of autonomy as it pertains to moral agents and rational subjects of action. Again, there is no neat way to cleave the topics apart, and recent writers have sought to bridge the different traditions. The key point to emphasize up front is that the notion of autonomy to which this thesis refers is that of the "path dependency" of social outcomes in relation to technological factors.[19]

My major source on this aspect of the topic is Preston (2013). I open the discussion of TD's autonomy thesis with a chapter on action theory and a prominent attempt to characterize artifacts and their functions in terms of the use plans of agents. Framing the discussion of agency in this way gets us thinking about action, about the status of agents and items of material culture in both current and historical relation to each other, and about the role of intention in reproduction. I suggest that Preston endorses the autonomy thesis, and the defense she gives of innovation and human ingenuity suggests that she also worries about the commitment of her account to this sort of idea.

[19] Kroes and Verbeek (2013), 18.

I will try not to prefigure my discussion of the autonomy thesis in too much detail. My strategy is simply to show that while an account like Preston's is well grounded, and despite an admirable effort on her part to avoid committing to ideas like TD's autonomy thesis, her defense of innovation fails to bar such a commitment. The power she assigns to material culture, coupled with the relative frailty and diminutivity of her countervailing notion of creativity, are simply too much to avoid it.

The final chapter is conceived as a practical continuation of the themes discussed in relation to TD's autonomy thesis. As someone who inclines toward methodological naturalism but who also has great interest in the status of our normative heritage handed down through the manifest image, I want to take stock of the practical ramifications of the case I develop. Technology is a topic thoroughly rooted in the normative dimension of the world through its connection with human agency, yet it is most commonly associated with scientific knowledge. I present a case for the crucial role technology has to play in guiding our normative affairs in the cognitive domain. The topic of cognitive engineering has attracted attention from moral philosophers but few have addressed the epistemic and cognitive issues. I offer an argument for a normative approach to the design of cognitive systems that I hope adds to such an agenda.

§1.2 The Varieties of Artifacts

Bimber's nomological reading of TD makes it clear that the phenomena at issue are artifacts and not other types of social or cultural phenomena. Yet Preston's account, as we shall see in chapter five, does construe them as cultural. Even if they are cultural, perhaps this can be reconciled. This is not an issue we can resolve either by argument or by fiat at this stage. To provide some perspective at the outset, however, I would like to give an overview of the sheer variety of phenomena studies of artifacts aim to treat.

Consider Sperber's example of an old footpath leading through the woods from a settlement to a nearby water source.[20] Is this an artifact? It results from human activity, but it is not clear that the creation of the path was an intended outcome. The intention was to gather water, or to fish. The activity that achieved the goal of this intention did generate the path, but the path may not have formed any part of that intention.

Does the footpath count as an artifact? It would not exist but for intentional human activity, yet not only does it seem to have been unintended, over time it can be seen to constrain the intentions of other people. Those who come later to fish or drink may see the trodden brush and form the very different intention of using the path to get to their destination more quickly. Again, the intention this time may not be to make a path, but merely to follow an existing one. Yet this activity also helps to construct a path.

There are a wide variety of phenomena like this. Sperber characterizes the notion of an artifact, in its home domain of anthropology, as a family resemblance concept.[21] It is probably a wise strategy. There exist many proposals for defining artifacts and their functions, and most seem to rely on intentional activity. But there are also many kinds of cases that seem artifactual yet do not easily fit under this design model.

Artworks are perhaps the traditional paradigm of an artifact. Art refers to a notoriously broad range of things. The notion of an artwork can perhaps be viewed as an umbrella notion for artifacts in general. But this is not ideal because many borderline cases do not fit the model. So perhaps a work of art is best used as an umbrella for the class of

[20] Sperber (2007), 125.
[21] Sperber (2007), 124.

artifacts that *are* intentional. But this is also far from ideal because art is sometimes regarded as a spontaneous output rather than a design. Theorists want to leave room for the fruits of innovation and improvisation.

Technology is probably the next most prominent kind of artifact discussed by philosophers. Most writers agree this is an old topic, having been raised by Aristotle in several spots but generally neglected until the last century. As Bimber intimated, many writers have conflated what are actually techniques and social capacities with material items. Aristotle is doubtless partly to blame, referring to craftwork and skills in general terms without drawing firm distinctions between efforts and their products.

Some animals are artifacts. Here I will mention two prominent kinds of cases: animals whose breeding has involved artificial selection, and food. Artificially bred animals are usually regarded as artifactual in the sense that natural processes have been hijacked to serve the ends of human design. So while it is likely true that wolves would have continued to breed independently of human activity, and have done so in many lineages, those that were bred into the species of dog we have today were in a sense co-opted over successive generations. Humans now largely direct their breeding, and the mechanisms of origination and propagation seem not to matter much.

Food is an interesting case. Here we have a large, diverse family of kinds of artifacts and found materials pragmatically grouped under a catch-all label. Food includes artificially bred animals, but also some that breed naturally, as well as plants by both methods. It includes substances designed and manufactured by humans. It also includes fully natural substances, like water; we will see that Dipert classifies such useful but unmodified things as *instruments*. "Food" is perhaps too loose a term to capture much, but the looseness raises interesting challenges.

Another borderline case from the material realm is that of scrap, waste, and by-products. This category refers to those material objects that merely result from ordinary production and other activities of humans and animals, like the footpath. Sawdust comes from both human sawmills and the damming activities of beavers. Garbage typically comes from humans, but such waste can formerly have been any other kind of thing, say a tool or item of food.

Is *waste* a robust enough notion to support a distinct designation as a category of artifact? Or is it merely a useful label referring to discarded items? Is waste a distinct category of artifact or something entirely non-artifactual? Byproducts also pose this dilemma. Here we have materials that, like scrap, are not waste exactly but, unlike sawdust, are not anticipated. I do not think bright line distinctions between scrap, waste, and by-products can be drawn, nor is there a pressing need to do so. It is enough to keep in mind that byproducts include unanticipated materials that we might call side effects, and these can be useful, whereas waste and scrap are usually unusable (a claim routinely challenged by the ingenuity of dumpster-divers).

Toys are yet another interesting category. Here we depart almost entirely from a focus on material affordances in defining an artifact. It seems that any material item can be a toy, since "toyhood" is a role served by an item in a session of play or as imagined by a child (or, let's face it, an adult). What makes something a toy seems to depend wholly on the intention of the user. No outward design is required. The manner in which objects are co-opted as toys is quite improvised. We might even say that toyhood is more about patterns of use than features of material objects. I think it is an important category to note because it gets to the heart of both standard "plan" theories as well as the most difficult objection for those theories, the characteristic lack of planning in play.

Perlman cites the example of *unique* artifacts such as the International Space Station, various high-technology systems, most artworks, and prototypes. These cases are problematic for approaches to artifact function that focus on reproductive history. Some writers argue that such items do not possess functions at all. In chapter five, we see how Preston defends such a claim for prototypes and "phantom functional" items.

Lastly, there is a large variety of artifacts whose status is wrapped up in the various social relationships and accomplishments of individuals and groups. Money, for instance, is a material item in the sense that it is printed on paper, minted on metal, or stored in digital bits. But what confers value on these instantiations are social arrangements or, as Searle argues, linguistic status functions.[22] He also wonders about social institutions like presidencies and etiquette. These are not material objects, but they have effects in virtue of human intentions. Similarly the large range of abstract entities like plans and techniques, mathematical relations (if devised and not discovered[23]), and language itself—words, texts, and ideas—are contenders for artifactual status. I largely ignore all the above phenomena in favour of a focus on technological artifacts.

Let me briefly cite a few of the consulted works up front which I have not had much opportunity to discuss in the text. Studies on function and artifact teleology are quite numerous (Hilpinen 1993; Houkes, Vermaas, Dorst, and de Vries 2002; Perlman 2004; Houkes and Vermaas 2004; 2010; Vermaas and Houkes 2003, 2006a, 2006b;

[22] Searle (1995, 2010).
[23] Thomasson, (2007). If discovered or naturally occurring, abstract entities perhaps count as instruments.

Kroes and Meijers 2006; Costall and Dreier 2006; Hughes 2009; Krohs 2009; Kroes, Franssen, and Bucciarelli 2009; Krohs and Kroes 2009; Kroes 2010, 2012; Vaesen 2011).

Issues about the metapysics and ontology of artifacts have been discussed at length (Devitt 1991; Meijers 2001; Elder 2007; Houkes and Vermaas 2009; Vermaas 2009; Franssen et al. 2014). Discussions specifically on the topic of artifact kinds and realism are also available (Thomasson 2003; Soavi 2009; Carrara and Vermaas 2009; Houkes and Vermaas 2014).

There are discussions on reference to artifact kinds and kind terms as well as general semantic studies (Putnam 1975; Schwartz 1978, 1980, 1983; Putnam 1982; Nelson 1982; Hughes, Kores, and Zwart 2007; Marconi 2013). Recently, philosophers who study artifacts have begun to consult the research on categorization in cognitive psychology (Thomasson 2007; Vaesen and van Amerongen 2008; Houkes and Vermaas 2013; MacIntyre 2013). A classic reading on this subject is Bloom (1996) and there are scores of cognitive psychology studies cited in each of these papers.

Discussions of classic epistemological topics concerning artifacts can also be found (Wilson 1995; Morton 2006; Hilpinen 1995, 2006). And there are, of course, countless volumes on the ethics of artifacts and technology; philosophers who study artifacts have made some recent contributions (Franssen 2006; Vaesen 2006; Illies and Meijers 2009; Verbeek 2011).

§1.3 Sketch of the Dissertation

Let me briefly sketch the structure of the discussion. The second chapter examines the historical theory of artifact reproduction. In advance of my discussion of TD's inexorability thesis and evolution (Ch. 4), here I consider some general criticisms about the role of natural selection in this kind of argument, with reference to Robert Cummins and functional analysis as well as to Millikan's exchange with Preston on functional pluralism. This expository chapter has the goal of introducing some basic concepts in the study of the functions of technologies and their reproduction.

In the third chapter I present the most prominent alternative to historical theories of function: the use plan theory that originates with Dipert (1993). The main idea is that artifacts acquire functions in virtue of agent plans, subject to rational assessment. I explore Dipert's basic argument and his classification scheme. I consider an elaborated version of this picture given by Houkes and Vermaas (2010). Their notion of use plan clarifies the basic picture and makes a notable gain with respect to function attribution.

In the fourth chapter I consider TD's inexorability thesis by investigating evolutionary models of reproduction of material culture. Does technology change occur by a process of natural selection? I first present Beth Preston's (2013) attack on standard intentionalist theories of function in material cultural reproduction. In §4.2 I outline a major objection to evolutionary models of technological innovation, namely the transmission problem identified in Lewens (2004). Then in §4.3 I confront this objection from two standpoints, the philosophy of memetics and the dual inheritance model of cultural evolution. I concur that memes are probably not adequate to defeat the objection, but the dual inheritance model is found to be more plausible than Lewens allows.

In the fifth chapter I present Preston's theory in more detail. In §5.1 I present Preston's view on action, proper function for material culture, and the sociogeneric stance she defends alongside it. In §5.2 I defend her sociogeneric stance in connection with

TD's autonomy thesis. In §5.3 I consider the compelling view of creative innovation Preston develops as a means of undercutting a reading of her account like the one I give (in §5.2). And in §5.4 I present reasons why her pre-emptive response might be seen as falling short of the autonomy thesis.

In the final chapter I confront some of the practical consequences of a technogenic account of cognitive agency. In §6.1 I motivate the normative dimension by examining the commitment to a narrow locus of control shared by most extended theories of cognition. In §6.2 I examine the recent appeal some of these theorists have made to responsibilist theories of knowledge to preserve their commitment. This gives me an opportunity to explore factitious intellectual virtue as a way to defend these sorts of appeals. And in §6.3 I argue that factitious virtue has several benefits as a normative design model for the practice of cognitive engineering.

Function and Reproduction

This chapter examines the historical theory of artifact reproduction. In advance of my discussion of TD's inexorability thesis and evolution (Ch. 4), here I consider some general criticisms about the role of natural selection in this kind of argument, with reference to Robert Cummins and functional analysis as well as to Millikan's exchange with Preston on functional pluralism. This expository chapter has the goal of introducing some basic concepts in the study of the functions of technologies.

The inexorability thesis ultimately comes down to an account of technology change and reproduction. The most prominent available account of the reproduction of artifacts is found in Ruth Garrett Millikan's *Language, Thought, and Other Biological Categories: New Foundations for Realism*. I outline the relevant threads in her main discussion and consider how artifacts fit into it. Her basic aim is to develop a naturalistic theory of intentionality, especially of language and thought; an account of artifacts turns out to be a bonus. I focus on the qualifications to her notion of proper function that are most pertinent for artifacts. I then consider criticisms of her argument, both at a general level, with reference to Cummins and functional analysis, and at a level that targets artifacts more directly, in reference to Millikan's exchange with Preston. This expository chapter sets some of the stage for a discussion of the inexorability thesis in chapter four.

§2.1 Intentional Devices

The express aim of Millikan's book is to give a naturalist account of a wide variety of phenomena we associate with human cognition: primarily language and thinking, but also the use of tools and other artifacts. The distinctive mark of her approach is to explain these phenomena in terms of what she calls proper functions, a technical phrase referring to the historical causes of the proliferation of an item's ancestors. Though my interest is in how Millikan's theory accounts for material cultural reproduction, it is prudent to present her theory first in a quite general way because it is intended to encompass so much of the psychological and biological phenomena which join to generate the items of interest. Let me begin by rehearsing Millikan's motivation in two respects: intentional icons/devices, and malfunction/misrepresentation.

One way to elaborate on the aim of Millikan's book might be to say she wants to give a comprehensive account of all the devices that arise in nature. We can profitably use the analogy of biological devices to explain the wide variety of phenomena at issue. All these phenomena, e.g. language, behaviour, cognition, etc., acquire an intentional property in virtue of their historical lineage rather than their actual capacities/dispositions as token items. Millikan characterizes this property in connection with the functional descriptors familiar from evolutionary biology. But as we shall see, what she calls "proper functions" differ somewhat from the usual biological understandings of function.

Millikan is principally concerned to explain language and meaning. She takes the characteristic feature of language to be its capacity for misrepresentation, i.e. the fact that it can fail in its various acquired roles (representation, denotation, etc.). This is in keeping with e.g. Frege's view that intentional contexts affect the validity of logical inference: for instance, identity substitution and existential generalization fail to hold in the context of propositional attitudes. (It is illicit to infer that someone who believes that the morning star is Venus also believes that the evening star is Venus, despite it being true that the morning star and the evening star are one and the same entity, namely Venus.)

Millikan suggests that this mark of the intentional is actually much broader than it might at first appear: it captures phenomena beyond just the psychological domain. In particular, she argues that it also characterizes biological structures which operate as copying devices. Intentionality is cast on her account as an acquired trait, something that accrues over time in virtue of the ancestry of a given device. Misrepresentation in language and thought turn out to be particular instances of biological malfunction generally.

This approach is neither reductive nor fully normative. It attempts to combine the best of each. Millikan characterizes linguistic phenomena as "intentional icons"—particular sorts of devices equipped with proper functions. Most of her book pursues an application of these icons to various problems in philosophy of language. But the account has much broader application and power. In this chapter I expound what is most important for getting clear on the reproduction of material culture or artifacts: how these devices manage to get made, to spread, and *to do* the things they do.

§2.2 Proper Functions

A device—biological or artifactual—may have multiple functions. Several of the functions of a device may be proper functions. According to Millikan, proper functions are a matter of history: those effects of a device's ancestors which are primarily or conjointly or disjointly responsible for having propagated the lineage can all be characterized as proper functions of that device, to varying degrees of proximity. A device may have the capacity to perform other functions and it may have several distinct proper functions.

Millikan qualifies her term "proper function" in a number of ways. Most importantly for our purposes, she speaks of *direct* proper functions, *derived* proper functions, *relational* proper functions and—as an elaboration of relational proper functions—*adapted* proper functions. Artifacts can have both derived and direct proper functions, but they always have at least one derived proper function because their proper functions are always acquired in virtue of the proper function of some prior device.

Millikan warns in a later paper that her various qualifications "do not widen or narrow the definition of proper function, they merely make it easier to talk about the phenomena it captures."[24] She forges one term, and then elaborates on it to help distinguish the various ways in which it can be useful. So at the genetic level, we might speak of the direct proper functions of DNA molecules to direct the chemical processes that ultimately generate organs, while at the phenotypic level it is more useful to speak of particular organs in terms of proper functions that derive from those associated with the chemical instructions. (Millikan also speaks of stabilizing and standardizing proper functions when discussing language devices, but I will not discuss these qualifications.)

Millikan begins with a notion of "reproductively established family" to help define proper function generally.[25] She gives definitions of first-order and higher-order reproductively established families. The basic idea is that items resembling one another as the result of some reproductive process count as families.

In the case of "first-order" reproductively established families, the reproductive process is "something like copying."[26] Millikan herself cites photocopiers as an example of something that produces straight copies in this way but, to remain clear about the different emphases in ascribing functions to artifacts and to biological devices, I will use an example from organic chemistry: gene tokens produce new tokens that are so similar to the original molecules that we can regard them as molecular copies of it. Cellular mitosis is another example of the sort of process which yields what Millikan has in mind by her

[24] Millikan (1999), 202.
[25] Millikan (1984), 18-25.
[26] Millikan (1984), 18.

notion of a first-order reproductively established family: a cell divides into two daughters and each resembles the original. The noun she uses informally is “a copy” but we could also supply “a replicate” to serve the same purpose.

More strictly, Millikan defines *B* as a reproduction of *A* if and only if the following three conditions are met:

> (1) *B* has some determinate properties p_1, p_2, p_3, etc., in common with *A* and (2) below is satisfied.
> (2) That *A* and *B* have the properties p_1, p_2, p_3, etc., in common can be explained by a natural law or laws operative in situ, which laws satisfy (3) below.
> (3) For each property p_1, p_2, p_3, etc., the laws in situ that explain why *B* is like *A* in respect to *p* are laws that correlate a specifiable range of determinates under a determinable under which *p* falls, such that whatever determinate characterizes *A* must also characterize *B*, *the direction of causality being straight from A to B*.[27]

What these conditions mean is that for one item *B* to reproduce—i.e. to be a copy of—another item *A*, we must be able to describe the causal history which led from *A* to *B* in such a way that it is clear why *B* had to resemble *A*. This causal history must be clear with respect to relevant determinate properties, i.e. those properties that are of interest in describing or explaining *B*. The two items had to be alike because, given some determinate property within a specifiable range of variation and the relevant laws for that situation or environment, if *A* varied with respect to that property then so too would *B*.

This definition of reproduction stipulates how causal history determines the properties of an item’s progeny, and it applies only to first-order reproductions. In Millikan’s words: “Any set of entities having the same or similar reproductively established characters derived by repetitive reproductions from the same character of the same model form *a first-order reproductively established family*.”[28]

[27] Millikan (1984), pp. 19-20.
[28] Millikan (1984), 23.

The notion of a higher-order reproductively established family applies to items produced not as direct copies of similar items, but rather in accordance with the operation of mechanisms that do produce straight copies. Millikan cites such organs as hearts and livers as items belonging to higher-order reproductively established families: "although my heart is not a copy of my parents' hearts, it was produced under Normal conditions in accordance with the proper functions of certain of my genes which *were* directly copied from my parents' *genes*."[29] Millikan offers a tripartite definition of higher-order reproductively established families which I will not rehearse here.

The definitions both of higher-order families and of proper functions follow from the notion of the first-order family conjoined with Millikan's biological notion of "Normal". A Normal explanation "is the *least detailed* explanation possible that starts by noting some features of the structure of members of [a reproductively established family] *R*, adds, some conditions in which *R* has historically been when [its members] actually performed [a proper function] *F*—these conditions being uniform over as large a number of historical cases as possible—adds natural laws, and deduces…how this setup leads to the performance of *F*."[30] The basic idea of a Normal condition is that it roughly specifies the conditions or relevant features of a structure's role in an environment in explaining the reproduction of the lineage. (It is not a statistical notion. She cites the example of sperm, which execute their proper functions in the presence of an egg, ovum, etc., even though the vast majority of sperm never actually encounter a situation with these conditions.[31])

Normal conditions figure prominently in Millikan's explanation of the reproductive processes of certain families of devices or items. Normal conditions bridge the causal

[29] Millikan (1984), 25.
[30] Millikan (1984), 33.
[31] Millikan (1984), 34.

gap in biological explanations between historically related items which are only indirect effects of prior causes. Direct reproduction of the kind found in first-order families is causally simple compared with such devices as hearts and other organs. The mitotic phase of a cell cycle decomposes into a variety of processes that ultimately terminates in the physical properties of the chemical kinds involved.

The thrust of the functional explanation of hearts in biology is that they exist in organisms today because similar devices in the past contributed to the replication and thus the survival of the mechanisms which produce hearts, i.e. genes. So the condition under which ancestral devices contributed to a lineage's survival, i.e. the effect of circulating blood, is itself part of the causal explanation of higher-order devices even though the intergenerational causal relation can be much less direct than in the case of items belonging to first-order families. Millikan asks, "How can a thing result from a prior *causing* as opposed to resulting from an effect produced via the causing?"[32] And she responds: "My suggestion is that when it is in part because *A*'s have caused *B*'s in the past that a positive *correlation* has existed between *A*'s and *B*'s, and the fact that this *correlation* has existed figures in an explanation of the proliferation of *A*'s, then it does make sense to say that *A*'s exist in part because *A*'s caused *B*'s."[33]

In addition to bridging this causal gap in biological explanation, the notion of Normal conditions makes it possible to define malfunction of token items and to include many other cases where a device fails to perform the proper function it nonetheless possesses. This feature constitutes the chief virtue of Millikan's theory. The key point is that even malformed tokens of a device family will have the proper function belonging to that

[32] Millikan (1984), 26.
[33] Millikan (1984), 26.

family. A poorly developed token of the heart family is still the product of an appropriate reproductive history, and so its function will still be to pump blood even if it lacks the structural capacity to do so. Specifying Normal conditions allows one to say, non-counterfactually, what has gone wrong in order that a token has malfunctioned.

The basic gain of replacing talk of functions with that of proper functions is that the latter captures not only the reproduction of material items but also both learned and novel varieties of behaviour, as well as malfunction and misrepresentation in the semantic domain. It allows Millikan to ascribe proper functions to such diverse phenomena as words and other language devices, social customs, viruses, artifacts, beliefs, mating dances, genes, organs, and the instincts of animals, while at the same time explaining how tokens of all these families can fail to perform their functions.

§2.2.1 Direct Proper Functions

I mentioned that proper functions can be attributed according to different explanatory aims. Derived proper functions will be attributed to organs, behaviours, and artifacts because the most proximate proper functions of these devices derive from mechanisms which produce them in the course of carrying out *other* proper functions (often also derived), these further proper functions themselves being derived from further devices with direct proper functions to copy. While for most purposes, biological structures could be satisfactorily explained by citing derived proper functions, direct proper functions are what ultimately account for reproduction.

Millikan provides both a formal and an informal characterization of direct proper functions. The intuitive, informal idea she puts as follows. "A function *F* is a direct proper function of *x* if *x* exists having a character *C* because by having *C* it *can* perform *F* [where ancestors of *x* performed *F* in the past due to having *C*].”[34] The idea is that historically, the effects of an item or structure which have tended to contribute to its reproduction can be identified not only as functional but as proper functions, i.e., effects which favourably increased the likelihood of further reproduction.

Combining the earlier notions of Normal conditions and reproductively established families, Millikan develops the basic notion of a direct proper function in two stages. First, she defines the notion “ancestor of a member of a reproductively established family”, and then she proceeds to define “direct proper function”.

Items or structures only acquire direct proper functions if they result from some history of reproductive activity. Thus the very first item (or even the first few items) in a given lineage will have no proper function—but its immediate offspring might have one, and later descendants would. Evolutionary processes tend to be slow and to spread across many generations. Many errors result from small mutations and interactions between environment and organism. Until reproduction recurs regularly, there can be no reproductively established family of items and hence no items with proper functions.

Millikan proposes to define direct proper function in terms of an item’s membership in some reproductively established family of similar items. She begins with the notion of an ancestor of such a member before defining direct proper function, and I will quote in full this important passage containing both definitions:

> (1) Any member of a (first-order) reproductively established family from which a current member *m* was derived by reproduction or by successive reproductions is an ancestor of *m*.
> (2) Any temporally earlier member of a (higher-order) reproductively established family which member was produced by an ancestor of the device that produced a present member *m* is an ancestor of *m*.

[34] Millikan (1984), 26.

(3) Any earlier member of a (higher-order) reproductively established family that a present member *m* is similar to in accordance with a proper function of a producer that produced both is an ancestor of *m*.

Now we define the notion "proper function."

Where *m* is a member of a reproductively established family *R* and *R* has the reproductively established or Normal character *C*, *m* has the function *F* as a direct proper function [if and only if]:

(1) Certain ancestors of *m* performed *F*.

(2) In part because there existed a direct causal connection between having the character *C* and performance of the function *F* in the case of these ancestors of *m*, *C* correlated positively with *F* over a certain set of items *S* which included these ancestors and other things not having *C*.

(3) One among the legitimate explanations that can be given of the fact that *m* exists makes reference to the fact that *C* correlated positively with *F* over *S*, either directly causing reproduction of *m* or explaining why *R* was proliferated and hence why *m* exists.

It follows that if any member of a reproductively established family has a direct proper function, all members of which this member is an ancestor have this proper function too.[35]

This definition entails that any item which results from the reproduction of members of a reproductively established family will have the same direct proper function as its ancestors, even if it never fulfills that function itself (or e.g. is malformed in some way).

The thrust of the second definition is in the second and third clauses. If we consider the members of *R* relative to the broader set of items *S* (i.e., the members of *R* plus items which do not possess the character or trait *C*), and we see that a positive correlation occurs between the fact that members of *R* possess *C* and the fact that *C* produces the effect *F*, then it makes sense to explain the existence of members of *R* (e.g., *m*) in virtue of their capacity to perform *F*. The correlation arises over time, so Millikan's theory is known as the historical or etiological theory of proper function. She is adamant that her theory is not a conceptual definition of "function".[36] "Proper function" is a particular term for a particular purpose, and while it is designed to capture the intuition behind the ordinary notion of function, it is neither an analysis nor a definition of this notion.

[35] Millikan (1984), pp. 27-28.
[36] Millikan (1984), 18.

To reiterate, Millikan emphasizes that this theory attributes proper functions to items regardless of token performance. Attributing a proper function to a token is not a matter of its actual dispositions or capacities. It need not be able actually to perform its proper function. It is important to make such attribution a matter of historical reproduction because if it were to depend on performance rather than history, then it would not be possible to attribute proper functions to malfunctioning tokens. And since Millikan wants to preserve the form of functional explanation in evolutionary biology, she must find some other way to attribute such functions to tokens without tying it to performance.

§2.2.2 Derived Proper Functions

"The proper functions of adapted devices are derived from proper functions of the devices that produce them that lie *beyond* the production of these adapted devices themselves. I will call the proper functions of adapted devices *derived proper functions*."[37] A derived proper function is always the proper function of an adapted device.[38] As adapted devices, all the proper functions we associate with human cognition (e.g. behaviour, language, artifacts) will turn out to be derived proper functions.[39]

Adapted devices—devices with relational proper functions where one of the relata is an *adaptor*, i.e. a specified feature of the environment—can acquire their derived proper functions in several ways. Either the adapted device derives its proper function from the producer alone, or from both the producer and the adaptor.[40] Millikan's example is the skin pigmentation in a chameleon. In general, this mechanism has the proper func-

[37] Millikan (1984), pp. 41-42.
[38] Millikan (1984), pp. 41-42.
[39] Millikan cautions that the term "'adapted' is not connected to 'adaptation'" (2002, 129) despite there being a "special similarity" between them (1984, 40).
[40] Millikan (1984), pp. 41-42.

tion to match the skin with a background colour. At this level of abstraction, without specifying a relatum to which the mechanism will relate itself, we can call this a *relational* proper function of the entire structure: skin plus unspecified colour. Say the chameleon is sitting on a brown rock when a predator approaches. The mechanism will generate a match between the skin and the rock. Once *brown* is counted as a fixed relatum in the relational structure, the rock itself is referred to as an *adaptor*. The entire structure—skin plus brown rock—is now considered an adapted device.

Adapted devices are tokens generated by other devices or mechanisms that have the (direct) proper function to produce relational structures, e.g. genes that code for matching in skin pigments. But adapted tokens have only derived proper functions because they are, strictly speaking, novel devices with no direct ancestries. However, they can begin to influence a reproductive lineage through the effects they exert. This is how higher-order reproductively established families become recurring features of organisms. Structures like organs, limbs, and behaviours, which make no direct copies and have no direct ancestors, nevertheless get replicated regularly, and can generate new adapted devices with further derived proper functions over time.

Millikan develops and emphasizes the distinction between direct and derived proper functions because this is how she proposes to explain misrepresentation in language and malfunction in artifacts—that is, as varieties of maladaptedness (keeping in mind that "adapted" does not refer to adaptation). Adapted devices can acquire further derived proper functions over and above the ones they already possess in virtue of their membership in reproductively established families. There will be a certain Normal explanation for how the adapted device is produced. If the derived proper functions it acquires in being Normally produced fail to fulfill their "intended" relation to the adaptor, the device is said to be maladapted. "Where adapted devices are maladapted, it can happen that one and the same device acquires *conflicting* proper functions."[41]

Cognitive and semantic misrepresentation, for example, turn out to be a malfunction in this manner. Millikan's example here is bee dances. A bee whose dance is maladapted, thus directing a gatherer bee to a site with no nectar, still uses a dance that has the derived proper function of representing the location of nectar. This representation does not hang on the adaptor itself, i.e., the thing to which the adapted device is adapted—the actual location of the nectar—because what confers its proper function is the history of its production.

Misrepresentations in language result from a conflict between the derived proper functions that produced an adapted device, and that adapted device's adapted proper functions, which for whatever reason are not performing in the way they are supposed to, given the device's reproductive history. For example, when "Stephen holds Millikan's book" fails to represent how things are (Stephen is actually sitting on top of her book, or nowhere near it), then a conflict exists between how these language devices are being adapted as a token device, and the derived proper functions—their senses—for which they have historically been reproduced. The verb "to hold" in its third person present indicative active form has historically been reproduced in accord with a particular sense (that someone is grasping or clutching something). But the person who spoke the above sentence, if spoken literally, did not adhere to this established sense. Hence there is a conflict between the senses for which its components were historically reproduced, and the newly minted derived proper function to which it is being adapted in this case.

[41] Millikan (1984), 43.

Another example might be reference: names are a useful unit of language and have proliferated in virtue of a capacity to denote fixed referents. But occasionally two names with nominal senses that have been taken as different will turn out to have been referring to the same entity all along, as in the well-worn example of the morning star and the evening star. In this case, the derived proper functions of these nominal phrases (i.e., to denote exactly one entity) conflicts with their adapted proper function as names which denote different entities, since they in fact denote the same entity. Millikan's theory allows us to account for misrepresentations that fail not just because they do not refer or represent (even if they do fail in these respects) but because of what the device is supposed to do in the first place.

§2.3 Cummins and Functional Analysis

In this section I offer a brief segue into Millikan's discussion of artifacts by outlining a competitor: functional analysis. I proceed in this way to make clear the heart of the dispute between Millikan and Robert Cummins. What Cummins emphasizes is that the distinctive use of functional characterization in the natural sciences involves a relativity of function attribution. He argues that functions only figure in scientific explanation relative to the capacity of some structure to contribute to an effect of a containing system.

Cummins designed his method of functional analysis to respond to a particular dilemma about functional characterization in the sciences. Sometimes functions are attributed as though it explains the presence in biological and other systems of traits that perform the function. Cummins cites a distinction between functions and mere effects: e.g., hearts produce heart sounds (an effect) but the function of hearts is to circulate blood and not to generate these sounds.[42] It is not clear exactly how the effects characterized as functions differ from other ordinary effects in such a way as to explain how the structure in question came to be.

"An attempt to explain the presence of something by appeal to what it does—its function—is bound to leave unexplained why something else that does the same thing—a functional equivalent—isn't there instead."[43] The problem, according to Cummins, is that "to 'explain' the presence of the heart in vertebrates by appeal to what the heart *does* is to 'explain' its presence by appeal to factors that are causally irrelevant to its presence."[44] A dilemma arises when we try to do so. On the one hand, we could try to say that some trait *i* non-exclusively contributes to the performance of a system. In this case, the explanation which cites *i* is invalid because some other trait or item could make the same contribution *i* does. On the other hand, if we say that *i* alone is exclusively capable of making that contribution, then the explanation is valid but not sound, because again it turns out that other similar structures could make the same contribution. So it is useless to appeal to function to explain the presence of a trait since any such explanation would not explain why this trait rather than another were present to perform the function.

(Cummins notes, not unimportantly for our larger purposes, that the only case in which it is legitimate to appeal to function in explaining why a structure exists is that of artifacts: the reason why some type of artifact exists can be that what it does has given agents a reason to create and to maintain it. But in a footnote he remarks that the items functionally characterized in the sciences "are typically not artifacts."[45])

[42] Cummins (1975), pp. 741-742. Cf. Hempel (1965).
[43] Cummins (1975), 745.
[44] Cummins (1975), 746.
[45] Cummins (1975), 747.

So functional explanation cannot generally be a form of or a substitute for the causal explanation of traits in evolutionary biology. The success or failure of organisms cannot be appealed to in explaining why they possess given traits.[46] What functional characterization helps to explain in evolutionary theory are the biological capacities of individual organisms, on the basis of which it is then possible to infer the fitness of populations of these individuals.

Attributing a function to a component trait is always relative to the analysis of an organism's capacities, specifically those capacities that in turn contribute to the fitness of the species. It is only legitimate to attribute function in virtue of a containing system we are concerned to explain, relative to an analysis: "To ascribe a function to something is to ascribe a capacity to it which is singled out by its role in an analysis of some capacity of a containing system. When a capacity of a containing system is appropriately explained by analyzing it into a number of other capacities whose programmed exercise yields a manifestation of the analyzed capacity, the analyzing capacities emerge as functions."[47]

Paired with a "subsumption strategy" for decomposing bottom-level functions into purely mechanical physiological, chemical, and physical processes under general laws, functional analysis gives us all the tools we need to preserve functional attribution as part of genuine explanation in the sciences. Attributing function is not the aspect of an explanation which does any causal heavy lifting, but it is an indispensable one. However, as we shall see, there are challenges that arise in the case of artifacts. The specification of systems that contain artifacts involves reference to intentional and psychological processes, which Cummins elsewhere characterizes as falling under ad hoc or sui generis ex-

[46] Cummins (1975), 751.
[47] Cummins (1975), 765.

planatory schemes, rather than lawful generalizations. And most critics agree that his account hastily dismisses the notion of malfunction altogether.[48]

§2.4 Artifacts

In this section I shift focus to discuss how Millikan's account treats the case of artifacts as a particular kind of intentional device. Artifacts are much like linguistic and cognitive devices, but one step removed in a lineage of derived proper functions. Millikan's theory offers a compelling picture of artifacts that is continuous with her clear remarks on cognition, language, and intentional devices. It deftly handles the curious transitions and transformations the proper functions of artifacts regularly undergo.

§2.4.1 Artifacts as Adapted Devices

Like most biological devices on Millikan's historical theory, artifacts derive their proper functions from other devices which themselves already possess both derived and direct proper functions. The devices which produce artifacts are, after all, human cognitive mechanisms and behaviours (often coordinated in groups). So artifacts already come equipped with multiple sources of proper function, and in the case of widely used tool types, some even acquire direct proper functions through their reproductive histories.

By focusing on reproductive history, Millikan's view explains why artifacts so often lend themselves to uses both novel and enduring: the reasons why a type of artifact exists in the design it does may be completely separate from the reasons why that design continues to proliferate in a social environment. A desire to smooth fabric with steam

[48] Cummins disputes that this is a consequence of functional analysis (personal correspondence).

may explain the origin of the clothes iron, but tokens of the original design now enjoy a new life spent propping books up on shelves. Plastic soda bottles are now reproduced to serve as containers of liquids, but were originally designed to store gases.[49]

What primarily explains artifacts on Millikan's theory is ultimately the same explanation that goes for language, thought, and purposive conduct: heritable structures produce adapted devices (e.g., psychological traits) that relate to features of the environment in useful ways. The human organism is an arrangement of delicately balanced organs and structures all operating in consort to exhibit a variety of traits such as coordinated muscular motion, neurochemical circuitry, sensory processing, etc. It may be misleading to consider the functions of the body's devices in isolation from each other given that they did not evolve in isolation. We have tongues for discriminating tastes and ears for hearing sound, but these organs also aid in the production of speech, and this is just one among their many direct and derived proper functions.

Which proper functions are more or less proximate (i.e., primary) is a matter of the specific evolutionary history of the device in question. We have thoughts and purposes with their own derived proper functions, produced by inherited devices in the course of performing their own direct proper functions. In accord with those purposes (e.g., mechanical designs) we produce instruments and tools to help carry them out. Put another way: the derived proper functions behind our intended purposes can be performed by creating sounds *or* tools, with similar success. "If the specific purposes of human behaviors coincide with derived proper functions of these behaviors, the purpose for which an artifact was designed and made is a derived proper function of that artifact, the

[49] The example is Preston's (1998).

artifact being itself an adapted device. So also with language."[50] What makes artifacts different from biological devices is the same thing that makes language devices different: they have complex sources of proper function, and those sources include purposive intentional phenomena.

Since Millikan attributes proper functions to artifacts in virtue of the adapted roles they play, derived from the proper functions of human cognitive devices, we may wonder whether the proper functions of artifacts are determinate. It would seem they are not determinate in the sense of being fixed once and for all, even though they are determined by historical lineage and the relation they bear to other devices with derived proper functions. But psychological and linguistic devices are not fixed in this sense either. They too can undergo changes in proper function as a result of changes in the environment. Artifacts are just more liable to change. Partly this is due to the fact that neither users nor designers are privileged in the attribution of proper function to artifacts. Designers, as Dennett points out, are merely other users.[51] The proper function of an artifact does not derive solely from the original purpose for which it was created. It derives from whatever most proximally and fully accounts for the reproduction of the lineage.

So it is important to note that for Millikan, no proper function is ever determinate in a fixed, permanent way, not even in the case of biological devices with direct proper functions. Things can change. But in the case of artifacts, the lack of determinacy can be even more pronounced. Instances of drift may be more common than in organisms. Focusing on the reproductive history of an artifact gives us a purchase on current proper function, even if it is more apt to change than (say) the proper function of a heart or kid-

[50] Millikan (1984), 48.
[51] Dennett (1990), 86.

ney. This raises the central question of the inexorability thesis: do artifacts *evolve* in the same way as the biological and psychological devices from which they derive their proper functions?

§2.4.2 Selection in Artifact Evolution

In ordinary systems in which cognitive processes play no significant causal role, it is generally agreed that variation drives selection. It is clear that (functional) traits get selected because when competition arises in a population due to changes or minor modifications, the variants with more effective designs or structures stand a better chance at passing on their features. The situation in the case of artifact function and design bears at least a surface resemblance but the selective outcome is achieved through less direct means than brute survival. The chief components of the environment inhabited by artifacts are social and psychological devices. These devices comprise the "survival context" in which artifact designs vary and compete. Let me use two examples to illustrate: domestically bred animals and simple hand tools like the common screwdriver.

Dogs descend from wolves that, it is thought, were originally drawn to human dwellings when foraging for food, etc. Humans began to raise wolves and eventually bred them for various purposes (labour, hunting aids, security, etc.). Some lineages served these purposes better than others, so methods of preferential breeding were gradually devised. Over time, the original direct proper functions of lupine biology ceased to explain the continued reproduction of the subspecies, and the practical purposes of human beings overtook the role of those direct proper functional effects in the causal history. A new explanation of the continued reproduction of these lineages arises. They continue primarily because of the intentions of human beings, so their ultimate, most proximal proper functions are now also our derived proper functions.

All this is grossly oversimplified, but it is more or less the kind of explanation that a historical theory like Millikan's would offer. A slightly different example: simple tools. A hammer, screwdriver, or a basic weapon like a spear starts off with a derived proper function even before it is ever reproduced. (This is different from the wolf-to-dog case only in that no pup is born with such a derived proper function unless it is the result of intentional breeding.) Whether the device is crafted with that derived purpose in mind or it is found on the ground and pressed into service for that purpose is not important: in both cases, the derived proper function is the same.

Some tokens of an artifact type will also acquire direct proper functions, and this is where artifact selection begins to play a role. Selection does not explain the production of unique or "one-off", non-reproduced devices (in chapter five, we see how Preston argues that we should accept that prototypes have no proper functions). If a design proves valuable enough, it may get replicated more often, generating more tokens of that variant.

Once the trait in question begins to help explain continued reproduction of a family of similar items within a population, that lineage acquires a direct proper function in virtue of that trait. A purposefully sharpened rock might make a better spear point than a rock that just happens to be found with a sharp edge, but we cannot say the sharpened variant has been selected until it begins to edge out its competitor within the spear population. So variation drives artifact selection, but the environment in which artifact designs compete is described by human intentions.

Artifacts can acquire their own direct proper functions in virtue of how well they serve our intentions because in the course of serving them, some designs will be favoured (from amongst a range of functionally equivalent but nonetheless varied competitors) and will account for their continued reproduction. Millikan writes:

> Tools that have been reproduced (as have traditional carpenter hand tools) because of their success in serving certain functions have these functions as *direct* proper functions. But all tools have as *derived* proper functions the functions that their designers intended for them. When a tool has both of these sources of proper function, they usually coincide. But the proper function that derives from the intention in design is always there in the case of tools. Tools simply *as such* have only *derived* proper functions. The intent with which a specific user uses a tool on a specific occasion corresponds to still a third proper function, a proper function of the user's behavior, however, rather than of the tool itself.[52]

What is clear from this passage is that both designers and users are sources of derived proper function in artifacts, even if the proper functions derived from the users do not help to explain how a variant originated. So I take it that what Millikan means when she says that "the proper function that derives from the intention in design is always there" is that the designer's intention will always feature in an account of how the artifact came to be. The proper function derived from the designer's intent is always part of the history of the device, whether it goes on to a long lineage or ends with the first token. An artifact may acquire other proper functions: *direct* proper functions in cases where the proper function most proximally accounts for the reproduction of the lineage, and other *derived* proper functions the sources of which are users. But Millikan seems to say that this third source of proper function is not carried by the artifact itself, but rather is attributed to the user's own cognition. I pick this point up below in chapter two (§2.5.4).

[52] Millikan (1984), 49.

§2.5 Criticisms and Replies

Millikan must answer to Cummins's criticism that she fundamentally misconstrues the role of natural selection. I then discuss Millikan's exchange with Preston on pluralism about function. Preston argues that Millikan backs herself into a corner, with artifacts to one side and exaptation to the other. Millikan responds that the corner is not so tight as Preston alleges. Both critical exchanges serve to clarify what Millikan's theory has to say about artifacts, their reproduction, and their abilities. And both objections are at root trying to clarify the role of selection in historical theories.

§2.5.1 Where Cummins Objects

Millikan's general conception of evolution comes under attack by Cummins on several fronts. Cummins regards Millikan's theory as a (neo-)teleological view. In the specific case of artifacts, this point alone would not bother many philosophers, since artifacts are traditionally cast as paradigm teleological entities. But as Perlman notes, not only does Cummins deny that functions play any role in the natural selection of biological devices, he argues for "the more extreme position that a structure's function *cannot* play a role in its etiology."[53] Logically no structure can be described as functional apart from its being explicitly specified as a component in some containing system.

Cummins argues that historical theories of biological function get the process of selection dead wrong. The processes driving selection are totally insensitive to function.[54]

[53] Perlman (2004), 13.
[54] Cummins (2002), 162.

It is driven not by the effects etiologists single out as functional but rather by the developmental processes of organismic structures. Selection, he argues, "is sensitive to the effects that are functions, but is, in the sense relevant to neo-teleology, utterly incapable of producing traits."[55] Since heritable traits are all that are subject to selection, function simply plays no driving role in the processes guiding it. Traits do not spread because of the effects that count as their functions.[56] Selection spreads traits due to variation between already functionally equivalent structures. Because selection presupposes functional equivalence between competing structures, the effects designated as functional cannot account for the spread of selected traits. All the variant structures on which selection acts will have the same function—e.g., all wings have the function to enable flight. "Functions just do not track the factors driving selection."[57]

Cummins's criticisms carry over to the case of artifact evolution since artifact selection is just a special case of selection in general. The major difference lies in specifying the containing system and the criteria by which its effects are explained. Artifacts contribute to containing systems comprised of intentional processes. Even though according to Cummins, these particular kinds of effects fall under the province of psychology and the social sciences to explain, the basic explanatory strategy is similar.[58] A tool such as a screwdriver contributes to a capacity of individuals to drive screws, e.g. in the pursuit of construction projects or repairs. The containing system is an intention, hence the containing systems for artifacts are biological and psychological (cultural, etc.).

---

[55] Cummins (2002), 163.
[56] Cummins (2002), 164.
[57] Cummins (2002), 166.
[58] Cummins (1983), ch. 1.

Cummins does argue, though, that the human sciences such as psychology, sociology, economics, etc., are "special" in contrast to such universal sciences as physics, chemistry, etc. What enables prediction about how individuals acquire concepts or draw conclusions tends to be of little use in predicting chemical reactions. Special systems like persons or economies are still described in terms of their effects.[59] But these effects are not thought to be universal: they are restricted to systems of a particular composition. The putative laws of the special sciences are more like local patterns: they specify effects as descriptions of regularities of mechanisms in limited classes of systems. By contrast, there is no supposition that e.g. the laws of motion describe merely local effects.[60]

Biology and psychology do share, however, the explanatory strategy of using functional analysis to understand system capacities "as a kind of complex dispositional property."[61] Capacities are acausal specifications of input/output, i.e., conditions under which some state or property would be realized. (Cummins offers an example in which he defines the solubility of salt in terms of a subjunctive conditional which states that if salt were placed in water, it would dissolve.[62]) The functions of artifacts are to be explained in virtue of the capacities of psychological and perhaps social systems. Cognitive psychological explanation of the intentions associated with artifact functions would be carried out in terms of a computational functional analysis (1983, 2010).

---

[59] Cummins (2010), chapter 15, 287.
[60] Cummins (2010), 287.
[61] Cummins (2010), 288.
[62] Cummins (2010), 288.

§2.5.2 The Crux of the Problem?

The basic "how it works" approach Cummins adopts appears satisfactory enough for the explanatory domain that interests him, but does he go far enough in criticizing the teleological underpinnings of evolutionary science? He derides etiological theories, and chortles that even when all the other natural sciences have excised teleological thinking, it survives in evolutionary biology, "or anyways in the philosophy of it."[63] But even if his criticisms are ultimately correct, he never explicitly seeks to challenge adaptationism, the prevalence of which everyone acknowledges; he only ever indirectly challenges e.g. optimization models. Cummins does not come down very hard against the ordinary biology which unabashedly trades in adaptationist thinking. The philosophers take a beating while the actual biologists get off relatively unscathed.

In the case of artifacts, this generates a problem for Cummins already identified by Millikan and Preston: it rules out in advance any notion of malfunction. This certainly runs counter to adaptationist thinking and thus to normal evolutionary biology. The exclusive emphasis on how a system works leaves Cummins with no way to explain systems that do not work, and in fact no reason to identify them as systems. A component acquires a function only in virtue of the role it plays in the explanation of some effect exhibited by its containing system. It follows that if a system exhibits no such effect, then its components do not acquire functions.

The reasons Cummins resists including any notion of malfunction in his theory are both logical and philosophical. Logically, it seems to him that in order to ascribe a function to any component, the effect in which it plays a role must actually be observed (or observable in the sense that it is a genuine capacity). If the effect does not obtain, then

[63] Cummins (2002), 160.

logically the component did not malfunction, but rather it merely did not acquire a function at all. A deformed lump of tissue where a heart would normally be is not circulating blood and thus there is no circulatory effect to which it contributes, meaning that there is no such function attributable to that tissue mass. It is not a malfunctioning heart because it is not a heart at all. There can be no such effect as a malfunction, and the primary interest for Cummins is always how to explain the effects.

Philosophically, Cummins resists the notion of malfunction because of its normative connotations. Normativity, however, is precisely what Millikan (and Preston) is (are) hoping to explain. For Millikan, explaining normativity is the key to explaining intentionality and thus thought and language while also unseating the rationalism about meaning she perceives in much naturalist philosophy. Normativity, for Cummins, is not simply bad science but bad philosophy, bringing back much of what the theory of evolution by natural selection had done away with. However, his distaste for the normative—he does not really bother to give an argument—raises a difficulty in the case of artifact evolution. For the intentional phenomena that comprise the containing systems of artifacts and their functions and are studied by the special sciences are paradigms of normativity.

§2.5.3 Preston on Pluralism

Preston characterizes pluralism about function as the view that there are at least two kinds of functions: what she calls "system functions", which are Cummins functions, along with Millikan's proper functions. Millikan herself readily admits to being a "staunch pluralist" in Preston's sense.[64] Millikan considers the notion of Cummins func-

[64] Millikan (1999), 193.

tions to be quite "open-ended,"[65] but potentially useful in explaining unique traits and the contribution they make to a system. However, "the interesting Cummins functions associated with the mechanisms and traits of an organism will be exactly the same as the proper functions."[66] Her own theory, Millikan maintains, captures the only Cummins functions that would be of interest.

Pluralism is motivated by the need any adequate theory of function has to address two distinct but complementary aspects of function. System function is a highly flexible notion used to explain the current capacity of a trait or device in terms of the contribution it makes to the generation of an effect by a containing system. But attributing function to a component is just one aspect of explaining how a containing system works.

System function is, in this sense, quite separate from any historical account explaining what generated the system in the first place.[67] Exaptations—structures existing for reasons having nothing to do with the role they currently play as functional devices—best illustrate the divergence of system functions and proper functions. "Each of these notions of function [*system* and *proper*] has associated with it a proprietary mode of explanation. System function is associated with explanations taking the form of compositional analyses of the capacities of containing systems in terms of their component parts. Proper function is associated with explanations taking the form of causal-historical accounts of why a thing is there in the first place."[68]

[65] Millikan (2002), 139.
[66] Millikan (2002), 139.
[67] Preston (1998), 220.
[68] Preston (1998), 225.

Preston alleges (contra Millikan) that Millikan rejects pluralism, and she aims to show this through a discussion of exaptation.[69] Because exaptive functions call for the functional-analytic mode of explanation, Preston posits that Millikan's hostility toward exaptation (as an alleged counter to her own theory) drives her to refine the notion of selection she uses to define proper function. Actually there are two refinements according to Preston: the first a reminder that selection differentially maintains rather than builds traits; the second an expansion of the definition of derived proper function.[70] The expanded notion of proper function, which now covers derived and exapted devices, employs weaker criteria. On Millikan's expanded definition of proper function, to possess a derived proper function, it is enough that a feature or structure be merely utilized by a device.[71] No longer is it necessary to generate the feature in question.

While this does cover exaptations, Preston argues that it commits Millikan to a far more bloated sphere of derived proper functions that damages her theory. It leads her to equivocate on the meaning of selection: "the introduction of derived and expanded proper functions means that proper function in general both does and does not essentially involve a selection history in the primary biological sense, and consequently it both is and is not normative."[72] Things will turn out to have proper functions Millikan would not want them to have. Preston's example is noses acquiring the derived proper function to support eyeglasses: noses will acquire this derived proper function because it is the *direct* proper function of the glasses to correct vision, which they do by utilizing noses.

---

[69] Preston (1998), 226.
[70] Preston (1998), 229.
[71] Preston (1998), 232.
[72] Preston (1998), 234.

The basic difficulty is that Millikan tries to expand the notion of proper function in a way that undermines the normativity which appears to be the chief virtue of her theory, yet if she does not expand it in this way, it leaves out not only exaptations but also most of the derived proper functions of artifacts. The kind of normativity Millikan postulates is achieved in virtue of the reproductive ancestry of a given device token. But exapted traits lack such histories; their functions are acquired in virtue of current capacities or dispositions in a system, i.e. Cummins functions.

Preston argues that the same is true for artifacts. Artifacts have functions, but they are system functions (Cummins functions) attributed in virtue of dispositions or features, and lack the same kind of normativity as biological or language devices with proper functions. They have ad hoc functions rather than proper functions.[73] Individuals find all sorts of novel uses for extant items. Once those novel uses diffuse through a population of users and begin to redirect reproduction, an exaptation has occurred. Plastic bottles manufactured not to contain beverages or gases, but for service as bird feeders are an example of an exapted artifact. The functions of artifacts are best construed as Cummins functions, and they are much more commonly exapted, so to the extent that Millikan fails to account for exaptations, she fails to account for artifacts.

Thus according to Preston, Millikan has backed herself into a corner and faces the following dilemma: either she can preserve her expansion such that the attribution of proper function becomes promiscuous and violates the historical normativity of her own theory, or she can reject the expansion and be forced to accept that exaptations, and hence

[73] Preston (1998), pp. 237-238.

most artifactual functions, fall outside the scope of her (original) theory. Preston urges Millikan to renege on her expansion and affirm pluralism about function.[74]

§2.5.4 Millikan's Replies

Ultimately what is at stake in these criticisms of Millikan's theory is a correct account of the role of selection. Her replies to both Preston and Cummins make this clear, and shed light in particular on how selection operates in the artifact case.

Millikan resists Preston's attempt to "assimilate" Cummins functions to exaptations, arguing that while "adaptation" and "exaptation" are exclusive terms by definition, some adaptations have Cummins functions. Hence there must be more to what Preston calls "system functions" than just exaptations.[75] Cummins functions have nothing to do with what distinguishes adaptations from exaptations, so Millikan can remain a pluralist about function regardless of where she stands on exaptation. On Millikan's theory, functional analysis still serves its same useful purpose.

The real problem with functional analysis is that the Cummins functions it assigns relative to analyzed containing systems are too liberal to do the kind of biological heavy lifting Millikan wants for her theory of intentionality. According to Millikan, "what counts as part of the 'species-maintaining' Cummins system…has no determinate answer."[76] Functional analysis presumes some other method of delineating the target systems we want to study, e.g. biological systems, cognitive capacities, systems of material culture or artifacts, planetary systems, etc. The method for describing a target is quite ob-

[74] Preston (1998), 239.
[75] Millikan (1999), 193.
[76] Millikan (1999), 196.

vious in the case of artifacts: the actual designs are often on hand. No such blueprints or flow charts are available in advance for biological systems.

How do we specify what counts as a containing system without some prior method of identifying the traits to analyze? Millikan continues, arguing that "because there is no such thing as *the* Cummins functions associated with maintenance of a species, there is no such thing as *the* exaptations associated with it…The only way to place nonarbitrary limits on what counts as part of a biological system is to bound it with adaptations, with proper functions."[77] Millikan agrees with Preston that it would be better not to treat her own "expanded" functions as proper functions in order to cover the exapted ones, and she insists she has not done so; Preston has misread her. What Millikan characterizes as expanded functions are not thereby proper functions.[78]

In the specific case of artifacts, Millikan believes Preston has placed entirely too much emphasis on users as sources of the derived proper functions of artifacts. Millikan replies that "artifacts have as derived proper functions the functions intended for them by their makers" and not by their users.[79] Artifacts typically acquire their derived proper functions from the intentions which generate them. These intentions, in turn, possess their own derived proper functions derived from types (inherited) and tokens (individual experiences) of neural and cognitive mechanisms. She cites a novel can opener as an example: "if I carefully design and make a new sort of implement for the purpose of opening cans, but nobody ever actually uses it, there is *no sense whatever* in which it is a can opener, or in which its function is to open cans."[80]

[77] Millikan (1999), 197.
[78] Millikan (1999), 197.
[79] Millikan (1999), 205.
[80] Millikan (1999), 205.

Recall here that Millikan originally suggested three ways in which artifacts can acquire proper functions: artifacts always acquire at least one proper function derived from the intention of the designer/creator; they often acquire *direct* proper functions when they are reproduced on account of the effect or use they generate; and sometimes they can acquire derived proper functions in a third way, derived from users and corresponding to a new use for the artifact that diverges from the purpose intended for it by its designer.[81] Taken together, Millikan's theory can account for all the shifts and conflicts in artifact proper function by distinguishing the sources of derivation.

This third way of acquiring proper function is the one Preston emphasizes, but she takes it to be the mark of derived proper functions in general, at least in the artifact case. But clearly Preston is wrong to suggest that artifacts that go unused would for that very reason lack derived proper functions. Unused artifacts always acquire derived proper functions just in virtue of being items generated at the intention of a designer. These proper functions are what maintain artifact "species" (i.e., types). When artifacts acquire proper functions in virtue of the actions of users rather than makers, e.g. someone decides to start using plastic soda containers as bird feeders, this can also be a source of derived proper function, but it is a secondary and perhaps less proximal one, and at any rate it may conflict with or represent a shift from the original one (i.e. the maker's). There is a deep issue here about the relation between use, function, and reproduction, which I examine more closely in chapter five.

In "Biofunctions: Two Paradigms" Millikan construes exaptations (and spandrels, the artifact equivalent) as varieties of Cummins functions.[82] She defines exaptations as

[81] Millikan (1984), 49.
[82] Millikan (2002), 114.

traits having Cummins functions but lacking proper functions. (The term is co-opted for the sole kind of Cummins function that does not correspond to a proper function.[83]) Her argument affirms her pluralism about function while reiterating why she thinks Preston is wrong in her claim about exapted traits. Millikan agrees with Cummins that functional analysis only serves to explain how a given biological system works. It does not explain how the system came about under selection, nor does it even describe the system. Further explanation of the origin of a system's traits, or of its historical performance, requires additional resources the notion of proper function was developed to provide.

Contrary to what Cummins concludes, the theory of proper functions is not an attempt to explain why traits exist as opposed to *how* traits come to exist. Millikan appeals to the notion of "selection *for*" to avoid Cummins's criticism that her theory gets selection wrong: "Any trait that was selected for performing its current function has historically *caused* an increase in fitness."[84] Thus proper functions are indeed useful in biology because they provide a way to appeal to historical conditions to identify and rank the various traits that describe a biological system. Proper functions bring history into the explanatory picture to give us a glimpse into the system that is only as arbitrary as our best hypotheses about that system's history. Cummins would have us demarcate systems on the basis of the effects which most interest us.

The basic disagreement seems to be that Millikan regards functional analysis as useful in the domain Cummins has carved out for it, but too general otherwise, while Cummins disputes that any further explanation of biological systems beyond describing how they currently operate requires an appeal to historical conditions. Millikan is fair to

[83] Millikan (2002), 139.
[84] Millikan (2002), 118.

suggest that we need something like a theory of proper functions to begin to describe functional systems not merely in virtue of their interest to us, but in virtue of the survival value they have actually had. She has essentially found a way to relate selection as a population-level phenomenon to the functional traits of token structures. Doing biology means talking about the reproductive, species-propagating fitness of organisms as well as their artifacts. This move not only responds in a satisfactory way to Cummins, but it blocks the objection Preston wants to raise about exaptation.

Furthermore, it suggests that the role of selection is unified across all the domains of interest to proponents of evolutionary explanation: psychology, language, biology, culture, design, etc. The inexorability thesis turns on the plausibility of this suggestion. In chapter four I return to this idea and examine whether it holds up to close scrutiny and, if not, whether other non-selectionist mechanisms of evolution are available.

Function and Action

In this chapter I present the most prominent alternative to historical theories of function: the use plan theory that originates with Dipert (1993). The main idea is that artifacts acquire functions in virtue of agent plans, subject to rational assessment. I explore Dipert's basic argument and his classification scheme. I then consider the elaborated version of this picture given by Houkes and Vermaas (2010). Their notion of ***use plan*** *clarifies the basic picture and makes a notable gain with respect to function attribution.*

The use plan theory of artifacts begins from action theory. It develops a particular understanding of action, and it attempts to incorporate artifacts into that understanding. It achieves this by tying the functions of artifacts to features of agents. In particular, the features which give rise to artifacts are the intentional plans of agents to use items to achieve goals. All this likely sounds obvious and intuitive, and it is important to see that this is so by design. The use plan theory tries to explain artifacts in a way that captures the common intuition that artifacts are derivative teleological entities. They are things made and shaped by the intentions of human beings.

§3.1 An Appeal to Action

"Our chief concern in this volume is understanding *attributions* of action-theoretic notions to agents through their behavior and especially their products."[85] Dipert pursues a fundamentally different investigation from Millikan. He regards the adoption of action theory in his approach to artifacts as a philosophical antidote to the naturalization project of e.g. theories like hers.[86] He denies outright that an item's material affordances or capacities determine its status as an artifact: "artifactuality does not consist in any present

[85] Dipert (1993), 57.
[86] Dipert (1995), 119.

physical qualities of a thing."[87] Contrast this with Millikan's theory, where the effects of an item's ancestors are the entire story. If an item's ancestors did not perform the effects they did on account of their affordances, that item would not be what it is.

But more importantly, the products agents make give clues to their actions and intentions. Dipert characterizes artifacts as a "residue" of intentional activity.[88] What intentionality artifacts display exhibits the intentions of agents. It manifests as an epistemic relation between designer, item, and user. In order for an item to count as an artifact, a user must be able to recognize the intended purpose or use of the item, as derived from its originating agent.

This action-theoretic approach also includes an emphasis on history, though it is history of a different sort than reproductive history. "An object is, or is not, an artifact in virtue of what its history was (which may be unrecoverable). Specifically, we might guess that an object is an artifact just when it was once intentionally modified, to some degree, by a human being or other finite agent."[89] What kinds of facts about the past does Dipert think determine an item's artifactual status? The epistemic relation between item and user must obtain through a process of practical reasoning. The facts must be psychologically intentional "in the sense that they are properties in virtue of the attitudes a cognitive agent has had toward the object…the entity [i.e. an artifact] must have once been conceived, intentionally used, or intentionally modified in some way."[90] This can be made clear by looking first at the classification scheme Dipert offers.

[87] Dipert (1993), 15.
[88] Dipert (1993), 15.
[89] Dipert (1993), 15.
[90] Dipert (1993), 42.

§3.2 Classifying Items as Instruments and Tools

The first division Dipert draws is between "fully natural" items and another category comprised of three subcategories of human-made items. Fully natural items are objects that have never been utilized by any agent to achieve any goal. In the other category are instruments, tools, and artifacts. "Tools are a subclass of instruments, and artifacts are a subclass of tools."[91] The broad category of "artificial entities" would also include such byproducts and unintended human-generated items as waste, scrap, side effects, etc., though Dipert never makes it entirely clear how these phenomena relate to his definitions or to the key role agent intentions and beliefs play in them.

Dipert gives us the following definition of *instrument*:

> An object O is an instrument with respect to property-set P for agent A and goal G just when:
> (i) O has properties P and is believed by A to have properties P,
> (ii) properties P are means to attaining G and are believed by A to be means to attaining G, and
> (iii) agent A has used O intentionally in order to achieve G, that is, has used A (at least in part) because of the belief that O (O's properties) were a means to G, and because A had G as a goal.[92]

This definition stipulates that in order to count as an instrument, an item O must have actually been used to achieve an intended goal. Furthermore, the agent who uses the item must believe that an efficacious way to achieve the goal in mind would be to use this item in respect of its properties.

Notice also that the only beliefs referred to in this definition are about the suitability of the properties of some natural item. No reference is made here to the beliefs or mental states of other agents. The instrumentality inherent in this definition could lend itself to formalization using material conditionals, though Dipert also mentions other

[91] Dipert (1995), 124.
[92] Dipert (1995), 121.

kinds of logical structures that might be suitable. Found or unmodified useful items would include stones used for striking or pounding, wild fruit and vegetables used as food, animal instruments such as the shells used by hermit crabs or the twigs crows use to retrieve food, and even the example from Cummins of the glacially depressed rock bowl used for holy water.[93]

The key criterion for instruments is rather intuitive: the user may not materially modify the item in the course of putting it to use; spatiotemporal adjustment or relocation does not count as a modification. While there are borderline examples, this actually covers a wide variety of items used from early pre-history down to today. (A borderline example: the material items used as fuel to create and maintain fire. These items may require no modification to produce a blaze, but it is not clear whether the resulting blaze which consumes the materials counts as "modifying" them. Similarly for food.)

Next up is the definition of a *tool*:

> An object O is a tool with respect to property-set P for the agent A1 and goal G just when
> (i) object O has properties P and is believed by A1 to have properties P,
> (ii) a (possibly unspecified, often different) agent A2 intentionally modified (or deliberately left alone) all of the properties in P in order better to achieve G, and this is believed to be so by A1, and
> (iii) agent A1 intentionally used object O because of beliefs about the intentionally-increased efficacy produced by A2 of O through properties P.[94]

This definition stipulates that an item must generate, on the part of a user, specific beliefs about the intentions of another agent besides the user: "It requires mental states *about* mental states."[95] So far the particular contents of the beliefs about the second agent are not very clear, aside from their pertaining to the efficacy of the item for a given goal. It is already obvious that these definitions are quite different from the ordinary colloquial

[93] Cummins (1975), 747.
[94] Dipert (1995), 123.
[95] Dipert (1995), 123.

notions of *instrument* and *tool*, being seemingly broader or perhaps a gradation removed. All we need to keep in mind for now is that the main difference between instruments and tools "is that instruments are considered with respect to their *usefulness*, while tools are considered (also) with respect to their having been *intentionally modified* for enhancing this usefulness."[96] Recognizing an item as a tool involves recognizing that, apart from having some apprehensible utility, it has properties which indicate that some other agent has in fact modified it with that utility foremost in mind.

(This can also include the recognition of deliberate inactions. Deliberately leaving natural items as they have been found can furnish them with tool-properties: "The left-alone properties achieve their desired effect against a background of intentionally modified properties."[97] This is the case with found art, but will not much concern us here.)

§3.3 Artifacts and Agent Plans

The basic idea behind Dipert's view of an artifact is that "artifacts *intentionally signal* (and more exactly, communicate) that they are tools."[98] They are items that wear their tool-purposes on their sleeves, as it were. In addition to their properties of being useful in the ways tools are, they also have features which telegraph or display that utility to users. They have additional modifications that are not enhancements strictly of function or utility, but rather of communicativeness. They advertise themselves.

[96] Dipert (1995), 124.
[97] Dipert (1995), 125.
[98] Dipert (1995), 127.

§3.3.1 Artifacts as Communicators

"An artifact is an object that possesses some self-communicative properties—that is, has properties intended to cause belief-like states, in certain ways, *about the object itself*—and specifically has properties that bring us to believe that the object has tool properties."[99] The properties or features of the item which communicate its enhanced function or modification as a tool can and often do coincide with the enhancements themselves, but this sort of case obscures rather than exhibits the aim of the definition Dipert wants to construct. The key point is that the features doing the communicating do not further enhance the item's functionality. Looking a little more at how the communication is supposed to work will help make this clear.

The beliefs and mental states of agents can be affected in a variety of ways. One such way is through signaling. Signals can be linguistic, perceptual, affective, cognitive, and so on. Artifacts are items that signal an agent's intention to another agent. The communicative features of artifacts "produce changes in behavior or mental states or processes by the intentional production of intermediate belief-like states in the target cognitive agent, which in turn produce the intended behavior, state or process (if it is not the belief itself)."[100] This passage is somewhat obscure, but roughly it means that the item will cause a dispositional belief to become occurrent. It will activate (or perhaps generate) a belief in an agent that an item is good for an appropriate use toward some particular end.

Dipert contrasts two traffic control measures as an example. A stop sign will cause an agent who is driving to bring his or her moving vehicle to a standstill. It does so

[99] Dipert (1995), 128.
[100] Dipert (1995), 127.

by getting the driver to believe that some other agent intended for drivers to stop at the sign, and that this other agent was an appropriate authority (e.g., local government). By contrast, rumble strips in the pavement can achieve a similar effect—the driver stops the vehicle upon recognizing an intention on the part of some other agent to stop drivers in a given spot—but it does so in a different way. The difference is that rumble strips do not rely on the belief-forming mechanisms of drivers to cause the action of stopping. "Although they are intentionally put there to stop us or slow us down, and we may in fact appreciate this…they just get us to stop or slow down by making it in our interest to do so, or merely startling us into slowing down. They are analogous to a barrier that actually stops a car; this is not a 'communication' that we should *believe* we should stop the car, it is just getting us to stop the car, somehow."[101]

It might seem that the distinction alluded to here suggests that artifacts must rely on language to communicate their intended uses. This is not so. While language (e.g., written instructions) is arguably the most effective and reliable method of causing agents to adopt beliefs about what courses of action they should take, there is nothing special or privileged about it. Artworks can achieve the same effect, and they tend to be much more visual than linguistic in their roles as epistemic conveyances. Language happens to be a reliably precise method of conveying intended use, and hence the designers of artifacts commonly employ it as a method of communication. But nothing about the view Dipert develops requires that the communication be linguistic.

Let's look at his definition of *artifact* to see if this can be made clearer:

> An object O is an artifact with respect to property-set P1 for agent A1 and goal G just when:

[101] Dipert (1995), pp. 127-128.

> (i, the TOOL CONDITION) O is a tool for Al with respect to properties P2 (intentionally modified by A2) and goal G and
> (ii, the TOOL-COMMUNICATING CONDITION) O has properties P1, and is believed by Al to have properties P1, and Al believes that A2 has added (or deliberately left alone) the properties P1 in order that an agent, in which category Al falls, come to believe that O is a tool for an agent (in which category Al falls) with respect to properties P2 and goal G, and
> (iii, the COMMUNICATIVE SUCCESS CONDITION) Al comes, in a certain way, to believe P2 are tool properties and that P1 intentionally communicates this, in virtue of the apprehension of properties P1 in a certain way.[102]

For an item to be an artifact, it must (i) be a tool, (ii) have certain features which indicate that some agent (typically) other than the user intended a use for it in pursuit of some goal, and (iii) the user must actually come to apprehend this intended use in virtue of those features. This communicative criterion (iii) is Dipert's central innovation.

The argument he offers for (iii) ultimately comes down to the point that artifacts ought to be recognizable for what they are intended to be. There are so many different kinds of tools, and purposes for which to use them, that their actual efficacy for a given use is liable to be lost on a potential user who is not familiar with the tool or the use in question. I do not know much about stereo equipment, e.g., the purposes of the various components, wires, and attachments. New designs for portable stereos do not look much like they did even ten years ago. They have no slots for inserting compact discs, no tape decks, no analog buttons or controls, and often no visible speakers. If it were not for packaging or other signals, I might not be able to tell that I was looking at a stereo. For this reason, the criterion for artifactuality ought to be the communicative effectiveness of the item regarding its intended use. Artifacts must advertise.

This communicative aspect of artifacts depends on features of those agents who would consider using the item in question. "An artifact is an intentionally modified tool whose modified properties were intended by the agent to be recognized by [another]

[102] Dipert (1995), 129.

agent at a later time as having been intentionally altered for that, or some other, use."[103] So according to Dipert, an artifact wears its purposes on its sleeve: it must exhibit its intended purpose in such a way that it could cause another agent to believe it to be a suitable item for achieving the end in question. Often this can be done through written instruction, and it need not generate a *specific* belief in every other agent. Suppose the designers of a nuclear facility intend for some containment equipment to generate in workers the belief that the item can be used to transport radioactive waste. They rely on the expertise of the waste handlers to recognize what the item is for, but not every agent who might encounter the item will possess such expertise. However, it is enough that other agents recognize that the item has some intended purpose, even if they can't tell just what that purpose is. (I take this to be a charitable construal of Dipert's definition.)

Adam Morton gives an interesting example of an environment in which the artifactual status of some items can fluctuate depending on the cognitive expertise of agents. Imagine a home where obscure ultra-modern décor furnishings (utensils, appliances, spots for sitting and sleeping, etc.) exhibit a counter-intuitive design appearance.[104] The intended use of a given artifact in this environment must be inferred. The inference can be justified by the relation of the item to other artifacts in the environment, or by methods of plain old practical reasoning like trial and error. In these ways I can determine what some items are for, but others remain obscure. The corkscrew, for instance, proves hopelessly elusive. Presumably I possess all the user expertise that is required to recognize a corkscrew (e.g., I have seen and used them many times before). Assuming there is an item in the kitchen (or perhaps in another room) that was intended for this use, does it

---

[103] Dipert (1993), pages 29-30.
[104] Morton (2006).

count as a corkscrew? Does it count as one only in the presence of an agent versed in the nuances of the décor? This bar seems unreasonably high for a common utensil. But maybe it still counts as an artifact even though it does not do a good job wearing its function on its sleeve. Or perhaps the item (wherever it is) should not count as an artifact on Dipert's view, but could still be considered a tool (so long as it does dislodge corks).

Dipert's way out of this predicament is to impose thresholds of intentional recognition for users: low-, middle-, and high-level intentions. "High-level intentions include the beliefs, emotions, behavior, and so on that the acting agent [i.e., a designer] wishes to cause in (other) cognitive agents [i.e., in users]."[105] At this level, the artifact generates a specific bit of procedural knowledge in a user, complete with a goal and a plan of action to achieve that goal. This is the level at which the nuclear waste handlers or aficionados of modern design would be operating.

"Middle-level" intentions deliver only on the goal, suggesting no plan or procedural knowledge. And minimally, "low-level" intentions are communicated merely in virtue of raw sensation: by sight, touch, etc. a user can tell that an item is for something, but even the goal is not clear. Typically an artifact delivers on all three levels at once for an intended user. But naturally, other users are less prepared. Morton's elusive corkscrew, if found, might not meet anything above the lowest level: it might be clear to a non-aficionado that it has some function (i.e., it is not purely decorative), but just what that function is cannot be ascertained. Perhaps reasoning by elimination or some other process, s/he can infer that it is a corkscrew, and hence ascend to middle-level recognition, yet still lack any actual idea *how* to use the item to dislodge corks. But if the agent

[105] Dipert (1993), pp. 54-56.

goes online, reads up on the style of design, and learns how to operate the item, s/he would attain high-level recognition by acquiring a specific plan intended by the creator.

The basic point is that agency is what makes objects artifactual or not. Agents are logically prior to artifacts because intentional action supplies the goals and plans in virtue of which objects can be recognized as deriving some purpose.[106] Intuitively, a plan is a chain of practical reasoning. It posits a goal or practical end—i.e. a purpose—that is realized by following a procedure. We can expand on this by looking at agent plans.

§3.3.2 Agent Plans

In elaborating on the attribution of plans to agents, Dipert lists three features that comprise an ideal understanding of an artifact:

> (1) we have a sufficiently full description of the creating agent to enable us to identify him or her uniquely among all present and past (actual) agents (i.e., a 'distinct' idea in the language of Descartes and Leibniz) and to enable us to interpret all culture- and context-dependent artifactual features; (2) we grasp all the steps of practical reasoning (including prior habit formation) that went into all the agent's actions upon the object; and (3) we grasp the history of the object's artifactual aspects (i.e., the loss or reconstruction of these features since its creation).[107]

"Understanding" an artifact in Dipert's sense could be compared to discovering the proper function of an item, as in Millikan. But here the work is interpretation rather than discovery. It is an epistemic task that investigates the deliberative history of the item. "By 'interpret' I mean coming somehow to regard or believe that something is an instrument, tool, or artifact together with coming to have assumptions or beliefs about the

[106] Dipert (1993), 61.
[107] Dipert (1993), 80.

deliberative history of that object."[108] Asking after an item's deliberative history helps show us how to think about using that item.

Attributing functions to artifacts means attributing plans to agents, as well as agency itself, and both of these under a causalist conception of action. Any arbitrary item could, in principle, serve the role of "hammer" in an agent's plan to drive nails. But obviously there are material limits to the efficacy of a large class of items, such as leaves, or rubber shoe soles, or any tool, instrument, or natural object that is not sufficiently hard, heavy, or flat. What determines whether an item fails or succeeds as an artifact in the plan of an agent is a set of rational standards. Rational evaluation must take into account such features as an item's material suitability to achieve a goal.

Dipert emphasizes the importance, even the indispensability, of the assumption that we really can come to better understand the intentions of other agents through their acts, and especially through their communicative acts such as we find in the creation and use of artifacts. Apprehending artifacts requires "both the attribution of meanings and intentions to others, as well as the possibility of their intelligibility to us," though Dipert mildly laments how this practice resembles transcendental argument.[109] The justification for this practice is not transcendental, however, but ultimately instrumental: "We must realize…that our success will not be in the pinpointing of precise beliefs or precise goals but at most in admitting or eliminating classes of beliefs and goals, and then only in those circumstances that are useful for the prediction of behavior."[110]

In other words, things count as artifacts when the attribution of function and artifactual status meets some set of normative standards. There are limits on the rational in-

[108] Dipert (1993), 77.
[109] Dipert (1993), 80.
[110] Dipert (1993), 86.

ferences users are permitted to draw about the proper or intended use of an item in making their deliberations. Even our modern home with its confusing décor is unlikely to contain five can openers; if there are no other candidates, surely one of them must really be the elusive corkscrew. Inferences by users to prescribed plans on behalf of creator-agents are governed by norms, particularly rational and epistemic norms but also moral norms. A set of kitchen knives could suggest homicide to a deranged mind, but a general moral prohibition on murder is evidence against it as a plan coinciding with intended use. In the next section, we look at plans in more detail.

§3.4 A Mature Use Plan Theory

A number of authors have recently sought to elaborate on Dipert's action theory approach to artifacts and function. Like Dipert, Wybo Houkes and Pieter E. Vermaas emphasize plans and agent teleology. They develop a much fuller theory of plans in order to account for both the use and design of technological artifacts. Beginning with a recognition that historical theories of function are not equipped to handle what they call the problem of alternative use, they turn to action theory. They develop a normatively robust technical notion referred to as a *use plan*. It also yields an intentionalist theory of artifact function that complements research in cognitive psychology.

§3.4.1 The Problem of Alternative Use

Houkes and Vermaas regard etiological theories of artifacts as something that comes as a sort of afterthought in the course of explaining organisms. Millikan's attempt to "naturalize" intentionality does not, in their view, capture the intentions that generate artifacts. The difference intention makes is that in the domain of artifacts, genuine novelty is possible, even common. They argue that etiological theories "cannot handle cases of innovation."[111] Biological devices inherit their proper functions either directly or indirectly. But artifacts seemingly can acquire them ex nihilo: devices with genuinely novel functions often appear on the scene with no discernible ancestors whatsoever.

Perlman points out that raising this concern "ignores the entire debate in Millikan and Cummins and others over novel organisms" (and Swampman in philosophy of mind).[112] Indeed, Houkes and Vermaas seem to be trading on an equivocation between ordinary function and proper function: Millikan is quite explicit that proper function rules out cases of novel, uninherited capacities in naturalizing biological function. It may be true that some new device has a particular capacity to perform a task but, like biological devices, the first device in any lineage does not have that capacity as a proper function. Reproduction and heredity are essential to the account of proper function we find in Millikan. (Houkes and Vermaas do seem to acknowledge this reply.[113])

However, Houkes and Vermaas develop another criticism. They argue that etiology fails to explain alternative use, an important factor in artifact reproduction and function attribution. Often "the distinction between standard and alternative use is rephrased as one between use that is and use that is not in accordance with an artifact function."[114] Relating artifact use to function leaves us with little means to distinguish the types of use it seems we must distinguish in the artifact domain. Some alternative uses, for example, are improper while others are acceptable: chairs make acceptable stools but improper pa-

[111] Vermaas and Houkes (2003), 277.
[112] Perlman (2004), 32.
[113] Vermaas and Houkes (2003), 278.
[114] Houkes and Vermaas (2004), 54.

perweights.[115] Contra proper function, standard uses are not historically fixed, according to Houkes and Vermaas: they constantly shift about.[116] The upshot is that what Houkes and Vermaas call the "dynamics of use" cannot, in their view, be adequately described in terms of proper and accidental functions alone, yet these ideas are the lone currency of historical theories. While the proper function account captures norms that determine malfunction, etc., it fails to capture others that set limits on (socially) acceptable use.

§3.4.2 Use Plans

The mere use of any item seems thoroughly goal-oriented. If I idly twirl a pencil while conversing with an office mate, it seems odd to say I am in fact using the pencil at all. "Using is using with an aim or goal."[117] Am I performing an action? Is an idle reflex or habit an action? The use of an artifact turns out to be a surprisingly complex phenomenon. Humans are raised to be proficient at using artifacts very early on. This hides the true level of skill involved in even the simplest tasks. To properly explain such use in terms of action, goals must be posited, knowledge of an item's capacities must be attributed, a stepwise sequence must be followed. I am not using my umbrella unless I first remove it from the stand, unbuckle and unsheathe it, step outside, extend it, and hold it at such an angle during a downpour that it actually serves to deflect the rain. The coordination involved is hidden partly by how mundane and routine it all seems.

Action theory offers the notion of a *plan* to help order all the steps that combine to form an intentional action. A plan is a complex mental state consisting of considered

[115] Houkes and Vermaas (2004), 55.
[116] Houkes and Vermaas (2004), 55.
[117] Houkes and Vermaas (2010), 16.

rather than actual actions.[118] Houkes and Vermaas qualify this notion in a mildly technical manner: a *use plan* is a plan that also considers the manipulation of an item in order to achieve its goal. Using an artifact, therefore, just is the carrying out or *execution* of a use plan. Use plans are mental states; executions thereof are actions involving the manipulation of items. They refer to the mental process of creating use plans as *design*.

Houkes and Vermaas address two limitations of plans: (1) that plans cannot specify in advance each detail of each action they include, and (2) that it is not practical to specify such details. To address (1), they appeal to knowledge-how or what they call skills, i.e. practical or procedural knowledge in the anti-intellectualist style. There is no need to specify in advance an action required to execute a use plan if the acting agent can rely on skill in executing that stage of the use plan. The response to (2) is that it is often counterproductive to specify rigid plans: "the less detailed plan is more flexible and more responsive to the specific, possibly changing situation in which it is realised."[119] In general, plans and use plans arise in the context of an agent's "belief base" or background knowledge about his or her own abilities, an artifact's capacities, an environment's challenges, the use plan's gaps, and the appropriateness of executing a plan given all of these factors. An agent's belief base figures also in the evaluation of use plans.

Use plans often need to be spawned in the minds of other agents. This is called design: "designing is primarily—sometimes even exclusively—constructing and communicating use plans."[120] Hence Houkes and Vermaas conceive using and designing as connected: "designers support users by constructing new plans for attaining new or existing

[118] Houkes and Vermaas (2010), 18.
[119] Houkes and Vermaas (2010), 25.
[120] Houkes and Vermaas (2010), 26.

goals."[121] An agent who has a goal but lacks a use plan to achieve it must either come up with a use plan on his or her own or, failing that, acquire one from a designer. The design of a use plan is basically similar to actually using one, except that there is an additional element: communicating the use plan to a user. The goal of design is to provide a use plan that enables a user to achieve a goal. I will not now go into the account of design in further detail, focusing instead on the communication of use plans to users.

§3.4.3 Communicating Use Plans

Artifacts, recall from Dipert, wear their intended uses on their sleeves. They must be capable of being recognized as items with which to achieve a goal. A use plan is in principle communicable. Plans are composed of explicit ordered steps that, when taken, are intended to achieve a goal; these steps can be communicated verbally, visually, etc.

A use plan simply incorporates some item in the pursuit of that goal. So the steps of a use plan ought to be communicable in just the same ways. And indeed written manuals are a common form of use plan communication, as are spoken instructions, visual demonstrations, and other practical lessons. On the accounts of both Dipert and Houkes and Vermaas, artifacts are always the product of a design process since their intended uses are always supposed to be recognized by or communicated to users.

Just what is communicated, however, differs in Houkes and Vermaas from what we saw in Dipert's theory: "although we share an emphasis on communication with Dipert, the content of this communication is the use plan and not the fact that the artefact

[121] Houkes and Vermaas (2010), 28.

has, somehow, been physically changed to enable or facilitate use."[122] So the criterion for an item to be an artifact here is that a specific use plan be communicated. For Dipert, it was enough that an intended use be conveyed. Houkes and Vermaas argue that there should be more: they seem to want a rather full-fledged use plan displayed.

I think the main difference here is in specifying a goal on the user's behalf. In Dipert, the artifact must communicate a clear use but the goal, for which the item is put to that use, could vary from user to user, or between user and designer. Here the goal is part of the use plan, and it is the use plan which an item must communicate to be an artifact.

But in Houkes and Vermaas, the exact means of communication is not clear. The references are to "transferring" use plans from designers to users, with the implication usually being that this transfer occurs in language, or through visual demonstration. A written manual seems to be the default means of communicating use plans. This is fair enough, but it is a shame that the flexibility of Dipert's signaling is not emphasized or even retained. There an item's artifactual status did not depend on whether a *user* really had the goal or adopted the designer's intended use. It was enough that the intended use be made evident to the user, or to some other agent besides the designer. Houkes and Vermaas propose a rather more stringent criterion for artifactuality than Dipert.

Partly the less flexible criterion is methodological: Houkes and Vermaas want to give a unified account of artifacts in terms of use plans, of use plans in terms of action and intention, of design and use, and of the norms governing each. The notion of a use plan does more overall work if tied to agent goals on each side of the user/designer divide. The communicative criterion perhaps should be relaxed if it yields a more efficient analysis. It has, in fact, yielded at least one elegant formalization of the notion of use plan

[122] Houkes and Vermaas (2010), 155.

(Hughes 2009). Let us turn to the normative aspect of the account—the standards governing use plans—to see whether the sacrifice is worthwhile.

§3.4.4 Evaluative Standards for Use Plans

This section began by citing the problem of alternative use. I mentioned norms governing use, e.g. rational standards and proper/improper use. Elaborating on this aspect of the action theory approach is perhaps the chief virtue of the use plan account. Houkes and Vermaas take care to erect an evaluative framework for assessing use plans. "The cornerstone of the evaluative framework is that use and design can be assessed on the basis of the quality of the plan that is executed or constructed, relative to the circumstances in which it is executed or constructed."[123]

Use plans exhibit *practical rationality* to the extent that they effectively realize target goals. But their effectiveness is an intrinsically relative standard. It is relative to the specificity of the goal, to the beliefs of the agent regarding either the use plan's prospects or its results, to the efficacy of the item used to achieve the goal, and possibly to other features of the environment in which the item is deployed.

Houkes and Vermaas also cite other standards pertinent to determining the rationality of a use plan: goal consistency, means-end consistency, and belief consistency. The standard of goal consistency is important because most use plans include not only subsidiary and provisional goals, but often multiple main goals. These goals must not conflict with each other in obvious ways. For instance, if an agent's goal is to arrive at work on time using her car, but she also knows that she must stop to service the car or else it

[123] Houkes and Vermaas (2010), 37.

will stall, her goals are inconsistent. She cannot both stop to service the car and use it to arrive on time for work. Goal consistency is a fairly minimal standard, intended to apply to not uncommon cases where we realize that some provisional task on the way to a main task cannot be accomplished, thus jeopardizing the accomplishment of the main task.

Another basic rational standard is means-end consistency. "A plan is means-end consistent if and only if the agent executing it reasonably believes that both the artefact and auxiliary items are available to her…a plan is means-end-inconsistent if the user knows that auxiliary items are unavailable."[124] Say I want to finish my laundry, but I know that the detergent is already used up. I lack the means to finish the laundry, and must formulate a new plan that involves a subsidiary goal of obtaining more detergent.

Belief consistency refers to the requirement that a use plan be based on justifiable or reasonable beliefs "about the world, ourselves, and the effects of [our] actions."[125] The norms of justification are not explicated, but they seem to refer to common sense and similar publicly available standards like folk physics, basic logic, etc. A use plan to ride my bicycle across the Atlantic ocean is not belief-consistent because the belief that a bicycle can propel one over or under the water is not justifiable. It is not justifiable due to common-sense facts about oceans and bicycles. This appeal to public standards of epistemic appraisal to inform the evaluation of use plans is intuitive, if a bit simplistic.

These rational standards yield further norms of use for artifacts. Since use just is the execution of a use plan, the use of an artifact is rational if and only if it executes a rational use plan.[126] It is rational to use a tablet computer as a paperweight, provided that there are no other available items suited to holding down paper, and that I believe the tab-

---

[124] Houkes and Vermaas (2010), 40.
[125] Houkes and Vermaas (2010), 40.
[126] Houkes and Vermaas (2010), 41.

let would effectively realize this goal, and that so using it would not interfere with some other goal of mine (e.g., I want something to hold the paper down so that I can leave the paper be while I use my tablet to go online). The rationality of use and use plans is fairly straightforward on this score.

However, the appraisal of artifact use as proper or improper is a somewhat different matter. This standard is characterized as independent of rationality: "artefact use is proper if and only if it is the execution of a use plan that is deemed acceptable within a certain community. This acceptance or privileging of a use plan may have nothing to do with rationality."[127] The claims seems to be that certain use plans are socially sanctioned and that this can also govern the use of the artifacts featured in those plans. In chapter five, we'll see some further support for this in Preston's argument that proper functions actually impose social constraints on how agents use items of material culture.

This turns out to be a rather arbitrary basis for assessing the use of artifacts. Houkes and Vermaas give the example of prisoners using bedsheets as a rope to escape their cells. This use is quite rational but also improper because it breaks the law. Presumably, then, *any* item used to escape prison is improperly used, e.g. a ladder or a key to the front door. The examples the authors give are not very illuminating. Is it that all non-sanctioned uses of an artifact are improper uses relative to a social context? Or are there specifically prohibited uses that are the improper uses, while others are simply neutral?

The aim in articulating a proper/improper standard of use seems to be that Houkes and Vermaas want to hold designers accountable to communities for the use plans they create and the artifacts they produce and make available for use. This is perhaps an admirable moral goal, and it deserves further discussion. But it also stands gravely in need of

[127] Houkes and Vermaas (2010), 41.

clarification. As I just mentioned, in chapter five we'll see more about how proper functions serve to standardize patterns of use. This basic standard of rationality, on the other hand, is a valuable general tool to assess the viability of use plans and agent uses.

§3.4.5 Function Attribution

Houkes and Vermaas ultimately propose combining elements of each of the major theories of function into one theory: the "ICE-function" theory. ICE refers to intention (I), to Cummins-style causal-role or system function (C), and to the "evolutionist" or etiological theory of function (E). They argue that functions can be ascribed relative to use plans, and that all of the existing theories fail individually to meet the desiderata of a theory of artifact functions which Houkes and Vermaas have established.

I will simply mention and not describe at length the four desiderata: (1) that artifacts should be able to have both proper and accidental functions; (2) that the concept of proper function must allow for malfunction; (3) that some (scientific) support exists for ascribing a function to an artifact, be it dysfunctional (or not); and (4) that we be able to ascribe "intuitively correct" functions to innovative or novel artifacts. These desiderata are intended to capture corresponding phenomena about the artifact domain: (1) artifacts are used with extreme versatility; (2) they can fail to succeed; (3) they are constrained by material conditions; (4) and they can be novel creations.

The ICE-function theory includes elements of each of the three major theories of function, plus a proviso about communication chains. The purpose of this added element is to secure a means by which a function attribution is transferred from agent to agent along the course of an item's production and transfer. Unfortunately, this element of the theory still does not clarify the nature of the transfer, the constraints (if any) on the method of communicating a use plan, or what determines successful communication of a use plan other than realization of a goal state that approximates the user's goal. All that Houkes and Vermaas say is that agent testimony plays a large role.[128]

The basic stroke of the theory is to establish a justificatory threshold for the adoption of a belief about the contribution an item with a given material capacity will make to a given use plan. "If a designer, justifier or passive user ascribes to an artefact a function to φ, he believes that the artefact has a physicochemical capacity to φ and that this capacity contributes to the effectiveness of the use plan for the artefact."[129] The *justification* for the belief that an item is suited to some use plan will come down to further beliefs about the material capacities of the item, its physical constitution and dispositions, its specific composition, etc. But once the beliefs about the item's capacities are justified, it is just a matter of judging the effectiveness of those capacities for the use plan.

This theory appears to depart from the action-theoretic approach inaugurated by Dipert. After all, the focus now seems to be on material capacities of items rather than the intentions and actions of agents. But this shift in emphasis is merely to clarify the justification of an agent's (designer's, user's) belief that some item is suitable for a purpose it either communicates or is otherwise attributed to it. The structure of the theory does not differ substantially from Dipert's basic approach.

What is improved here is the basis on which an agent judges an item to be appropriate to a use in a use plan. Users no longer rely solely on their own epistemic muster to judge that an artifact will be effective. Users can now rely on the communication chain to

---

[128] Houkes and Vermaas (2010), 83.
[129] Houkes and Vermaas (2010), 88.

justify the belief that an artifact has the function it is attributed, since the chain represents a set of adjudications that it does, and not just the intention of a designer. Houkes and Vermaas have managed to replace the communication of an intention with a more robust attestation that an item really does have a given function, ideally according to scientific and professional sources with relevant social standing. This is a clear gain.

The Reproduction and Evolution of Material Culture

In this chapter I consider TD's inexorability thesis by investigating evolutionary models of reproduction of material culture. Does technology change occur by a process of natural selection? I first present Beth Preston's (2013) attack on standard intentionalist theories of function in material cultural reproduction. In §4.2 I outline a major objection to evolutionary models of technological innovation, namely the transmission problem identified in Lewens (2004). Then in §4.3 I confront this objection from two standpoints, the philosophy of memetics and the dual inheritance model of cultural evolution. I concur that memes are probably not adequate to defeat the objection, but the dual inheritance model is found to be more plausible than Lewens allows.

§4.1 Preston on Reproduction in Material Culture

A recent landmark in the cross-disciplinary study of technology, agency, design, and innovation is Beth Preston's *A Philosophy of Material Culture: Action, Function, and Mind* (2013). It addresses theories of function in philosophy of biology and theories of action and agency in philosophy of mind, utilizing aspects of each as well as empirical results from anthropology and other fields to present a comprehensive study of material culture. In this section I present Preston's account of reproduction.

§4.1.1 Function Theory: Displacing Intentions

How do items of material culture acquire their functions? How does a hammer acquire the function to drive nails? Whether the hammer comes straight from the casting mould of a solitary blacksmith or off the assembly line in a tool factory, presumably the answer must be that someone somewhere wanted nails driven and devised a plan to produce items for this purpose. But what if in the case of the manager of a hammer manufacturing plant, there never was any actual intention to drive nails, only the intention to generate profit from the sale of hammers? Is the intention to drive nails then a proximal or mediated intention of the manager? Or does the hammer get its function from the customer who is ultimately the one using the item to drive nails?

The question of how material cultural items acquire their functions is bound up with the question of how such items are produced and reproduced. It does seem that such items owe their propagation to the presence of agent intentions. A stone does not become a paperweight unless some agent puts it to use in weighing down paper. A footpath emerges anywhere people tend to walk, regardless of intent. Are intentions what perform the feat of establishing functions for material culture? If so, how does this work, what role do intentions play in the process, and what is involved in further propagation? The answer we saw in the historical theory of function (ch. 2) was that agents directly select variants of item types and hence confer functions by a process of guided selection. The historical theory thus assigns to designers (and sometimes to users) a privileged role in establishing the proper function of an item type. In this respect, historical theories and use plan theories of function are quite similar. Preston attacks the privileged role attributed to the intentions of designers (or users) by each of these two major theories. She argues at length that items of material culture acquire their functions through patterns of use which defy the appeal to intention that is made by both prominent theories.

We also saw (ch. 3) how in the use plan theory, there is a pecking order among agents with respect to function. Houkes and Vermaas argue that a designer is the agent who determines the proper function of an artifact. Users who adopt a use plan for an artifact are bound by a standard of proper use set out by the designer. In adopting that use plan, users propagate the designer's intended function for the artifact, which thereby becomes its proper function. For Houkes and Vermaas, "the intentions of designers are efficacious in a way the intentions of users are not. A use plan devised by a designer is *ipso facto* a standard use plan, and thus immediately establishes a proper function," once it is communicated to and adopted by users.[130] The communicative aspect of use plans is a transmission vehicle for agent intentions, from designer to user and also between users. This is what allows proper functions to remain stable.

Innovative users can also establish new functions as a result of non-standard uses that are not intended by designers. This was a motivation for the distinction in use plan theories between standard and improper use plans. "Designers' intentions are sufficient to endow their productions with derived proper functions, which are full-fledged proper functions with all the Normative characteristics of direct proper functions. Users' intentions, on the other hand, are not sufficient to establish derived proper functions."[131] It is important that any theory of function for material culture be able to explain how non-standard uses can also establish new proper functions, since this often occurs precisely in the absence of designer intentions and communicative use plans. Old-fashioned irons acquire a standard/proper use as bookends, spittoons become doorstops, and pipe cleaners become material for craftwork. Preston picks up on this issue and addresses it using a pluralist vocabulary of proper function and system function.

§4.1.2 Function Theory: Prototypes and Phantom Functions

Novel prototypes are an important example because the "history of reproduction" criterion in Millikan's theory is usually taken to mean that items with no ancestors also have no established proper function. These items do not meet the historical requirement

[130] Preston (2013), 167.
[131] Preston (2013), 168.

for proper-functional attribution, yet they can sometimes generate progeny or inspire copies that will later acquire proper functions. The International Space Station is a construction that has never been copied. Custom tools often have no ancestry or template from which they are copied. The standard response of intentionalist theories is to say that some form of intentional selection establishes the proper function but, in displacing designer intentions, Preston has jettisoned this source of function.[132]

The solution for Preston is simply to deny that novel prototypes have proper functions. It is the intentionalist theories themselves that have promoted this idea that prototypes must have proper functions.[133] They suggest that designers are somehow infallible regarding the functions of their creations. But this is clearly untenable, since the possibility of innovation depends on some creators overturning the functions established by earlier creators, thereby contravening those earlier design plans. For example, children's Play-Doh was originally invented as a cleaning putty for wallpaper, but later repurposed as a children's toy by the inventor's son. Both uses have had a claim to account for the reproduction of the lineage and hence both are proper functions to some proximity. So are there two distinct proper functions, or has one now superseded the other? Nothing requires that a prototype have a proper function in virtue of its designer's intention any more than putty must retain its original proper function when it acquires a new one. In the case of putty, there are competing legitimate claims to proper function, but one is more recent and more successful. In the case of prototypes, there is no legitimate claim in the first place because there is no reproductive history for which to account.

[132] Preston (2013), 165.
[133] Preston (2013), 173.

Items with phantom functions pose a different but related problem for historical theories of proper function. These items are reproduced in virtue of some function they do not perform and never actually could perform, e.g. in virtue of some magical effect that does not really exist (an "empty" function). Hence they threaten the criterion that items belonging to a proper-functional lineage have been reproduced on account of some propensity of their ancestors to actually successfully perform an alleged function.

Here Preston cites such examples as amulets or medallions thought to ward off evil and bug zappers thought to reduce local mosquito populations. One could add the metal bracelets that promise to direct the flow of magnetic currents in the body so as to promote health, or acupuncture pins that redirect chakras to alleviate pain. What is remarkable about these items is that none of them can perform the alleged function in virtue of which they are reproduced. They have proper functions in the sense that there are functions cited as the cause of their continued reproduction. But no tokens ever in fact perform the cited functions because these functions are not real effects.

What leads the proper function account astray in the case of phantom functional items is its criterion for successful performance. Historical theories require that some effect have actually been achieved down through the lineage to account for the existence of the token before us, even if that token cannot itself exhibit the effect. This criterion can never be met in the case of items with phantom functions because no ancestor has ever been capable of actually performing the effect in question; no item is. "The phantom function problem is…the more general problem of how any account that involves *reproduction* could work without appeal to successful performance."[134]

[134] Preston (2013), 179.

§4.1.3 Preston's Pluralism about Function (Redux)

If intentions are not what ultimately generate functions, just how do items of material culture acquire their functions? Preston tackles this question in her final chapter. What she calls the standard view has it that agents confer function by utilizing raw materials in accordance with their own (logically prior) purposes.[135] Without agent purposes, no item counts as an artifact. The standard view accepts a firm contrast between the way in which functions are acquired in material culture with, for instance, how biological items acquire their functions. "Biological functions are natural characteristics precisely because these functions do not depend on the intentions of an external, intelligent creator. In contrast, functions in material culture are non-natural characteristics, because they depend entirely on the intentions of intelligent human producers and/or users."[136]

Preston's is a pluralist view about function because she defends both proper function and system function, which might seem difficult without the standard view. System functions exist in virtue of the role an item plays in an embedding system. What is the embedding system that would confer function on e.g. a basket woven by an artisan producing it according to a ritual activity?[137] Must the embedding system be reproduced in order to account for reproduction in the lineage? If so, must this be intentional? Preston argues that what is important is that a structure (e.g., plant fibres for a wicker basket) interact with other components of a system. These other components can but need not be agent intentions. They can be existing tools and other items with proper functions, e.g. a stone placed next to an entrance can be understood to be a doorjamb. Components can also be non-intentional varieties of agent attitudes, or even social relations. Reproducing

[135] Preston (2013), 188.
[136] Preston (2013), 189.
[137] Preston (2013), 193.

a system-functional item is a matter of getting contextual arrangements right, just as re-producing proper-functional items is a matter of having the right history.

Preston recognizes that it is difficult to let go of the very plain and appealing idea that material cultural items acquire their functions in virtue of the privileged, consciously represented intentions and purposes of agents. The standard view seems to have a simple rejoinder to all this, namely that the functions of such items "do seem to depend on purposes human beings have and explicitly represent to themselves."[138] Nothing about the structure or raw material of a stone is modified when it gets made into a paperweight. What else could reasonably explain this transformation if not an agent's repurposing of the item? How do even the most simple tools and instruments get made if not by the purposes of agents who design and use them? What does Preston's pluralism tell us here?

Preston traces historical examples of this standard view in her introductory chapter, focusing on Aristotle, Marx, and (more recently) Dipert. The sort of objection for which the standard view naturally seems to give support here is basically that it is the sharing and transmission of purposes among agents which seems to account for function propagation in material culture.[139] Even the patterns of use which renew and sustain these functions, over and above the specifiable intentions of individual designers, might be said to depend on some form of collective intention.[140] This basic point is captured by the much more fully articulated intentionalist use plan theory of Houkes and Vermaas. Recall from §3.4.3 that they argued that artifacts have functions by specifying goals on behalf of *users*, i.e. by communicating the intended use plans of designers. This strikes Preston as an idealization. Real world communication is always sketchy. One's ability to use a

[138] Preston (2013), 199.
[139] Preston (2013), 200.
[140] Preston (2013), 201.

toaster does not depend on the successful transfer of an intended use plan from the designer. One can often infer a function on the basis of tradition and locally available knowledge,[141] or perhaps discover it by trial, or just guess at it. It just isn't the case that people always learn about functions in the way the use plan model depicts. Transmission of use plans "is overwhelmingly a matter of transmission from user to user,"[142] and even then the kind of thing being transmitted isn't always a specific use plan for a specific item, and the functions of simple tools are often inferred from local know-how.

Preston also cites research in cognitive anthropology to further deflate this objection from the standard view. The patterns of use that generate function in material culture propagate by a system of apprenticeship learning. This is not only true of children, who acquire their progressive training in the material culture "under the tutelage of already competent adults…working up through mastery of these subtasks to their own full-fledged competence," but also of adult peers encountering new objects.[143] Adults acquire facility with new items and their functions through general capacities such as imitation and other messier, more flexible methods such as trial-and-error demonstration, sink-or-swim adoption, etc. Use plan communication is obviously a useful method but is it really the most accurate model for learning in material culture? Even the training of adults "bears little resemblance to the explanation-based communication of fully formed use plans from designer to user that Houkes and Vermaas take as their paradigm."[144] The propagation of proper functions does not seem to occur on the model of this paradigm,

[141] Cf. Morton (2006); §3.3.1 above.
[142] Preston (2013), 202.
[143] Preston (2013), 203-204.
[144] Preston (2013), 203-204.

either. Preston argues that the intentional use plan aligned with the proper functions of items is itself actually a *result* of the culture's regimen of apprenticeship learning:

> …apprenticeship learning is the fundamental process through which proper functions come to be explicitly represented [and] learning by observation or explanation are secondary processes parasitic on it. More importantly, though, the representations of proper function acquired in this fundamental way are representations of items that already exist in the surrounding culture. And the point of insisting on the ubiquity of apprenticeship learning of everyday skills is to emphasize that such representations of proper function are acquired as much through direct interaction with material culture itself as through interaction with other agents…we encounter [e.g.] wastepaper baskets as always already proper-functional.[145]

So if agent intention is not the primary privileged source of function for items of material culture, just how do items acquire their functions? And how should the function of an item be determined? Preston's advice is that to settle a question about an item's function, "what we would need to do…is look at the specific patterns of use—where and under what conditions do people acquire these items, and most importantly, what do they do with them after acquiring them?"[146] Her pluralism affirms that for proper function the crucial determinant is history, but this is the history of reproduction understood as actual use as opposed to guided selection or directed variation. "It is not a matter of users' intentions, but rather of patterns of actual use that contribute significantly to demand for an item of material culture, and thereby drive its reproduction."[147] This approach is consistent with the problem cases of prototypes and phantom functions. For prototypes, Preston simply bites the bullet in accepting that they lack functions. For items with phantom functions, patterns of use can still explain why the item propagates: "the proper function is the effect the item of material culture would have to have [in order] to make sense of the pattern of use to which it is put."[148]

[145] Preston (2013), 204.
[146] Preston (2013), 186.
[147] Preston (2013), 186.
[148] Preston (2013), 186.

Preston's pluralist account of how material culture is reproduced places the emphasis on patterns of use over history or embedded role. Is it enough to explain the occurrence of technology change in a way that would endorse the inexorability thesis? Recall that the thesis advocates an evolutionary explanation of technology change, so as to avoid explaining it in terms of the guidance of human agents. It is not clear on the basis of this introductory sketch whether Preston's view would endorse the thesis. To assess the plausibility of the thesis, we must look more closely at applications of evolutionary theory to the study of culture, artifacts, and technology change. I will argue in §4.3 and §4.4 that cultural evolution does support TD's inexorability thesis and that it does so in a way which lends credence to Preston's account.

§4.2 Lewens on Technological Evolution

What would make the inexorability thesis viable would be to show that some retentive process other than guided selection could drive technology change. We saw in the previous section how Preston argues that agent intentions do not drive material cultural reproduction in the way the standard instrumental view or "use plan" theory has made out. To make this central point more vivid, I want to turn now to how material culture has been studied using the theoretical and explanatory resources of evolutionary biology. It is a common enough refrain and perhaps even an obvious idea that technology change occurs by an evolutionary process. If this were true in the same sense as it is true for organisms, then the inexorability thesis would stand affirmed. For it is the great hallmark of evolutionary theory that it explains how organisms can change and adapt in abundant diversity without the intentional guidance of designers.

Tim Lewens (2002; 2004; 2009; 2012) has argued that while artifacts do evolve, this does not occur in the same sense as it does for organisms. The basic reason is that organisms acquire their adaptive traits through the power of full-blooded Darwinian natural selection, whereas in the case of material cultural change there appears to be no suitable transmission mechanism to support a similar process. The version of evolutionary theory which can be applied in this domain is so diluted, Lewens thinks, that it lacks virtually any explanatory power.[149] He urges theorists of technology to consider what is actually gained over traditional models of change by an application of evolutionary theory in domains that study material culture (rather than e.g. homegrown theories derived from Marx or elsewhere).

In this second section I explain how Lewens approaches this question in order to understand the criticism he presents against the notion that technology change is profitably viewed as a process of natural selection. The view of selection he offers does not support the standard rejoinder, for instance, that technology change occurs by a process of intentional selection. And the more basic question about how to explain the acquisition of function is deflated across both the biological and material cultural domains. According to Lewens, the major obstacle facing the proponent of "full-blooded" Darwinian technology change is the problem of inheritance, i.e. the lack of an appropriate transmission mechanism. In the third section I consider replies to this objection from memetics and models of cultural evolution.

[149] Lewens (2004), 18-19.

§4.2.1 Lewens on Adaptation and Selection

Adaptationists have found it informative to compare organic form with the design of artifacts. This is an old strategy inherited from natural theology. Even the comparisons Darwin himself makes to artificial selection and breeding are rooted in it. Lewens wants to bring out the disanalogies and relevant analogies between organisms and artifacts to clarify how selection differs from design in explaining adaptation. To understand how he objects to evolutionary theory as an explanation of technology change, we should first acquaint ourselves with his account of selection and function.

The major difference between intention as it gives rise to design and selection as it gives rise to complex organic forms is that the two concepts operate at distinct explanatory levels. Intention refers to events which occur at an individual level. A farmer elects to use a new model of plough to harvest her crop, placing strain on the old plough's prospects for future utilization. Selection, by contrast, is a probabilistic, population-level explanatory concept. Its explananda are compositional changes in populations rather than causally discrete events at the individual level.

Lewens attempts to make clear how selection as a population-level category can explain adaptations in organisms at the individual level through an effort to distinguish selection and selective forces. Selection and drift are statistical evolutionary forces whereas selective forces, e.g. predation and birth, are individual events that cause changes in the composition of a population—i.e., ordinary physical forces. Evolutionary forces are not like ordinary "microlevel" physical forces. To illustrate this difference, Lewens compares the microlevel forces which can determine the outcome of an isolated coin toss with the statistical forces which might affect the probability that particular sequences of individual tosses will be observed. He claims e.g. there are no grounds for the idea that drift is caused by individual events like lightning strikes, or selection by instances of predation.[150] This implies that evolutionary forces must be seen not as outcomes or results but rather as processes, to preserve their status as genuine forces.

Lewens makes it clear that adaptation, while not unaffected by the distinct "microlevel" forces in an environment such as scarcity of water, an influx of predators, or increased fertility, is driven not by these but by the evolutionary force of selection:

> There is a multiplicity of forces that act on individual members of a population, which may give rise to changes in the frequency of types in that population. These forces do not have the peculiarities of the force of selection. They can act on one type of individual alone, and they can act to produce changes that are identified with drift or changes that are identified with selection…unless we are careful, it is easy to slide from "selection pressures" or "selective forces" thought of as a heterogeneous collection of forces that act on individuals, to selection thought of as an evolutionary force acting on populations.[151]

Selection must be distinguished from selective (i.e. material/microlevel) forces even though they are not wholly unrelated: the ways in which the material forces of causally discrete individual events affect trait fitness has an impact on whether selection acts.[152] Yet the effect of selective forces on the variant individuals in an environment might be exactly similar to those of drift, and selection will still be said to occur.

How then does selection contribute to the explanation of adaptation? To answer this, Lewens asks: "What is the relationship between selective forces thought of as causes of the differential reproduction of certain variants already present in a population, and selection thought of as a cumulative process whereby new adaptive designs are efficiently created? How does the first conception of selection relate to the second?"[153] His answer is that, while selective forces play an important role in enlarging or trimming the set of in-

[150] Lewens (2004), 25.
[151] Lewens (2004), 26-27.
[152] Lewens (2004), 28.
[153] Lewens (2004), 31.

dividuals on which selection acts, and making combinations of genes more likely, selection itself acts by retaining those variant traits which promote fitness in a population. Selective forces will always be present even when no variants exist in a population, whereas selection ceases to act in the absence of variation. This is what makes selection a "peculiar" force next to ordinary physical forces. Selection is the accumulation of useful variants from the pool of trait-bearing individuals that is produced and trimmed by selective forces operating at a causally discrete level, e.g. births and deaths that affect the composition of the population but do not actually act to *promote* traits.

This account has the consequence that "selection only sometimes explains adaptation, and that selection is not required for the explanation of adaptation."[154] Selection only serves to make adaptation "more likely than a random search, because if one is limited to a finite number of trials, one is better off trialing the already fitter variants in the hope of finding ones that are fitter still, than picking variants at random."[155] However, the complexity that makes existing variants fitter also contributes to the explanation of adaptation. Lewens argues that development in organisms makes adaptation a result of the increased likelihood that new traits will retain and that adaptations will accrue in a population. "Developmental organization itself is instrumental in generating complex adaptation,"[156] and "we should conclude first, that selection alone does not explain adaptation, and second, that selection only explains adaptation in rather tightly circumscribed contexts. Selection is not by itself 'cumulative'—the outcome of selection processes have

[154] Lewens (2004), 32.
[155] Lewens (2004), 32.
[156] Lewens (2004), 19.

this feature only when the items undergoing selection have certain properties," namely the right sort of developmental organization with which to retain useful traits.[157]

Emphasizing the role of development in retaining useful traits and helping generate new ones runs counter to adaptationism and its artifact model. The major obstacle to the adaptationist view is that in highly integrated complex organisms like multi-cellular creatures, cumulative evolution is unlikely to occur because of the many deleterious effects even a slight variation in the makeup of an organism can have. Selection alone has a difficult time explaining how the right sort of developmental organization that is receptive to trait retention can arise in the first place.[158] So the upshot is that selective forces, selection, and development all play some role in explaining adaptation in organic forms.

§4.2.2 Lewens on Function

Lewens argues against historical or "selected effects" (SE) accounts of function such as Millikan's and in favour of causal role (CR) accounts, i.e. the "system function" theory we saw from Cummins in chapter 2, as the best way to make sense of function attribution in biology.[159] A function of a trait is the contribution that tokens of the trait make to fitness.[160]

If that is what is meant by function claims, why make such claims at all? Lewens gives this question an historical answer which cites the artifact model and the legacy of natural theology: any account of function that appeals to the history of a trait to make sense of normative and allegedly explanatory connotations of function claims in lip serv-

[157] Lewens (2004), 34.
[158] Lewens (2004), 36.
[159] Lewens (2004), 87.
[160] Lewens (2004), 17-18.

ice to this legacy is probably misusing the notion of selection. Lewens argues "that selection does not, in fact, meet these [normative and explanatory] connotations particularly well," and that "the nonhistorical concept meets them well enough, and that in any case, it is not clear that all of the connotations that have been thought to be marks of teleological function claims should really be accepted."[161] Selection does not underwrite the connotations of function claims derived from the old artifact model, as Lewens shows by demonstrating that inorganic sorting processes also support these same connotations (examined below).

Lewens describes the three connotations we invoke when ascribing functions to artifacts via an agent's intention: (1) that function ascriptions are explanatory; (2) that they have normative content; (3) and that functions can be distinguished from "accidents".[162] This "intended effects" account of function for artifacts is, of course, straightforwardly our old friend the standard view so well elaborated in use plan theories. The SE accounts (e.g., Millikan) are designed to preserve these connotations with appeal only to the process of natural selection: "The function of a trait T is *F* iff T was selected for *F*."[163]

One objection to this is the disanalogy between selection for some property in an orthodox biological sense, and in the intentional sense where something gets selected "under some set of criteria."[164] No appeal to selection is required in the latter case, and SE theories of teleosemantics take on (from adaptationism, perhaps) the mistaken or at least unnecessary idea that natural selection plays some distinctive role in creating nor-

[161] Lewens (2004), 18.
[162] Lewens (2004), 88-89.
[163] Lewens (2004), 90.
[164] Lewens (2004), 91.

mative states even though “selection is not necessary for the imposition of historical norms in this sense.”[165] Another major problem for SE accounts is that no appeal to history is required. Lewens develops a “naïve fitness” (NF) account of function that he argues can support all three connotations in virtue of the fact that “current environment” can have a plausibly extended meaning.

Hence the two concepts of selection—natural and intentional—are quite different: “The intentional concept is that of picking out some item in virtue of some perceived capacity it has. The concept of natural selection is of one type outcompeting another in virtue of some causal capacity.”[166] He cites three major differences: (1) “natural selection relies on the existence of different competing types, whereas intentional selection does not;” (2) “natural selection involves at least one of the competing types having the selected capacity, while an agent can select an item on the basis of a merely perceived capacity;” and (3) “natural selection is essentially a population-level phenomenon, while intentional selection is not.”[167] Intentional selection operates at the individual level, the way selective forces like births and predation do for organisms. Technology populations are subject to gain or lose variants without any population-level selection occurring, such as when Microsoft elects to discontinue Windows XP or Toyota recalls and destroys a line of a defective car model.

Lewens claims that “when we attribute a function to an artifact, what we really do is attribute a set of beliefs and goals to an agent who comes into contact with that artifact. The capacities of the artifact itself—past and present—need not feature,” in stark contrast to the practice in biology, where the capacities and effects of earlier items of a given type

[165] Lewens (2004), 97.
[166] Lewens (2004), 109.
[167] Lewens (2004), 110.

are central.[168] The adaptations that inspire the design talk of adaptationism in biology are not "shaped" by selection in the way artifacts are designed to satisfy an intention; evolution always acts by modifying available and hence ancestral structures, and so all adaptations start off as exaptations, even if we wish to depict a distinction according to which traits were the result of "sustained cycles of gradual mutation and selection under constant selection pressures."[169]

Adaptationism's artifact model, as the idea that probably accounts historically for the presence of function claims in biology, is only attractive in virtue of the complexity exhibited in those organisms with "the right kind of developmental organization," i.e. ones that accumulate useful traits. But Lewens argues that the same artifact vocabulary can be justified in the case of other sorts of phenomena. His main example is inorganic sorting processes: "A sorting process is one where there is variation across a collection of items, and differential propensities among the items to survive some kind of test, but no reproduction."[170] Lewens gives a few examples of a sorting process: longshore drift, in which pebbles along a beach are sorted by the tide according to size, weight, or sometimes shape, resulting in regular separation of different kinds of pebbles. Another example is ion bonding in mass spectrometry: a catalyst fragments a compound substance when some of its ions bond more easily than others to its surface. This process sorts the ions by mass and can be seen to support functional generalizations about why one kind of ion bonded more readily than some other kind.[171]

[168] Lewens (2004), 112.
[169] Lewens (2004), 117.
[170] Lewens (2004), 127.
[171] Lewens (2004), 129.

The fact that it is false that only selection gives us license to speak of "genuine" teleology, as the SE theory would have it, is why Lewens characterizes his account of function as deflationary. Lewens argues that "there is no nonarbitrary way for the proponent of the SE account to say why sorted functions are any less genuine than biological functions."[172] Hence any conclusions about a normativity-conferring property of natural selection inferred on the basis of "the wide appearance of teleological language in biological inquiry" are mistaken because sorting processes also confer normativity in an exactly similar way.[173] Both sorting processes and selection processes give rise to equally genuine functions "because both processes support the three connotations widely thought to be the marks of genuine teleology."[174] Hence neither is privileged with respect to the teleology that an artifact vocabulary invokes in function attributions. Are the attributions themselves legitimate? Neither is more so than the other, but Lewens argues that there is a difference between function claims originating in intention and those originating in a selection process: "Functional specialization, and the assignment of different design problems…*precedes* the process of modification and testing in many cases of artifact design. In nature, functional specialization is instead an *outcome* of the action of selection pressures across whole organisms."[175]

§4.2.3 Material Culture's Transmission Problem

So in Lewens we have an account of adaptation that incorporates both individual-level selective forces, the population-level evolutionary force of selection, and the devel-

---

[172] Lewens (2004), 128.
[173] Lewens (2004), 120.
[174] Lewens (2004), 128.
[175] Lewens (2004), 136.

opmental organization required to support the selective retention of useful traits found in complex organic forms. And we have a deflationary account of function attribution that shows how the historically-oriented SE approach, e.g. Millikan's theory, gets selection wrong in casting it as an exclusive normativity-conferring process. Other kinds of non-selective processes can also satisfy the conditions for the teleological attribution that characterizes the artifact vocabulary. Even though the complexity of organic forms appears to invite the application of an artifact vocabulary, applying it in biology carries serious risk of misapprehending the actual causes of adaptations.

What about the prospects for an evolutionary approach to the explanation of technology change? Can incremental adjustments in technological design over time be seen as adaptations, as the inexorability thesis invites us to see them? Lewens thinks such an approach is possible but ultimately uninformative. "Artifacts do evolve, yet only a very abstract version of evolutionary theory that declines to comment about the broad character of selection pressures and the nature of cultural inheritance systems can be made to fit. The price for this abstraction is a corresponding lack of explanatory and predictive power when we try to apply evolutionary models to specific technological changes."[176]

This section explains the main criticism Lewens presents and the challenge it poses for TD's inexorability thesis. At its core, the inexorability thesis is an attempt to challenge the rather attractive idea that guided selection drives (i.e., predominantly causes/explains) technology change. An appeal is made to evolutionary theory because evolution in organisms is not guided by intentions. However, the form of evolution used in biological explanation is selectionist. Lewens raises a direct challenge to the viability of a selectionist explanation of technology. Later (§4.3) I consider how theories of cul-

[176] Lewens (2004), 18-19.

tural evolution can respond and (§4.4) how other prospects for an evolutionary approach to technology change might slide past the wedge in the door left by Lewens's own vitiations of selection's role in adaptation.

§4.2.3.1 Sorting Processes and the Utility of Evolutionary Models

Lewens seeks to operationalize (to some extent) the question of whether technology change is an evolutionary process by focusing on "the question of whether evolutionary models are likely to give us insights that non-evolutionary models could not have provided."[177] He claims that evolutionary models are legitimate but that they currently offer little explanatory impact not available in non-evolutionary models. It is possible, he advises, for theorists who find evolutionary models attractive for thinking about technology change to "make do with a very modest model that says that technology change is a sorting process, not a full-blooded selection process."[178]

This is enough to give such theorists the full range of functional teleology attributions. And it is enough to support "crude ideas of artifact fitness," as well as simple stories about how well particular lineages of items perish or propagate according to e.g. market conditions, social organization, etc. The fit between items of material culture and their environments, and their reciprocal effects on one another, can be described with the same level of granularity and detail as the story about the ionic bonds. Some of the mathematical tools of evolutionary models will even be available here to describe the values of trait fitnesses, e.g. "selection coefficients" for describing ions and artifacts.

[177] Lewens (2004), 139.
[178] Lewens (2004), 140.

However, just because models in one domain can be exported with limited success doesn't mean there is a strong affinity. "What all this shows is that we cannot take either the appearance of the vocabulary of evolutionary biology in describing some non-biological system, nor the usefulness of the statistical mathematics of population genetics for predicting the behavior of such a system, always to be indicative of deep similarities between the system in question and evolving natural populations."[179] But Lewens thinks the grounds for borrowing concepts like fitness from evolutionary models are much stronger in the case of e.g. primitive tool reproduction than in the case of ion bonding. For one thing, he acknowledges that in tools there appears to be a copying process:

> it is quite straightforward to think of a stone tool as an evolving entity. A tool type is copied by observers according to certain criteria, which may be functional, ornamental, and so forth. These criteria will partly determine the chances that a tool will promote the production of further tokens; that is, the criteria determine the tool's reproductive fitness. Some tools will be copied more often, while other tools which do their jobs poorly will be discarded. So long as the copying process is fairly faithful we can see that some tool types will tend to increase their frequencies in the population of tools according to these reproductive fitnesses.[180]

Nothing about the ion case involves copying of this sort.

The case of mass-produced technologies in material culture is less hopeful. Lewens points out that "successive 'generations' of [mass-produced] artifacts typically do not give rise to each other through chains of reproduction, but instead owe their production to a common cause."[181] Last year's television models do not generate this year's line in any direct fashion. The resemblance here with organic populations is still better than it is for ions but a new problem is raised, namely the fact that there is no way to trace lines of descent between tokens of material culture belonging to distinct generations. Lewens uses the example of automobiles, whose properties in tokens in one year may influence

---

[179] Lewens (2004), 141.
[180] Lewens (2004), 141.
[181] Lewens (2004), 142.

the next year's model's properties, yet the exact causal mechanism by which this influence occurs is difficult to describe. There is no possibility of assigning offspring to token items of material culture, hence estimating the reproductive fitnesses of such tokens will be likewise quite difficult.[182] (Lewens does suggest that related concepts such as inclusive fitness—a measure that "includes the contributions an organism makes to bringing organisms into existence that are not its own offspring"—do not face this same difficulty.[183])

§4.2.3.2 Material Culture and Replicator Status

On this basis Lewens claims that "so long as we do not insist that an item can possess the property of fitness only if it has some number of identifiable offspring, then artifact populations can be said to possess heritable variation in fitness, and they can be thought of as undergoing selection processes as a result."[184] One problem is to determine what role artifacts could play in this selection process. What units of analysis should be the focus of investigation, e.g. artifacts, ideas, a process, an artifact/idea complex, memes, or something else entirely?

Lewens broaches this problem in a preliminary way by discussing a distinction from Hull between *replicators* ("an entity that passes on its structure largely intact in successive replications [e.g., genes]") and *interactors* ("an entity that interacts as a cohesive whole with its environment in such a way that this interaction *causes* replication to be differential [e.g., organisms]").[185] Lewens argues that this distinction is unavailable in the technological domain, and that any of the proposed units of analysis can act as replica-

[182] Lewens (2004), 143.
[183] Lewens (2004), 143.
[184] Lewens (2004), 144.
[185] Lewens (2004), 145.

tors, such as behaviours, ideas, or items themselves. Hence the sort of replicative process required for a genuine selection process is not to be found because any component can play the part.

This can be seen by considering whether artifacts meet a simple test for replicator status: were an entity's structure changed in some way, would that altered structure appear in the next generation?[186] "The test provides evidence in favor of the view that it is the structure of the element in question that is, at least in part, responsible for ensuring the resemblance in structure across generations, hence that the item is, indeed, a replicator."[187] Using this test to determine replicator versus interactor status reveals that assigning this status is largely a contextual matter. Other mechanisms besides genes are likely to turn out to count as replicators (Lewens mentions Jablonka's epigenetic inheritance examples, e.g. methylation in mammals); perhaps some DNA won't count. And the nature of the alteration can make a difference. There might also be instances where one item triggers the production of a new item with a similar structure "yet does not serve as a *template* for it."[188] Faithful reproduction might occur without a robust causal connection.

For example, consider a 3D printer equipped with (1) a computer that can direct the design of new 3D printers, (2) a robotic appendage for assembling the printed components of new printers including copies of the computer and assembly appendage, and (3) an indefinite supply of energy and raw materials. Suppose that this first printer can create and assemble all the components required to make copies of itself, and that its copies will in turn be able to create more copies, etc. It can print all the circuitry and parts with which to construct and assemble new printers. Does it count as a replicator? Presumably

[186] Lewens (2004), 146.
[187] Lewens (2004), 146.
[188] Lewens (2004), 147.

not according to the test, since the instructions for printing and assembling the components for subsequent tokens reside with the computer, and the computer does not have any way to respond to structural changes.

Now imagine that an accident occurs during duplication of one of the daughter 3D printers. A misprint of the computer chip changes the printing instructions so that the new token does transmit changes in parental structure to the structure of its offspring. Would the members of this new lineage count as replicators? Likely they would. What this example shows is that it is by no means always the case that items of material culture will be ambiguous with respect to replicator status. Lewens does not consider cases, for instance, in which machines are purposely designed to operate as replicators with no ambiguity, as it is presumably possible to do.

The candidates for replicator status in cultural reproduction at large are exponentially greater in number. But the salient question for Lewens remains whether artifacts in particular are in fact replicators with respect to technological reproduction, rather than e.g. mental states or individual behaviours. "My answer here is a cautious 'sometimes,' although again with a wary eye on qualifications about the possibility of very indirect copying, about the failure of the counterfactual test to distinguish template copying and triggering, and about the relativization of replication to certain kinds of substitution in certain contexts of copying machinery."[189] There will be many cases, we may grant, where artifacts are merely interactors, and even that there is no principled way to determine which status to assign in a given instance. "The answer, then, to the question of whether artifacts are replicators is that artifacts of all types can sometimes be replicators in some contexts. Sometimes, however, they act as interactors without also acting as rep-

[189] Lewens (2004), 149.

licators."[190] But it is impractical to distinguish cases where they acquire genuine replicator status because other elements of agents are often implicated (Lewens cites beliefs, intentions, and manufacturing processes in particular).

§4.2.3.3 Heritable Transmission

Thus Lewens concludes that a question about the proper units of selection for material cultural reproduction is likely ill-posed since (1) it glosses over the replicator/interactor distinction and (2) it does not adequately account for the contextual nature of either of these statuses.[191]

> What this means is that for any given area of technological reproduction, the shifting attention of agents involved in the reproductive process will cause different items to become replicators at different moments. Worse, chance alterations to different types of item may cause agents to pay attention to them, thus resulting in their becoming replicators for a short time. There is little chance of finding any stable series of replicators, hence little chance of establishing a general, informative, theory of cultural inheritance, in virtue of the role of reasoning agents in the processes of technology change.[192]

However, there is a further obstacle. Exactly what is the mechanism of inheritance in material culture? How can there be uniform transmission if it is so difficult to determine which entities are replicators? Genetic transmission in biology takes many different forms, but even this range of forms pales in comparison to the ways in which material culture proliferates and alters. So the problem facing proponents of evolutionary models of technology change is that "the ways in which technological inheritance is ensured, and the many ways in which technologies combine with each other, are likely to depend on fine-grained contextual factors in a way that will make the discovery of any general rules

---

[190] Lewens (2004), 150.
[191] Lewens (2004), 150-151.
[192] Lewens (2004), 151.

of technological transmission very unlikely."[193] At any rate, proposals for such general rules are not likely to fare better than traditional theories of technology change, e.g. Marx's, or Freud's, or economic models.

Lewens notes that mutation in biological evolution is not unlike directed mutation found in material cultural innovation. Genetic mutation is certainly not random "in the sense that all items are equally likely to arise through mutation at any locus."[194] If the idea of directed mutation just means "that there is a bias in mutation in favor of mutations that increase fitness rather than decrease it,"[195] then ordinary genetic mutation can be seen as directed "without posing any threat to the traditional Darwinian view of mutation, since the fact that some mutations are more likely than others says nothing about whether it is the fitter mutations that are more likely."[196] His earlier comments about the role of developmental organization in permitting and retaining novel organic forms in fact makes such a view quite probable. The upshot for heritability is that even if directed variation exists, as it is generally supposed to exist in material culture, Darwinian selection can still play a role in explaining why one variant goes to fixation. Even guided variations can vary in fitness in a way that makes innovation subject to selection processes.[197]

§4.2.3.4 Design and Innovation

Lewens returns once more to the question of the role of intention in artifact evolution. Is the fact that human designers have goals what makes evolutionary models of technology change uninformative? Selectionist models are required just where there are

[193] Lewens (2004), 156.
[194] Lewens (2004), 154.
[195] Lewens (2004), 154.
[196] Lewens (2004), 154.
[197] Lewens (2004), 155.

no intelligent designers. This is the traditional view, according to which agent plans must exhaustively explain such change. "When we have real designers, the objection goes, with real intentions and real plans, no such theory is needed. This would be a poor objection to applying selection theories in the technological domain because it is not clear how even real intentions and real plans explain the emergence of good design."[198] Lewens points out that both poor designers and good ones alike have plans and intentions to design their creations, yet only some of these plans manifest as good designs. So do we need natural selection to explain good design for innovation too?

Lewens thinks that to the extent that design processes are explicable at all, "evolutionary developmental processes look to be quite promising explanations of success—at least for some artificers at some times."[199] He agrees that much anecdotal evidence suggests that "the process of invention follows an algorithm where a set of variants is created and some are selected for further modification."[200] But he reminds us that there are alternatives to selectionist construals of this alleged algorithmic route. Simple sorting processes are one. Another is that in some cases there is no sorting or selection process, only knowledge and reasoning on the part of agents, to explain the utter absence of variation. These causes themselves may have proximal selective causes, he acknowledges, but "if what we are interested in is proximal explanations for design success we need not always invoke Darwinian mechanisms."[201] For TD to have a strong case, however, the inexorability thesis should not be characterized such that intelligent agents are intentionally utilizing selection processes to innovate.

[198] Lewens (2004), 160.
[199] Lewens (2004), 161.
[200] Lewens (2004), 161.
[201] Lewens (2004), 161.

Lewens also considers areas in which selectionist evolutionary models of technology change can offer useful insights. The relative looseness of function talk which Lewens has established through his deflationary account can support a number of fitness claims that could prove beneficial to explanation in the social sciences. Artifacts could be said to have functions in virtue of many different kinds of effects, from economic and political to psychological and sociological.[202] Each domain might lay claim to a fitness value tied to a function attribution, and different levels of selection could be distinguished by the variety of selection processes through which artifacts pass during their design and production phases.[203] This can be useful in explaining both good and poor design.

Finally, as a corollary to his main point about the way development constrains the range of morphological variation available in mutation, Lewens makes a similar point about material culture:

> …when we think of the developmental program of an artifact, this concept needs to reflect the likely possibilities for how an artifact's form might change at a moment in time. So this program needs to be interpreted very broadly to include parameters for change fixed by all of the processes that go into its formation. These will include factors relating to how the artifact is conceived by its designer, or, in the case of a corporate research effort, by the group of designers that fashion it. The claim that only those artifacts with the right developmental organization will evolve to show complex adaptations is supported by the common observations that to solve a problem one must learn how to think about it in the right way, how to organize parts so that functional subsystems do not interact with each other in detrimental ways, how to represent design parameters in preliminary drawings, how to measure performance and so forth.[204]

Success in artifact design and technological innovation is also a result of benefitting from the right developmental organization as well as the right selection pressures.[205] How processes of design are themselves modified in response over time is also part of the picture.

[202] Lewens (2004), 158.
[203] Lewens (2004), 158.
[204] Lewens (2004), 159.
[205] Lewens (2004), 160.

§4.3 Memetics and Cultural Evolution

Lewens points out some general drawbacks to the intuitive appeal of the sort of evolutionary models that would otherwise tend to affirm TD's inexorability thesis, and he raises at least one specific trouble, the transmission problem. Evolutionary models for technology change are not wildly implausible, only apparently uninformative and in need of defending. How they fare in comparison with existing approaches from the social sciences is not yet clear. Selectionist models must face the lack of a suitable transmission mechanism by which changes in technologies could be treated as heritable. In this section I consider two sources from which to draw broad lines of reply to this particular objection: memetics and the study of cultural evolution.

§4.3.1 Memetics and Material Culture

The study of memes proceeds more or less as an affirmation of the idea, contra Lewens, that it is possible to distinguish replicators in non-genetic domains, e.g. material culture, psychology, etc. If correct, then TD's inexorability thesis still stands. The fundamental question here is about the status of memes as replicators.

§4.3.1.1 Memes as Replicators

Memes are defined on analogy with genes. They are imitable patterns of information in the same way genes are replicable patterns of information. Memeticists like Blackmore (1999) and Dawkins (1976) want to push the analogy because they think that what is special about genes is also what is special about memes: both are replicators in the sense (from Hull) discussed in Lewens: a replicator is a structure that interacts with its environment in such a way as to cause duplication of its own structure. Memes are information structures that consist in sets of instructions for duplicating their own "phenotype", typically through imitation.

Blackmore is well aware of the sort of objections that Lewens raises about transmission and the status of artifact as replicators, and she goes to some length to address them for memes in a general way. She sometimes refers to the meme as "the second replicator" because its method of replication is indirect. Memes replicate by imitation rather than by direct material duplication. Imitation produces replicates of low fidelity and are thus much more prone to error in transmission. This can give the impression that they do not succeed in transmitting their information. However, just because imitations tend to fail at replication, it does not follow that no transmission occurs. Other factors can facilitate replication and transmission.

For instance, some memes are "catchier" than others. The tune of "Happy Birthday to You" or "O Canada" is probably easier to hum to oneself than a more elaborate arrangement like "Stairway to Heaven". The point is that structural affordances exist in patterns of information just as they do in material forms. Some memes possess affordances that catch our attention or stick in our memory as cognitive attractors. We could speculate that this even has some psychological explanation: the ways in which we learn by "chunking" information to retain it just might happen to coincide with the structures of certain memes better than others. Those structures would therefore be more likely to replicate regardless of imitability, and hence have a higher propagation rate.

If there is replication even at low fidelity, there can be transmission. The objection Lewens raises against evolutionary approaches to technology change is that items of material culture tend either not to count as replicators or to count as replicators along with a host of other items, with no identifiable point of transmission. But if memes can in fact duplicate their structures at low fidelity and achieve transmission, this objection can be put aside. Perhaps what Lewens means, when he says that the replicator status of artifacts cannot be distinguished from that of other items that are also implicated in an alleged transmission of structural information, just is that the transmission will be low fidelity because no source is identifiable. But I would suggest that what we need to identify is not a source, but rather an isomorphism between a source and a target. Knowing the structure of the target helps us to determine the structure of the source. For instance, somebody walking by an art studio observes a small class of students helping each other carve birds. After observing for a few moments, the pedestrian goes home and carves a (low fidelity) bird based on what was observed. However, no one student in the art studio was observed carving a bird from start to finish, and hence there is no singular discernible transmission source. Yet having a completed target should allow us to claim that a transmission occurred, and perhaps to declare one student arbitrarily, or even to cite a composite source.

§4.3.1.2 Material Culture as a Memeplex

Blackmore clearly thinks of items of material culture as memes. She cites artifacts as one of the earliest examples of higher imitation.[206] And this is borne out by anthropology and ethology. Humans are distinct from other animals in two major respects: language and sophisticated tools. Both traits are present in other species to some extent. But in humans both of these traits are, by comparison, highly developed. Stone tools in par-

[206] Blackmore (1999), 76.

ticular recur throughout the prehistoric record in higher proportions because their matter is more durable, but also presumably because this greater durability affords more opportunities for imitation. This helps anthropologists explain why stone tools ultimately form the basis for so many other technologies (as I discuss in the next chapter).

Sterelny proposes modifying the notion of memes to better explain transmission. He argues that the locus of the meme should be reconceived, and that we should regard material cultural lineages themselves as memes.[207] This may not initially appear to help the proponent of the inexorability thesis. The suggestion might seem to be that we solve the transmission problem by identifying artifacts as their own units of selection, but this is not quite the idea. Rather than thinking about memes in broad terms, as ideas or behaviours, Sterelny suggests construing them as tangible artifacts and very concrete learned skills. "In contrast to ideas, artefacts and skills are public and copied…we can make better sense of competition between artefact variants than between ideas."[208]

Sterelny places this view opposite one he attributes to Sperber, on which it is prior features of human cognition that explain why certain kinds of artifacts, designs, skills, etc. are the best propagators. According to Sterelny, Sperber's "innate cognition" view does not explain why some artifacts are so robustly evolvable, since we can suppose there are variations in social contact, learning ability, and other factors between cultural populations. What Sterelny thinks must explain the "takeoff" of convergent designs across cultural lineages are the identifiable affordances technologies have. These affordances make technologies "readily and accurately copiable" and hence underwrite not only the fitness of these items as memes but also "an account of meme fitness that is independent of the

[207] Sterelny (2006b).
[208] Sterelny (2006b), 156.

actual replication rate of the meme."[209] Artifacts do not directly copy themselves, but their structural features do make them ideal candidates for replication.[210]

One significant consequence of this view is that a sort of co-evolution of humans and material culture occurs.[211] What this means for the question of the replicator status of items of material culture is that it may not be possible to distinguish separate transmission mechanisms for cognition and memes-as-material-culture. But this does not mean that definitive replicator status cannot be assigned to items of material culture, as Lewens claims. What it means is that more study is required to determine the role of cognitive mechanisms and memes as material culture in the replicative descent of information patterns. Perhaps there is a transmission dynamic in which both cognitive mechanisms and items of material culture participate, in which case the worry Lewens raises about how to assign replicator status remains. But perhaps the mode of transmission does exhibit preferential tendencies, in which case it might be possible to identify genuine replicators.

As Boyd and Richerson note, this issue will ultimately be decided by cognitive, biological, and anthropological evidence: "It may be that cultural information stored in brains takes the form of discrete memes that are replicated faithfully in each subsequent generation, or it may not. This is an empirical question that at present is unanswered."[212] Sterelny's co-fitness view, with its emphasis on affordances, directly challenges a view

---

[209] Sterelny (2006b), 156-157.

[210] "Fit memes are those with properties that make them apt inputs into template copying or imitation. In particular: (i) artefacts are more easily used for template copying if they are physically robust: for they will be stable and enduring. (ii) It helps if the physical materials from which they are made are readily available, and if the costs of making them are not too high. (iii) It is very important that the artefact be easy to reverse-engineer, that is, that the final form of the finished artefact does not conceal its history of production. Pots and compound bows are hard to reverse-engineer. It is hard to tell how a pot is made from examining the finished product. A fish-spear, on the other hand, can be readily reverse-engineered. (iv) Finally, artefacts are fitter if their design is error-tolerant: if crude, early versions of the technology give its users some reward" (Sterelny 2006b, 157).

[211] Sterelny (2006b), 158.

[212] Boyd and Richerson (2005), 421.

like Sperber's on which the fitness of human cognitive psychology alone is responsible for rapid and convergent takeoff in culture. Humans actually enhance their cognitive fitness through the construction and inheritance of prostheses.[213] Ereshefsky identifies a similar example in evidence from the field of primatology:

> Boyd and Silk assume that the units of transmission in cultural evolution are memes in the head that are transmitted from primate to primate. Yet memes can also be artifacts outside the head, such as nut-cracking tools, that are transmitted from one primate generation to the next. Furthermore, these artifacts can display cumulative evolution. Consider a study by Mercader et al. (2002). They report that a set of cracking tools has been used by a band of West African chimpanzees for over a hundred years. Moreover, they suggest that this set of tools has evolved over time. Some stones have been physically altered, and some stones have been replaced. Over time, the tools have improved with respect to the task of nut cracking.[214]

We can thus resist a view like Sperber's on two grounds. First, Sterelny insists on the importance of causal interaction between technology and human cognitive biases. The picture he attributes to Sperber has it that cognitive biases are causally prior to artifact proliferation and extinction, but that these biases—and indeed the whole mind—are products of co-evolution with technologies that have shaped these cognitive features.[215]

Second, the technologies themselves are dynamic in that certain simple designs crop up again and again in different cultural circumstances. Robust artifact designs appear to share four characteristics conducive to convergence: (1) they are reverse-engineerable, (2) incrementally improvable, (3) useful in simple variations, and (4) always prone to at least partial success. Spears, for example, are robustly evolvable because they possess affordances that are invariably beneficial for humans.[216] Some technologies are just more likely to arise and propagate in virtue of their affordances, given the ease

---

[213] Sterelny (2006b), 150.
[214] Ereshefsky (2004), 916-917.
[215] Sterelny (2006b), 160.
[216] Sterelny (2006b), 161.

with which their designs lend themselves to copying and their usefulness. In this way, we can begin to characterize the convergence of designs without appeal to replicators.

§4.3.1.3 Objections to Blackmore

Blackmore's memetics casts human minds in much the same way that TD does, as products of some larger force. In Blackmore's case that producer is a memeplex, while for proponents of TD it is a technological milieu. These views fit well together. Blackmore seems to think of the mind quite literally as just a meme vehicle, much the way organisms are just vehicles for genes, being driven by but never driving change.

One image for her idea is of memes as parasites and human minds as hosts for their reproduction. Meme production can be quite detrimental to the aims and desires of hosts. A proponent of TD might tend to regard technology and its designers in a similar way. However, the individual agents who carry out the imitations, replications, and material cultural reproductions do play an integral part. It is both true that they are being driven by and also that they are crucially important cogs in the machinery. A TD proponent would merely add that, overwhelmingly, memes take the form of material culture.

Eve Jablonka (2005) offers some responses to Blackmore's view. She worries that the copying process—predominantly imitation and emulation—cannot be distinguished from what it copies.[217] This worry is similar to the one Lewens raises about replicator status. Her concern is that imitation is sensitive to both context and content, whereas the sort of copying observed in genetic transmission is not. A lion cub will carry its parents' genes no matter whether they are copied in the wild or in a zoo, or whether the parent un-

[217] Jablonka (2005), 210.

dergoes changes in its phenotype over its lifetime. But the upshot of Jablonka's objection is the same: imitation is not a true replicative process, hence memes are not replicators.

Another worry Jablonka raises is that memeticists have not adequately explained how culture originates or how its changes affect the yield and uptake of memes, yet this all seems central to their account of meme propagation. The ways in which economic and political forces affect social organization and cultural creation require more nuance to explain than neo-Darwinian theory can offer. She scolds this theory's impoverished conception of the environment and its orthodoxy about variation, both of which lead it to fail to appreciate the environment's role in generating cultural traits.[218] The theoretical resources of Darwinian selection seem just not to apply in the case of memes and material culture—"selection acting on discrete units that are not altered during the process of transmission and that are random with respect to the factors that affect their generation and subsequent chances of spread"[219]—and so all of the language of evolution in memetics can be dismissed as posturing.

Jablonka does agree that basic concepts of evolutionary theory must be recast for the study of culture. But the impetus for her view here is that she thinks this is required in biology as well. Like Lewens, she emphasizes the role development plays in inheritance, and hence in selection, and this leads her to propose a limited role for epigenetic inheritance. This role is presumably amplified in the cultural domain because how societies grow and enculturate individuals goes much further toward shaping the mechanisms of cultural inheritance than it would appear to go in the process of genetic transmission. Ultimately Jablonka (and Lewens) are probably correct that memeticists like Blackmore

[218] Jablonka (2005), 222.
[219] Jablonka (2005), 227.

simply have not given us enough to erect a suitable theory of transmission and propagation of memes to underwrite TD's inexorability thesis. But there is another more promising approach that calls for a closer look.

§4.3.2 Material Culture and the Dual Inheritance Model

Boyd and Richerson (1985; 2005; Richerson and Boyd 2005) have long pursued the proposal that group selection drives macroevolution in culture, based on a mathematical model of dual inheritance. The basic difference between humans and other animals (even other species who manufacture and use tools) is that group selection plays a tangible feedback role in genetic selection via a secondary mechanism of cultural inheritance.

Cultural variations are heritable in human groups, whereas in virtually all other observed species they are not. New Caledonian crows manufacture at least three variants of stick tools for food retrieval, but each has a distinct function and so far as is known there is no cultural transmission.[220] Selection at the group level can also affect genetic reproduction in humans. Individuals who do not conform to group behaviour regarding e.g. hunting weapons can face exile or be killed, and either outcome will decrease the likelihood that the offspring of such individuals would continue as part of the group. In this way, human genetic lineages can build up around items of material culture.

Material cultural variations can be a catalyst for group selection because in many cases it can be easier to copy a tool than to learn a new language or idea. The contact with European societies experienced by indigenous peoples in various regions around the world was in most cases transformative not because of the new languages or ideas the

[220] St Clair and Rutz (2013).

Europeans introduced, but more often by the tools such as guns, maps, writing implements, and medicine. Agricultural methods in earlier eras are thought to have had similar effects on group survival and extinction. Written language also helped Asian societies better preserve information for cultural longevity.

The dual inheritance theory cites some of the same basic mechanisms as memetics—imitation, social learning, etc.—but supplies a precise mathematical model of cumulative cultural adaptation. The basic insight is to recognize the simple idea that innovations build on what came before. But this is not simply the notion that new inventions are only possible because of materials or inspirational models that were not previously available to an innovator. An enriched cultural milieu actually creates an environment more hostile to individuals who resist or refuse to utilize new technologies. (Preston talks about how proper functions are "policed" by individuals in a material culture, a point to which we return in the next chapter.) Group selection influences psychology by retaining more and more individuals over time who possess those psychological/cognitive traits that are adept at using the group's traditional technologies. Culture has largely obviated the need for human speciation in occupying new habitats.[221] People can and have lived in the Nile Delta or on Baffin Island so long as they are outfitted with appropriate cultural resources for those environments.

§4.3.2.1 Adaptation in Material Culture

How do variations in material culture accumulate to generate adaptations? Boyd and Richerson argue that variations are transmitted and retained in several ways. Social

[221] Richerson and Boyd (2005), 243.

learning processes such as imitation or emulation are the most common means of transmission, but items of material culture themselves also serve as models. For example, "designs that are used to decorate pots are stored on the pots themselves, and when young potters learn how to make pots they use old pots, not old potters, as models."[222] Variations result from individual guidances, or from different forms of bias in the transmission process, e.g. direct bias (responding to affordances), frequency bias (how common a variant is), or model bias (imitation).[223]

One of the primary mechanisms these authors identify is conformist bias, which acts to correct errors in replication at the individual level and thereby to preserve useful traits at the population level, providing a basis for group selection. Variants of an item that are more common tend to be replicated more faithfully because there are more attempts made at replication. Adaptations occur when cultural variants are differentially maintained either at individual or group levels. The forces of guided variation (i.e. the results of individual departures) and biased transmission are responsible for many traits, but according to Boyd and Richerson they can be overcome by natural selection. Certain conditions must be present for direct bias and guided variation to cause successful innovation in the first place, e.g. initial adopters must be receptive and influential. More often a new cultural variation is spread with less conscious guidance.[224]

As for biased transmission versus selection, they distinguish the former by comparing it to "meiotic drive" in which some genes are more likely to get carried forward because of their effect on genetic copying. Biased transmission also works over and above the effects a cultural variant might normally have on the propagation of its lineage

---

[222] Richerson and Boyd (2005), 61.
[223] Richerson and Boyd (2005), 69.
[224] Richerson and Boyd (2005), 174.

in virtue of direct affordance bias.[225] So individual choice in technological innovation is to some extent an illusion. It is guided by the same forces that guide genetic reproduction, and even where it is not the influence of selection can be strong. The force of natural selection on material cultural evolution is so strong, in fact, that its effects show up in the archaeological record as studied by neuroanthropologists. Material cultural tools have affected both the evolution of psychological mechanisms, which in turn influence social learning styles and mechanisms, as well as the human body itself. Marzke (2013) compares the grip styles and musculature in humans and chimpanzees and concludes that the heavy use of stone tools likely played some role in the shaping of the hand. Plausibly, hand morphology for such tools was preserved through the success of those using them.

Boyd and Richerson argue that cognitive psychological mechanisms are one result of this feedback effect between the two forms of inheritance. The sorts of cognitive capacities modern homo sapiens have acquired in a short evolutionary time frame would not likely have arisen in a population that had not also been enhanced by cumulative cultural adaptations.

> To understand the evolution of the psychology that underlies culture, we must take this population-level feedback into account…Under the right conditions, selection can favor a psychology that causes most people most of the time to adopt behaviors 'just' because the people around them are using those behaviors…the ability to imitate can generate the cumulative cultural evolution of new adaptations at blinding speed compared with organic evolution. A population of purely individual learners would be stuck with what little they can learn by themselves; they can't bootstrap a whole new adaptation based on cumulatively improving cultural traditions. This design for human behavior depends on people adopting beliefs and technologies largely because other people in their group share those beliefs or use these technologies.[226]

The idea that genetic evolution produced the brain, and that "*then* culture arose as an evolutionary byproduct" is not tenable.[227] There is a feedback effect between the mechanisms

[225] Richerson and Boyd (2005), 79-80.
[226] Richerson and Boyd (2005), 12-13.
[227] Richerson and Boyd (2005), 12.

of our psychology "and the kind of social information that this psychology should be designed to process."[228] Competition and selection among groups has the general effect of reducing genetic variation in the species as a whole, thus leaving us with peaks of cognitive phenotypes clustered around particular material cultures.[229] The focus of these authors is culture in general and so they emphasize the social aspect of this feedback dynamic. But our interest is in the dynamic between the material products of a culture and the cognition of its individuals. To what extent are we adapted to our material culture?

That material culture does exhibit adaptations is probably not highly controversial, but it does require some justification. Boyd and Richerson acknowledge that the adaptationist stance has come under fire in philosophy of evolutionary biology from the criticisms of Gould and Lewtontin. But even if exaptations and other mere accidents can mitigate the need for adaptive explanations in biology, it is probably unreasonable to think that cultural and technological traits are generally like spandrels. On the one hand, the accumulation of useful traits occurs and spreads too quickly to be attributed to the purely genetic evolution of human cognition. On the other, it does not occur quickly enough to be attributed solely to the guidance of individuals.

Likewise, Boyd and Richerson believe that anti-adaptationists such as Gould and Lewontin are surely too conservative in their hesitancy about adaptive explanation: "we should be equally cautious, perhaps more cautious, about casually accepting nonadaptive just-so stories that invoke mysterious unspecified events or tradeoffs."[230] They reason that if adaptive explanation is useful in the study of organisms, it may have some utility in the study of culture as well, where the explanation of cumulative useful traits/variants

---

[228] Richerson and Boyd (2005), 12.
[229] Boyd and Richerson (2010), 3794.
[230] Richerson and Boyd (2005), 103.

faces similar problems. And the emphasis Lewens places on development's role in adaptation is not an automatic objection to adaptation in the cultural domain either.

§4.3.2.2 Additional Objections from Lewens to Dual Inheritance

The challenge Lewens raises in his (2004) can, I think, be met by the preceding considerations. Even though e.g. Godfrey-Smith thinks that cultural variants need not be replicators to have successful variation and selection,[231] the replicator condition can be satisfied and a viable transmission mechanism does appear to be available. Lewens (2009; 2012), however, characterizes the population approach as nothing more than aggregative thinking with no distinctive evolutionary flavour, i.e. "the kind of thinking one engages in when one explains the behavior of a unit composed of varied parts in terms of the properties of those parts and their interactions."[232] He contrasts the dual inheritance model with memetics, and acknowledges that their proposal of the mechanism of conformist bias is more promising than imitation as a mechanism of propagation.

Lewens seems to be chiefly concerned that the model they propose will not be enough to animate the social sciences.[233] He also rightly points out that the features of their characterization of individual psychology can be opposed, though this will be on comparably minor empirical grounds. Oddly, Lewens questions their model's application of "population thinking" despite the emphasis he himself places on such thinking in his exposition of selection. But his major complaint seems to be that the tools of population thinking are not strong enough in the cultural domain to displace already established

[231] Godfrey-Smith (2012), 2166.
[232] Lewens (2009), 249.
[233] Lewens (2012).

theories: "population thinking is not distinctively Darwinian, nor will population thinking displace traditional forms of historical inquiry, ethnographic investigation, and so forth."[234] The fact the these statistical tools merely add to the toolbox, whereas in the case of biological theory the toolbox was virtually empty, somehow weighs against counting them as properly evolutionary. It may be true that population thinking has not displaced traditional theories in the social sciences, but it seems a double standard. The explanations put forward in biology on this same basis are also weak, but they appear relatively strong because there are no major competing explanations. This seems a rather arbitrary basis on which to count or discount this approach.

Godfrey-Smith makes virtually the same objection against population thinking, insisting that the collective process involved is not smarter than the individuals in the required sense.[235] But what his objection seems not to address is the feedback process at the heart of the dual inheritance model. Contrary to what Godfrey-Smith seems to state, the claim of the model is that population-level feedback processes produce smarter individuals over time by enhancing the material conditions for cognitive development. It may be true that the process is not smarter than its members, remaining utterly unintelligent, but it also accumulates in individuals those useful cognitive changes that further sustain the process. The aggregation objection does not speak to this aspect of the model.

§4.4 Adaptive Material Culture and Technological Determinism

The dual inheritance model of cultural evolution is promising but also peculiar. It posits a secondary transmission mechanism by which cultural variants are transmitted and retained. The cumulative adaptations of this secondary inheritance feed back into the ge-

---

[234] Lewens (2012),
[235] Godfrey-Smith (2012), 2167.

netic inheritance which reinforces the evolution of cognitive and psychological traits in individuals. Hence the dual inheritance process produces individuals and material cultural lineages over time which are co-adapted. However, Boyd and Richerson do not presume that this endorses TD:

> even the strongest skeptics of culture's significance must make an exception for the culturally transmitted knowledge that produces technological differences in the same environment. Many might be comfortable with technological determinist explanations granting that aspect of culture important causal power. But cracking the door of dispute this far greatly weakens the environmental determinism argument, because there is no clear dividing line between technological knowledge and other forms of knowledge."[236]

A great deal of further empirical study would be required to support the claim that technological items of material culture in particular are causally privileged in the process of human cognitive evolution (over, say, other environmental factors) or even that the culture is predominantly adapted in its material as opposed to its social or other forms (a weaker claim).

So why should we accept TD's inexorability thesis? There is no knock-down empirical evidence to support it, and the more suggestive elements of the available evidence rely on charity and interpretation to make a compelling case. But it should be clear that the thesis is viable. All it requires is the possibility that material culture can exhibit adaptations that are not produced under the direct guidance of intentional agents. If adaptive technology change can result from evolutionary forces such as selection, then the thesis is viable. The major obstacle facing this viability was the objection concerning transmission which Lewens raised. But the dual inheritance model of cultural evolution seems to provide the makings of a plausible response.

Part of the concern that writers like Lewens, Gould, Lewontin, Jablonka, etc. have with adaptationists such as Dennett and Boyd and Richerson is that the idea of universal

[236] Richerson and Boyd (2005), 29.

Darwinism is almost certainly wrong. Despite how rich and useful the theory of evolution by natural selection has been, there is no good reason to think that all systems exhibit adaptation. And, as Lewens argues, it is not difficult to show that many systems simply are not evolving. But the universal Darwinist is not a substantive intellectual opponent, but rather a straw caricature. Adaptationists in the social sciences are better treated as though they are invoking the idea that evolution occurs only when particular conditions conducive to selection and adaptation are present. Is it clear that systems under such conditions are confined to the domain of organisms? What about artificial systems designed as virtual recreations of the biological world? It seems arbitrary to close off applications of evolutionary theory on the grounds that their viability has been difficult to establish. It is not entirely clear even among orthodox applications just what the exact nature of the conditions for selection and adaptation are, as is familiar from many debates in philosophy of biology in recent decades.

TD's inexorability thesis is just the idea that the required conditions can obtain in material cultural reproduction. No author reviewed here is currently in a position to deny that these conditions could or do obtain. The dual inheritance model provides a strong example of a viable approach to social scientific occurrences of the conditions in question. It also provides a good basis on which to study Preston's notion of patterns of use. Her account can be challenged either on its pluralism regarding function or on the imprecision of the notion of patterns of use as the key to material cultural reproduction. But if the latter can be reinforced using dual inheritance, then her account as a whole is strengthened.

Suppose, however, that nothing about material culture or artifacts is felt to be on firm enough ground to support the sort of robust empirical work necessary to make headway on the proposed application of evolutionary theory. Cultural phenomena are just too vague and slippery to support such analysis, especially given the sheer range of material items (at best a broad family resemblance notion[237]). What then?

Ideally we would like to find a singular naturalistic vocabulary and theoretical stance for classifying and predicting the outcomes of systems in which the conditions for selection and adaptation obtain. However, it may be that the conditions themselves turn out to be poorly defined. The furor over adaptationism, and the fact that recent theorists like Jablonka and Lewens can point to such basic discrepancies in the nature of the dynamic between selection and adaptation (e.g. the role of development, the possibility of epigenetic transmission) indicates quite clearly that there is room for more fundamental research in biology and on how to export these ideas to other domains. Before dismissing application in the cultural domain outright, it may be prudent to consider whether further modifications can be made to the basic tenets of evolutionary theory that would accommodate more domains under the relevant conditions.

With respect to heritability, for example, consider Sperber's more relaxed notion of propagation. This notion is not tied strictly to replication. It is more vague than heritability, but it can be made to include replicators as a special case. And it is flexible enough to accommodate the apparently non-transmitting but self-reproducing cases we associate with the cultural domain. The true problem cases are those items which straddle both domains, such as people, domesticated species, weaponized biological entities, synthetic organisms, and even the titular example of Sperber's paper, seedless grapes. This fruit

[237] Sperber (2007), 124.

has had its reproductive apparatus intentionally bred out of it, and is utterly dependent on human beings for its propagation. The point here is that we should not presume there are bright lines between biological and cultural phenomena, or that because selection and adaptation are components of a successful theory, they are currently understood well enough to be treated as reaching broadly beyond their domain of origin without significant modification and reappraisal. The phenomenon of the co-evolution of cognition presents challenges even for the vocabulary of biology.

Sperber wonders how an ordinary approach to biological function can explain the fact that the seedlessness of seedless grapes "is also an optimal biological adaptation."[238] Add to this the fact that our material cultural creations are transforming all the time, becoming more and more similar to biological creatures as human engineers adopt more and more inspiration from the design solutions of natural selection, and it is even more prudent to think broadly about how to unify the domain divisions in an intellectually satisfying way. This is what is at stake in the question about TD's inexorability thesis. What kind of modifications might this require? Sperber suggests both a more general concept of propagation and a deflationary attitude toward *artifacts*: "What all this suggest is that, in taking artifacts as a proper category for scientific and philosophical theorizing, we are being deluded by a doubly obsolete industrial-age revival of a Paleolithic categorization."[239] So we are misled on both fronts: "artifact" is an increasingly poor approximation for material culture, while biological theory is under growing pressure to clarify the blurry lines between nature and culture (perhaps insurmountable pressure).

[238] Sperber (2007), 136.
[239] Sperber (2007), 136.

Autonomy and Innovation

In §5.1 I present in more detail Preston's view on action, proper function for material culture, and the sociogeneric stance she defends alongside it. In §5.2 I defend her sociogeneric stance in connection with TD's autonomy thesis. In §5.3 I consider the compelling view of creative innovation Preston develops as a means of undercutting a reading of her account like the one I give (in §5.2). And in §5.4 I present reasons why her preemptive response might be seen as failing to avoid the autonomy thesis.

Preston describes her approach to action and agency in the social domain as a *sociogeneric* one, which means roughly that the cognitive competencies of individuals are thoroughly social in their origin and nature. On this approach, "the theoretically significant facts about individuals are always already social facts, because they reflect the internalization or instantiation of pre-existing social practices."[240] The kinds of things people can do, the kinds of skills we tend to associate with agency and action as guided by individuals, are virtually never novel. The learning and development phase of our growth as socially-scaffolded neural organisms guarantees that we obtain many of our activity and decision patterns from the ways in which we are trained to behave.

As we will see, Preston does not think such an approach is deterministic in the hard sense that the outcomes of individual decisions are utterly fixed by what others do or have done in the presence of the individuals under prior circumstances. What it means is that the basic influence is such that individuals acquire training largely modeled on these sorts of behaviours, and then apply those models in new situations. This does not mean that there is no intermediate causal step between influence and action where the acting agent does not make a decision to act in a deliberate way; only that the range of available choices is heavily constrained by social factors.

[240] Preston (2013), 78.

What Preston argues on the basis of sociogenerism is that the local material culture also has a very great impact on individual competences. Tools and technologies are one of the primary repositories for social learning. Some of the evidence for this has already been mentioned in the previous chapter. In this chapter, I examine sociogenerism and then the available evidence in its favour. I argue that the sociogeneric approach does indeed fall in line with the spirit of TD and in particular the autonomy thesis. I also consider Preston's attempt to escape this conclusion: her theory of innovation.

§5.1 Action, Sociogenerism, and Proper Function

What is the relation between agency and technology? As we saw earlier, the standard view treats technology as the instruments and tools—and artifacts—of agents. This view of material culture is an example of a derived theory, positing that the functions of technologies are an instrumental extension of the purposes and intentions of agents. Agents are logically prior to their derived creations, traditionally called artifacts in virtue of a derived ontological status. I do not intend to dwell on the metaphysics of these entities on either side of the agent/artifact dynamic, but it is important to be clear about the content of the standard view and the picture it offers.

It is interesting to investigate the historical origins of this view, which seem to be an emulation of creationist worldviews gleaned from religious traditions and elaborated through an engagement with ancient philosophy. Human beings are regarded on the image of a creator, who made everything in the world according to the shape and form of its own thought. Artifacts are conceived on analogous terms in relation to human agents. Individual human agents conceive a thought or plan for some purpose and then fashion a tool with which to achieve it. The intention/purpose/plan of the agent imbues the material item with an ontologically derived purpose for achieving the agent's goal.

Preston aims to break with this view once and for all. It has saddled philosophy and much social science for far too long. The derived view has made artifacts seem like uninteresting, unproblematic phenomena, as though there is nothing especially difficult to understand about them that is not already a difficulty for a more robust agent entity. The presumption that there is simply nothing at stake in the study of artifacts holds back a good deal of progress on issues concerning the role of technology in contemporary life, the ideals by which it can and should be designed, as well as the study of human agency itself. Even Preston's adoption of terminology reflects the depth of the break she wants. "Artifact" is simply an inadequate term for the variety of phenomena worthy of study in this domain. She borrows the more comprehensive term "material culture" from anthropology, where there has been greater recognition of the complex role that our tools and crafted material items play in our biological, cognitive, and social history.

In previous chapters (2 and 4) we saw how Preston's pluralist approach to function informed her views on reproduction in material culture. In this chapter I examine this theme in relation to what she calls her "sociogenerist" approach. This approach looks at reproduction in terms of the study of action. Action theory is a prominent discipline for the philosophical study of agency. Agents are entities capable of action, and moreover of intentional action (so the story goes). Preston argues that two somewhat neglected aspects of action must be attended to in order to make headway on the question of material culture and its reproduction, in connection with the action of human agents. By jettisoning the notion of artifact and the derivative dynamic on which it has been conceived by most philosophical and scientific traditions, Preston does not thereby wish to abandon the associated notion of agency. But she does think that we must look more carefully at two aspects of human agency not always emphasized in other treatments, namely improvised action and collaboration. My central question in this chapter is whether and how this sort of account can, under the banner of sociogenerism, avoid the autonomy thesis by emphasizing these aspects of agency.

Preston's criticisms of intention in function theory are echoed in her approach to the study of action. She has characterized the standard, derivative, intentionalist stance toward material culture/artifacts as a "centralized control model" of what artifacts are and how they work. This characterization establishes a two-pronged attack on the traditional standard intentionalist view of function and action theories of material culture. The centralization component of the traditional model is countered "with an emphasis on collaboration as the typical mode of human action," while the control component is attacked "with an account of improvisation as the predominant structure of human action."[241] Both counterproposals are basic elements of her own philosophy of material culture: the focus on collaborative action leads to her basic stance of sociogenerism and to its aid in the account of proper function, while the focus on improvisation gives her an account of innovation that preemptively defends against the sort of objection I aim to develop with respect to TD's autonomy thesis.

[241] Preston (2013), 222.

§5.1.1 Improvisation

When we think about design, e.g. the engineering of buildings and new computers and cars and furniture, the images that occur to us tend to be of blueprint specifications for prototypes and researchers inventing things in laboratories. There is an ideal associated with this image: the creation of a set of instructions for producing an item or outcome. A plan is made in advance for some thing's production. The process of design resembles an algorithm or effective procedure, or perhaps more informally a cooking recipe. But the recipe image betrays the inadequacy of this ideal, according to Preston. A recipe, like any design, is at best a guide rather than a comprehensive set of instructions for achieving some goal. Following it requires contextual knowledge. What it means to add a "dash" of salt, to omit ingredients to taste, or to what extent one can change the order in which ingredients are mixed is usually not specified. Similarly, every design contains wholly unstated steps that are both crucial for correct execution and also entirely superfluous when specified. The process of design is not the creation of an effective procedure. (Preston reminds us of the vast graveyard of AI projects that have died on the shores of the quiescent environment.)

People do not think creatively in a stepwise fashion, so it is odd that our ideal of design would suggest that this is how innovation occurs. The stereotype is of the inventor as an isolated problem-solver labouring away in a workshop or laboratory, planning a design for a machine yet to be constructed. In reality, innovators tinker. Preston counters this ideal with a view of human ingenuity as rooted in improvisation. How fundamental this sort of activity is for us is not apparent precisely because it is so deeply ingrained in all our activity, and hence usually quite invisible to conscious reflection.

This ideal calls back to the control component of what Preston calls the centralized control model that underwrites and guides most traditional thinking about the relation human agents have to their material culture. The presumption, lacking any real explicit argument in its favour to be found outside action theory, is that material culture results from the intentional planning of human agents. A designer comes up with a prior mental design for some item, in the form of a set of instructions or effective procedure which could, in principle, be carried out unintelligently, and then executes that plan to produce the item. She conceives of some new widget to solve a given problem, and then prepares a production process to actually build it. The design process is presumed to be similar for groups, as when a team of car designers works for months to come up with a blueprint for a new model of sport utility vehicle, adding features or removing them from last year's model. This is exactly the sort of picture that is developed in Dipert and in Houkes and Vermaas. And Preston argues that it is wholly backward.

People do not innovate by becoming better planners. They innovate by becoming better improvisers. But none of this is apparent from contemporary action theory. The study of action identifies intention as the crucial feature of action. Intention in turn is theorized in terms of planning. The orientation of the entire field appears to have ignored or severely downplayed the rich range of unplanned action. The focus on planning, Preston says, has been a response to concerns that action theorists initially made intention a simple combination of volitional and representational components.[242] This was inadequate because desires do not require that agents satisfy them. An agent can desire to go to Hawaii without ever acting on it, yet still have the desire. A plan, by contrast, contains

[242] Preston (2013), 47-48.

both beliefs and desires and also expresses a goal and the means to achieve it. But this approach to action leaves out an entire range of human activity:

> If all intentional action is plan-based, the very idea of unplanned or improvised action seems conceptually incoherent. Perhaps reflex actions, such as blinking, and unintended side effects, such as leaving a footprint in the mud on the way to the mailbox, might in some sense still be regarded as unplanned. But reflexes and side effects are not the domain of improvisation any more than they are the domain of planning. If improvisation is an interesting or important phenomenon it is so because it inhabits the same domain as intentional actions—most likely because it *is* a kind of intentional action. But there is really no conceptually coherent way of talking about such a phenomenon in the context of planning theories of intention and intentional action.[243]

Improvised action can be volitional without being intentional, but action theory tends to regard it as little more than an uninteresting spontaneous species of action. It tends not to countenance unplanned volitional activity. In addition to the example of cooking recipes, consider driving. Drivers have a volition to arrive at a destination by a particular route. But all along the way they make instantaneous decisions that are totally unplanned yet serve as a model of successful coordination in human activity. People change routes mid-trip, they slow down for other drivers or accelerate to pass them in anticipation of making an upcoming turn, they respond to the motions of other vehicles on the road around them, and they adapt their piloting of the car on the fly at every step of the drive. Such activities are impossible to conduct on the model of planned control. You cannot explicitly plan each step in a drive down the street, let alone across town. Both trips hinge on responsiveness to aspects of the situation as they arise. These examples suggest that improvisation is an important, fundamental aspect of human activity.

How do such "merely purposive" agents operate? "Here we are met by silence from action theorists…the dominance of the planning paradigm has generated the background assumption that any use—or, at least, any interesting use—of representations of

[243] Preston (2013), 52.

the world in the pursuit of goals counts as planning, so the question of how else representations of the world might be employed by human agents just does not arise."[244] And most attempts to address this deficiency have misconstrued the phenomena, either by casting it as merely spontaneous action or by focusing on cognitive processing rather than agency. What is required to identify how improvisation works is an approach to action that pays attention to the actual strategies and forms that action takes. How do improvising agents coordinate their activity over time? How do they achieve such clear gains without formulating explicit plans?

Preston argues that cultural practices are ultimately what implement such activities—practices habituated into individuals through apprenticeship learning and other methods. "Improvised action sequences are coordinated through the simultaneous use of multiple resources at multiple levels."[245] Action theorists have been inclined to treat habits and the cultural practices they implement in terms of plans. But if we treat them instead as resources, as raw material for the departures of unplanned individual actions, then we have the makings of an account of improvised action. "[In] our improvisatory picture, action is creative in an ongoing way as it turns a changing array or available resources to good use. This contrasts sharply with the planning picture, which relegates creativity to the mental construction of plans, thus depriving the action itself of any real, theoretical interest."[246] Preston thinks plan theories of action tend not to capture the sort of activity that especially characterizes the uses of material culture. Plans are always partial, but plan theories, especially the use plan theory, do not appreciate how much tradeoff there can be between structure/practices and execution in purposive human activity. No

[244] Preston (2013), 63.
[245] Preston (2013), 129.
[246] Preston (2013), 130.

such improvised activity is "formless or haphazard."[247] It takes its cue from the local culture, from tradition, from the ways in which individuals learn to use the items available to them.

Preston discusses three kinds of resources that facilitate improvisation: strategies, practices, and habits. She discusses practices and habits mainly in general terms: practices are "procedures for generating and concatenating actions that are specific to particular cultures or subcultures," while habits are more idiosyncratic.[248] She outlines three kinds of strategies for improvised action: *appropriating-and-extending*, *proliferating-and-selecting*, and *turn-taking*. Often, improvised action is a mixture of each.

According to Preston, *appropriate-and-extend* occurs when some unintentional course of action arises or presents itself, and is seized upon. A musician plays a wrong note, but regards the mistake as an interesting sound worth pursuing. An engineer notices that the code she has written solves a different problem in another application. Or a medical researcher notices that an unanticipated effect of a drug she has designed is useful (e.g., the libidinal properties of Viagra, originally a heart medication). The key for Preston is that the misstep must be unintentional, not simply vaguely intended. Usually an agent does not intend to make a mistake, but sometimes making one turns out to be useful, i.e. a lucky accident. The key for this and each of these strategies is that a situation presents itself in ways that are conduct-guiding "just in virtue of the constraints on future action it generates."[249] Such situations do not result from vague intentions, as when someone means to try something on the off-chance it might work. And the "contributions

[247] Preston (2013), 94.
[248] Preston (2013), 91.
[249] Preston (2013), 97.

are not so much accepted as *appropriated*. That is, improvising agents make previous contributions their own by interpreting them, revising them or rejecting them."[250] (103)

Proliferate-and-select first generates multiple options for using some structure or doing something, and then selects the best one. No plan arises in this case because the options are generated by "doodling" about, and only when they are laid out, compared, played around with, etc., is a final selection made. A painter follows their gut and makes a range of different landscapes, with different colours and depicting different times of day, and then chooses one. An engineer facing a difficult coding problem does not stop to think about the solution, but simply keeps trying different and unrelated methods in the hope that a solution can be found. Proliferation can be guided or free, but the important thing is that it be prolific. Generating a variety of options is the key in this strategy. It can take stimulus from anywhere, but typically the environment is the main source.

This strategy makes extensive use of material culture in two ways: (1) the materials serve as external memory for cognitive effort, and (2) the instruments themselves have affordances which contribute to the prolific output. Material cultural practices make doodling at least easier (and perhaps possible in the first place). This strategy can also extend over long periods of time: storage "makes possible a larger pool of items on which [agent] selection can operate," and "it makes possible retention of items for appropriation and extension, especially when the ongoing process is intermittent."[251] This dimension of output, of just how prolific human action tends to be, is identified as being even more important than how it is directed by the express intentions of an agent.

---

[250] Preston (2013), 103.
[251] Preston (2013), 114-115.

Finally, in *turn-taking*, we have a simple but highly versatile strategy for improvising action that is documented extensively in the scientific study of conversation. This is an ordered but not usually rule-bound form of activity. It allows for robust coordination without an explicit plan. Here we might also think of brain-storming in small groups, or even playing games like “Tag” without explicit rules. The idea is that coordinated action is easily achieved on the fly, without advance planning. Preston cites the extensive studies conducted by ethnomethodologists on conversational turn-taking, which have revealed micro-level structures at work in how people trade turns in speech.[252] There are many implicitly recognized transition points in every conversation, all of which are introduced by agents. These structures are not planned in any way, yet they structure speech all the same. Material cultural practices can often work in just the same way.

§5.1.2 Collaboration

Centralization in the centralized control model of material culture and action is a view about the locus of action. Much early modern philosophy presumed that individuals are self-sufficient with respect to capacities like reason and foresight. But many thinkers today would argue that these are thoroughly social competencies. The emphasis Preston places on collaboration is crucial because the locus of action and of material cultural production has to be completely reconsidered. She says that collaboration “has a special significance for a philosophy of material culture” in that it is “the most concrete manifesta-

[252] Preston (2013), 120-122.

tion of human sociality" and the "growing point of human sociality in general, and of specific cultures in particular."[253]

Paleoanthropology supplies some evidence here. Preston cites the claim of Peter Reynolds that material cultural production is not observed to be a particularly individual activity but rather is typically social in human cultures.[254] Primates do make tools but are more inclined to do so individually whereas humans, by contrast, produce tools almost exclusively in collaboration with each other.[255] This conflicts with the common presumption that innovators simply apply their reason and ingenuity to create new items. "Reynolds casts suspicion on the idea, explicitly endorsed by Dipert, that individualist reconstruction of what are in fact collaborative activities constitutes an appropriate methodology for the study of production."[256] It is not likely that designs are ever stored in minds in any full sense, due to task specialization and role complementarity. Action theorists have studied shared agency and group action, but again the approach has been rooted in individualist models. There is some attention to responsiveness between agents by these theorists, but little has been done "to include responsiveness to aspects of the action situation that are *not* other agents… [e.g.] nonhuman animals, items of material culture, and natural features of our environment."[257] A robotic system might be the best most current example of something responsive with which we interact and collaborate on a regular basis.

Collaboration is less a requirement than it is an advantage.[258] It is undeniably present in performances such as plays or concerts—examples cited by both Dipert and Pre-

[253] Preston (2013), 36.
[254] Preston (2013), 36.
[255] Preston (2013), 37.
[256] Preston (2013), 38.
[257] Preston (2013), 71.
[258] Preston (2013), 33.

ston, to different ends. Even in the deceptively simple case of carving a stone tool, the anthropological evidence suggests that only in primates is this sort of task carried out by lone individuals; in humans, it is virtually always collaborative. The sorts of tools produced also tend to differ quite a lot, of course, but the greater sophistication and design flexibility of human-crafted tools only speaks more strongly to the advantages of collaborative production. Preston argues that much of the action theory about social activity assumes a cooperative paradigm and then works to distinguish such action from competition between agents. Of course, the deeper paradigm here is the rational individual. Preston prefers to emphasize the collaborative nature of social action so as to reduce attention to individual agent intentions. But this collaborative enterprise she emphasizes is not the notion of group agency to which action theorists appeal.[259]

Collaborative action is characterized by mutual responsiveness and lack of a shared plan. Consider driving on a busy road. To some extent, there is a shared plan in the sense that there are explicit rules governing the flow of traffic. But for the most part, how drivers obey these rules is a collaborative enterprise. Drivers watch each other for cues, proceeding or yielding as caution and the actions of others permit. There is little expectation that coordination by set traffic rules will reliably control anyone's conduct, yet drivers manage to get around relatively safely all the same.

§5.1.3 Preston's Sociogenerist Stance

The underemphasized collaborative dimension of action points to the general importance of the social. No one denies that social existence is an important facet of human

[259] Preston (2013), 76.

agency, of course. But there has been a tendency in modern thinking to adopt a stance on which the individual is somehow prior to its participation in the social world, already replete with rational capacity, intention and planning, etc. The terms Preston uses to contrast this stance with her own are *suigenerism* versus *sociogenerism*.

§5.1.3.1 Suigeneric Individualism

The social sciences have no broad consensus about how to describe the nature of society or its influence on and relation to individuals. Many theorists are methodological individualists who do not take society as presenting a different object from the study of individuals, thus taking "society" as an aggregate phenomenon.

One of the better known and philosophically grounded recent versions of this view is Searle's philosophy of social ontology, which casts social phenomena as a result of the biological capacity individuals have for language.[260] The fact that we can create facts about social institutions such as currencies and presidents is due to the power of language to generate representations that do not correspond to states of affairs, but rather create new states of affairs. Social phenomena exist in virtue of status functions available to every speaker. Preston does not entirely oppose individualism in this vein, but she does want to modify it so that it better reflects psychological theory and is more coherent. The basic idea has been either that the social is constituted through the agreements or coordinated activities of individuals who are already formed to carry out such coordinations, or that the social is a collective activity (e.g., Searle, Gilbert, etc.).

[260] Searle (1995, 2010).

On the view Preston calls suigeneric individualism, the social world is created entirely through the actions of already-competent individuals. These competencies are therefore acquired in some non-social way.[261] We already saw how this sort of idea leads action theorists to untenable views about shared agency, on which multiple-agent cooperation is emphasized at the expense of competition and collaboration. This might seem like a straw caricature of the Enlightenment focus on the rational capacity of individuals, but this is quite evidently the argument of many prominent action theorists (e.g., Bratman and Gilbert), and by this argument individuals somehow arrive in the social world already endowed with the means for full participation. But agents cannot be congenital Crusoes, i.e. already self-sufficient.[262] There is simply no way to account for the competencies required to participate in social arrangements and activities without appealing to pre-existing practices and social resources. This does not beg any questions about the origin of social practices because, as we shall see, the competencies and the social practices which underwrite them have a common origin.

§5.1.3.2 Sociogeneric Individualism

How can we begin to explain society by appealing to individuals who are already socialized? Preston calls her own stance *sociogeneric individualism*. It says that we simply have no theory of pre-socialized individualist competencies. None of the important traits we associate with agents exist apart from the social conditions in which they are born and trained. Action theorists take an approach to collaboration that does not appreci-

[261] Preston (2013), 76-77.
[262] Preston (2013), 86.

ate this, and so they end up with untenable theories of group agency, collective intentions, multi-agent phenomena, and the like.

The basic insight missing from suigeneric accounts is that the shared dimension of action is ultimately primary. Trying to build up a theory of shared action from isolated components is doomed from the outset because it starts from a backward premise. It might seem circular to invert that premise, and to claim that an individual is a result of socialization and the circumstances of inherited group practices. But it is otherwise hard to say what is implicated. Acquisition of competence in these practices is facilitated by other individuals who have already mastered them. And subsequent action is always guided by the context these practices create around the action of individual agents.[263]

The radical prospect of this stance is that it reconceives the locus of action itself, casting it as a partly external phenomenon. The competencies of individuals are tendencies or dispositions which instantiate practices not internal to the agent.[264] This is a somewhat awkward way of speaking, but the basic idea is that the practices which enable the skills of agents are not already innate within individuals, and must therefore be obtained from elsewhere. But the dynamic is not entirely unidirectional; new individuals can make adjustments and small changes that recreate the social environment as they are enculturated. Individuals "transmit this social environment to the next generation of individuals,"[265] and, in so doing, introduce alterations.

The key consequence of the sociogeneric stance, for our purposes, is that material culture can be seen as a source of socialization. "[Material culture] is inherited—often quite literally—along with the social practices involved in its production and use, and a

[263] Cf. Preston's discussion of Baier, 78.
[264] Preston (2013), 78.
[265] Preston (2013), 206.

large part of what is involved in the development of full-fledged individuals is training in the use and production of the material culture peculiar to their society."[266] What this means for the production of human agency deserves closer comment.

§5.1.3.3 Material Culture and Sociogenerism

The role of material culture as a resource in the generation of agency and the enculturation of competent individuals is crucial. Items and lineages of material culture serve as repositories and storage for many kinds of practices. They serve as external vehicles and training manuals for learning to adopt and carry out those practices. And they serve as enforcers of cultural norms (as in e.g. Latour's example of the automatic door closing mechanisms that assert the norm of keeping doors shut[267]).

How does this occur? Mainly this is a consequence of the reproduction relationship that exists between the individuals who sustain a society's material cultural heritage and that heritage itself, which turns out agents well suited to the tasks required for the propagation of that heritage. Reproduction is a dynamic feedback process: "the reproduction relationship between human beings and material culture is not a one-way street… we are reproduced by our material culture just as surely as we reproduce it."[268] Once individual and society are regarded as mutually constitutive, and the role of material resources is acknowledged as a significant source of social training, it becomes clear why non-intentional forms of action like improvisation and collaboration are important. Notions like "agency" and "individual choice" (e.g., innovation) owe much to the scaffolding

[266] Preston (2013), 79.
[267] Preston (2013), 75.
[268] Preston (2013), 223.

support found in technology and other material resources. This isn't a strict causal claim, e.g. "toasters make you want to eat singed bread." But it is a claim about how desires for singed bread and intentions to singe it to finely controlled degrees tend not to arise in traditions where toasters are not available.

Proper functions track a standard use pattern for a given material cultural lineage. Dressers have the standard role of storing clothing, and the fact that they are often used in garages to store cords or bottles of screws does not upset this standard role. The proper function of an item of material culture tracks the patterns of use which best explain the history of its lineage's reproduction: "A current token of an item of material culture has the proper function of producing an effect of a given type just in case producing this effect (whether it actually does so or not) contributes to the best explanation of the patterns of use to which past tokens of this type of item have been put, and which in turn have contributed to the reproduction of such items."[269]

But crucially, it is not *only* the items and their functions that are reproduced. Preston argues that this same process also in a very literal sense reproduces features of individuals. Moreover, she repeatedly claims that a material culture generates the purposes of its agent inhabitants: "what is reproduced in material culture—or perhaps more precisely, along with material culture—are the intentions and purposes of the human individuals born into that material culture, and so in a very real sense those individuals themselves as they result from the developmental processes they undergo in the context of their material culture."[270] No inventor arrives on the scene capable of building something as intricate as a computer, automobile, or even a clock or wristwatch without a long prior tradition of

[269] Preston (2013), 187.
[270] Preston (2013), 160.

underlying technologies on which to base their craft.[271] The kind of agent required to invent a watch ex nihilo, wholly apart from a background history of "incremental variations" on which to tinker, is quite difficult to imagine.

In many cases, novel artifacts have no inventors. Preston discusses how the electric guitar was nearly simultaneously invented by at least three different people in different versions, with the aid of different groups and resources. Pinpointing a determinate individual as an inventor is a useful fiction: what really happened was that a combination of components became available for assembly from a variety of pre-existing resources, and over time several agents began jointly to experiment with different combinations. The "invention" of this device was neither sudden nor centralized, but rather diffuse.[272] Sociogenerism's central claim is that the purposes which individuals appear on the standard intentionalist view to generate ex nihilo and to bring to bear in various situations are not generated from within, but rather from without and, in particular, have their origin in material cultural resources. Preston pushes this theme even further: it is not just the purposes of individuals, but the very ranges and types of individuals found in a society that are brought about by its material culture. Thus the standard intentionalist view of artifacts

> ignores the full extent of the reproduction that takes place in material culture. It is not merely the reproduction of material structures, but of functions, of the human purposes corresponding to those functions, and of particular types of human agents corresponding to the particular configuration of the material culture into which they are born and within which they develop into fully competent adults…it is just as correct to say that our purposes are imposed on us by the material culture we inhabit, as that we impose our purposes on it.[273]

Proclamations like this recur throughout Preston's book. The whole historical context of a proper-functional lineage requires the reproduction not only of items but also

[271] Preston (2013), 157-158.
[272] Preston (2013), 216.
[273] Preston (2013), 188.

of the sorts of individuals who would have use for them. Practices and social roles are replicated alongside the artifacts to which they are wedded and fed back. A material culture is a sort of cognitive milieu, and indeed Preston remarks in her conclusion on the resemblance this sort of picture bears to extended mind accounts—a comparison I'll pursue in the final chapter. She cites work in developmental psychology and cognitive anthropology to support her suggestion that apprenticeship learning is a prime mechanism by which material culture propagates itself. Individuals acquire and pass on the proper functions of items because this is part of an inherited training regimen.[274] People acquire the kinds of representations of items they do because these functions are already established learning resources in each agent's surrounding environment.

Proper functions further propagate not only because that is how people are trained, but also because there are forms of social authority at work. The proper functions of material cultural items are policed by social convention. Here Preston cites Foucault's discussion of the Panopticon to illustrate the social effects of material and environmental structures, a point not unlike that of Winner's mantra that "artifacts have politics."[275] Proper functions eventually come to be all but synonymous with social constraints in Preston's discussion. "The only viable view is one that sees human purposes and the proper functions of items of material culture indissolubly linked in patterns of use and reproduction."[276] Sociogenerism begins to elicit a very radical picture of material cultural autonomy. Individual agency exists largely in virtue of social practices and competencies, while those practices and competencies exist largely in virtue of the historical patterns of a society's material culture. Proper-functional lineages of a culture's items require for

[274] Preston (2013), 204.
[275] Preston (2013), 205-206; cf. Winner (1977, 1986).
[276] Preston (2013), 206.

their maintenance the reproduction not only of particular kinds of purposes, but of the sorts of individuals who would have use for those purposes. From what I have said so far, it does not seem that Preston leaves much room for individual agents to assert control over the material culture in which they are raised and trained. Their purposes and aims appear to derive principally from the requirements of the historical inheritance of their material culture.

However, Preston foresees this consequence. She recognizes how sociogenerism "problematizes individual cultural creativity and innovation by emphasizing the extent to which agents are themselves products of their material culture."[277] She thus works to build into her account of material cultural reproduction a solution to this problem: system-functional innovation. This is a function theory correlate to her action theory notion of improvisation, and it is the key to her argument for creative freedom.

§5.2 Agency, Material Culture, and the Autonomy Thesis

How does all this relate to TD's autonomy thesis? It could be objected to TD, at least in the form I have presented, that it is still a product of the standard view, of traditional thinking about agency and its creations. Preston goes to great lengths to break with that view, yet she still ends up addressing similar themes. TD could be characterized as a version of the standard view (of the artifact/agent dynamic) which simply reverses the valence of the relation of derivation, i.e. so that human agents are the derived products of technological activity. The basic outlook is the same in that one side of the equation still determines the other. Is this all that TD amounts to?

[277] Preston (2013), 188.

The theme of technological control over society and individuals has a long and sordid history in social theory, in political philosophy, in literature, and in science fiction stories. More than one academic author/theorist has concluded that its longevity has more to say about us and our own worries than it does about the actual state of affairs in the world and in our social environment. I think that even if no person has ever actually held the strongest form of this view, as seemed to be the case after our examination of the possible formulations laid out in Bimber, investigating it is still very helpful in learning how to think about the influences and impacts of material culture. The fact that it remains a touchstone for so much musing and colloquial thinking about technology's impact on contemporary society suggests that sound ways of thinking about these impacts continue to elude us. We just don't yet know how to think very well about technology.

In this section I argue that Preston's sociogeneric stance and her account of proper function in the propagation, via material cultural reproduction, of individual agents comes as close as anyone to endorsing TD's autonomy thesis. It is quite interesting how earnestly Preston commits to these same kinds of ideas that have bounced around in social theory since the 1700s. But it is also interesting how she proposes to guard against a charge like the autonomy thesis. I investigate the extent to which her ideas affirm a claim like the autonomy thesis, supporting the connection with empirical evidence. I then look at the pre-emptive defense Preston mounts against a charge like this one: her account of creative innovation in material culture. Are system-functional deviations from proper functional uses enough to escape the charge of autonomy/TD? The view Preston suggests is explicitly ambiguous about the status of individuals and their material culture with respect to such notions as autonomy. In the end, this may work against her.

It could also turn out that Preston has endorsed an account of material culture on which individual agent autonomy is severely constrained by technological impacts but does not by the same token ascribe autonomy to those technologies. This is a subtle point and it is only obscured by conflating it with questions about artificial intelligence. What is at stake here is the status of the material cultural contribution to the phenomenon we have been calling “agency”. Does technology play no role in forming agency (i.e., it’s merely an instrument?), does it play some limited role, does it dominate but not control, or does it genuinely control out of some internal volition of its own? This last seems ludicrous, but as a colleague once remarked, if there’s a seat at the table you can be sure that some philosopher will sit down in it. I am not the one to sit down, but I am willing to pull out the chair to inspect its sturdiness.

§5.2.1 Proper Function and Agent Purposes

Here is the revised provisional definition of proper function for items of material culture that Preston proposes:

> A current token of an item of material culture has the proper function of producing an effect of a given type just in case producing this effect (whether it actually does so or not) contributes to the best explanation of the patterns of use to which past tokens of this type of item have been put, and which in turn have contributed to the reproduction of such items.[278]

When we examined this idea in the context of her discussion of reproduction, we found that it displaced the role of intention and the privilege and authority accorded to designers in establishing functions for such items, in favour of their being established through patterns of use. This implies that conferring proper function is a long-term social accomplishment without any simple relationship between the purposes of agents and the

[278] Preston (2013), 187.

proper functions of material culture.[279] Both are intricately linked and mutually implicate each other in patterns of use and reproduction.[280] Sociogenerism blends the two formerly opposed poles into a single "self-same social process" by which material cultural functions and the agents who wield them are produced, reproduced, and fed back again. The material culture literally generates the kinds of agents required for its own propagation. This may be why specialized knowledge of tools for navigating environments tends to remain localized even though weapons cross-pollinate through invasion and trade.

We might agree that material culture can *produce* types of agents, but how does it thereby *constrain* individual action? Preston argues that functional norms are enforced by a policing mechanism at the heart of material cultural propagation. What enables tool-equipped action also holds it to particular standards. Sometimes this takes the form of legal edicts or displays of political power. Individuals are encouraged to adhere to proper functions displayed by their confederates, and sanctioned for departing from them.[281]

§5.2.2 Policing Purposes: Sociogenerism and Material Cultural Constraints

How does this add up to an endorsement of TD's autonomy thesis? Recall that TD's autonomy thesis suggests that individuals lack robust suigeneric agency due to technological influences: the role that material cultures have played in generating the important features of cognitive agency over time has been decisive. A view like Winner's technological somnambulism might be thought to count on this score, but the norms driving technology while human agents are "asleep at the wheel" are, in his account, still

[279] Preston (2013), 206.
[280] Preston (2013), 206.
[281] Preston (2013), 213.

conceived as norms that humans put into place. However, perhaps by invoking *culture* and treating technology in terms of its material expressions, I have done nothing more than affirm a normative account of TD, which as we saw in the introduction, was dismissed by Bimber as being not sufficiently technological in nature.

Preston's sociogenerism would also appear to affirm a normative reading of the autonomy thesis: material cultural conditions tend to produce particular agent types. Even if the resulting agents had fallen asleep at the wheel after setting the cruise control, it was never their novel idea to set it in the first place, or at what speed, etc. Furthermore, the traditions and practices themselves conscript and police individual actions. To illustrate this point, Preston uses the image of Foucault's Panopticon, which is meant to show "how features of the built environment, just in virtue of their physical form and the uses to which they are put, inculcate beliefs and purposes that define types of individual agents suited to live, work, and play in precisely these built environments."[282] One of the capacities of the presence of material culture in a society is to circulate power.

So rather than claims about technological agency or attributions of autonomy, what emerges from the idea of Preston's relation to TD is that agency itself receives a thorough questioning/problematization. We certainly lack the sort of strict criteria that would endorse a suigeneric individualist conception of agency. What "agency" ought to mean just for the basic study of cognition, etc. must be evaluated anew. Latour's actor-network theory, for instance, has it that material culture or organic form can each (both) participate in agency, and that by "agency" all we can mean is some efficacious property of an organized system. The nature of agentic or causal power in this dynamic can mean that the effects of material culture can indeed count as exercises of agency.

[282] Preston (2013), 205-206.

§5.2.3 Defending Sociogenerism in Connection with the Autonomy Thesis

One of the most striking things about human evolution is not simply how drastically it has split away from that of the great ape relatives, but the pace at which this has occurred: "No other great ape lineage seems to have undergone such a profound transformation: as far as we know, living chimps and gorillas are broadly similar in habitat and ecology to their ancestors of five million years ago."[283] A multitude of factors are doubtless responsible for this feat, including language, social and moral cognition, etc. But the role of material culture, especially technology, in shaping our environments and in supporting growth and learning can be singled out apart even from these staples of theoretical study. The longevity, cumulative prowess, and hence the sophistication of technology provides an obvious and perhaps too easy answer to questions about how hominin lineages acquired such distinctive cognitive traits.

Sterelny, for one, argues for a "feedback dynamic" as a more plausible explanation of the acquisition of these distinctive traits, over the standard social intelligence hypothesis which he thinks fails to capture the dynamic integration of social organization, foraging skill, technologies, and learning environment in the explanation of human cognitive evolution.[284] In this section I want to emphasize the role technology has played in this special feedback dynamic. I argue that Sterelny presents strong evidence that compels us to take Preston's sociogenerism and its emphasis on material culture seriously. If by the autonomy thesis we understand that the conditions of material culture facilitate the production of agents, especially our purposes and distinctive competencies, then the thesis would seem to find support in anthropological and archaeological evidence.

---

[283] Sterelny (2012), 19.
[284] Sterelny (2012), 21.

§5.2.3.1 Apprenticeship Learning

There are two major ways in which learning reveals the critical, driving role of material culture in human cognitive development: apprenticeship learning, and the manner in which pedagogy accumulates what Sterelny (2012) calls cognitive capital. He discusses both, emphasizing that apprenticeship learning is founded on the latter. Preston only mentions Sterelny's account in passing, but it offers a strong defense of her appeal to apprenticeship learning in the rise of material culture.

Sterelny argues that apprenticeship learning is at the core of what has driven the transformation of our species into its unique modern form. Competent adults train novices in the skills required for group life, and these skills are then passed on by the students once they attain competence. But he explains that the techniques of apprenticeship have evolved over time, increasing in sophistication. The manner in which these techniques have evolved is through manipulation and organization of the social and material environment. To initiate novices in the training and deployment of increasingly difficult skills, it is not enough to simply rely on direct instruction. The master employs a range of tools and aids to impart the lesson.

For example, practice artifacts come to be relied upon more and more when the sorts of skills observed in the archaeological record increase in complexity and abstraction. Even stone tools must be chipped using particular techniques in order to get the proper shape and sharpness. Evidence shows that novices have been made to practice on stones that are prefabricated in various stages, for demonstration purposes, as well as on blanks that can be wasted.[285] This sort of approach to learning takes on an increasingly important role as the range and functionality of material culture increases. What it shows

[285] Sterelny (2012), §2.3

is that there has been an attempt by competent masters to offload some of the task of instruction onto objects that are both available and plentiful enough to be wasted. It reflects a stable abundance and commitment to learning processes.

The key shift in learning processes, though, is in what Sterelny describes as the feedback dynamic that occurs within the social and material environment. This is what drives the growth in human cognitive evolution over the past 50,000 years. The striking thing about Sterelny's account is that he insists we must not cite any one factor in particular as being responsible for the transformation. Basically his idea is that processes of social learning and transmission, coupled with strategies for modifying and cumulatively organizing the environment in which these processes occur, produced the remarkable cultural and cognitive adaptations that characterize modern humans. This thesis is opposed to the more familiar story of culture as a by-product of cognitive adaptations or of linguistic capacity. Language certainly plays an important role, but it is adapted minds and adapted environments that have done the heavy lifting.

Sterelny is clear (and Lewens, in his review, concurs[286]) that no single factor—e.g. cooperation, technology, cumulative social learning, ecological expertise, etc.—are solely responsible for the explosion in the evolution of human cognitive agency. Rather this can most plausibly be explained in terms of a feedback dynamic involving each of these factors. But it seems clear to me that technology does play a special "first among equals" role here. It is the key soil without which reliable and durable social learning tends not to grow and intensify. It is the core of the cognitively adapted environment.

An emphasis on technology seems to find support when in §5.3.4 Sterelny discusses the evidence against unidirectional technology change in the distant past. Innova-

[286] Lewens (2014).

tions can be seen to have been repeatedly gained and lost throughout the vast stretches of deep pre-history, with corresponding blips in the archaeological records of different peoples in different eras. Now technologies can be lost or gained as a result of social calamities, such as disease, war, or drought, etc., but what is clear is that once they are lost, they are difficult to regain. The Australian aborigines, upon migrating to that continent, lost much of the material culture they had previously accumulated. During this period their social evolution appears to have stalled. Instability of innovation is also observed in the Middle Stone Age period, plausibly a result of smaller bands of hominins dispersed over a large area with a correspondingly higher risk that innovations would be lost from one generation to the next.[287] This suggests that even though the feedback dynamic which underwrites the gain and reliable transmission of high-fidelity cognitive capital involves each of the above factors, the material cultural basis is especially important in that it provides the stable background against which the feedback effect can intensify. Material culture can be lost for a people in a way that existing neural and morphological adaptations, language, and social behaviour cannot or are far less likely to be lost once gained (provided the group still reproduces).

The upshot of this account is that, if correct, the usual stories about individual cognitive adaptations or sudden bursts of innovation driving the evolution of modern humanity are at best incomplete. Apprentice learning requires more than individual cognitive adaptations. It also depends on "adaptively structured learning environments" that both retain prior innovations and generate fertile ground for new ones to be produced by novices and competent adults alike. A single innovation can take several generations to

[287] Sterelny (2012), §5.3.4

become reliably integrated into skills training. But apprenticeship learning plausibly explains both how novelty emerges and how it is retained and transmitted.

This point extends to the construction of the learning environment itself as well. Experts know how to create a proper setting in which novices can practice. Earlier in the evolutionary sequence, this begins gradually, presumably without a conscious intent to construct any such environment. Young novices happen just to notice things about how expert adults exercise their skills, but receive limited direct training. Nevertheless, they retain the more passive/contextual aspects of a craft, and perhaps learn from fellow novices as well. As craft sophistication increases, so too does the kind of agent who practices it, selected for greater dexterity in complex crafts. Once this sort of lineage begins to build up, new forms of learning emerge. These depend on "supervised and organized trial and error…in an environment seeded with props and other cognitive tools," and in which finished and practice variants of tools "are available as sources of inspiration and comparison."[288] As Sterelny puts it, children "learn by doing, but what they do is engineered by adult experts via their equipment supply."[289]

This evolution of pedagogy is crucial to evaluating TD's autonomy thesis. It suggests how the material culture plays a critical role in the learning environment that grows in influence over time. It provides a stable proving ground that helps both to produce and to work in tandem with other changing circumstances, such as the extension of both biological lifespan and childhood development. The more skill that is required for members of a band to contribute as competent adults, the longer and more intricate the training phase must be for skill acquisition. Adolescence emerges as a distinctly human life phase,

[288] Sterelny (2012), §5.2.3
[289] Sterelny (2012), §5.2.3

a new developmental period devoted exclusively to training. "Those changes in human life history, in turn, depend on ecological innovation…So human life history characteristics coevolve with technological competence and cultural learning. The technological and informational bases of cooperative technological foraging typically require deep educations."[290] The autonomy thesis invites us to infer that while technology's role is perhaps not privileged over that of social learning mechanisms or cognitive adaptations, it is at least equally important, and seems to play a key role in intensifying the feedback dynamic that has driven the evolution of modern human cognitive agency.

§5.2.3.2 Material Culture and Cognitive Niche Construction

The key difference between humans and our predecessors is not simply that our social learning methods are more faithful or efficient, but that the social and material environments we construct support "high-volume, high-fidelity cultural learning" and do so through the increasingly sophisticated organization of the learning environment itself. The feedback dynamic depends on the construction of an appropriate cognitive niche in which individual cognitive adaptations interact with environments that have been organized and increasingly optimized for social learning. The material environments themselves are engineered and refined over generations, accumulating ever more optimal resources for training agents to operate effectively in local ecologies.

It is evidence for the autonomy thesis that if groups are displaced suddenly, losing their material heritage, they often quickly lose (in a generation or two) those operational skills. The material culture—the technologies—accumulates to become a foundation for

[290] Sterelny (2012), §5.2.2, 45.

both cognitive development of individuals and cognitive adaptations across generations. Nobody set out to design it; it was likely a side effect of early incremental innovations.[291]

The construction of a cognitive niche is a long-term, multi-faceted project that exceeds the effects of neural adaptation. The entire ecology is affected. The effects of technology include not only impacts on how cognitive agents think and solve problems in their environments, but also on morphology in other dimensions. Features of our limbs, gait, and posture can be seen to reflect adaptations to particular types of tools. Marzke (2013) reviews the evidence linking manual morphology with the use of stone tools through successive iterations of this technology, and finds it likely that our hands have indeed evolved to wield them: "it is clear from the prehistoric record that hands were exposed increasingly to large, repeated, prolonged stresses associated with tool making and use of the tools."[292] Likewise, bipedalism has also been refined in hominins in connection with hunting practices, especially weapons. It is no real challenge to imagine that these features of the human body evolved in conjunction with the performance of these tasks, e.g. cutting and hunting, etc. Why shouldn't we think that cognitive performance also evolved to adapt to available material culture?

Effectively the claim of the autonomy thesis here is that the technological environment has selected for cognitive adaptations that tend to further intensify the gain and accumulation of additional innovations. "There will be selection in favor of mutations that increase the reliability and accuracy of learning from the parental generation (unless these mutations come with other, unaffordable costs). Such mutations can adapt mor-

[291] Sterelny (2012), §5.2.2
[292] Marzke (2013), 6.

phology as well as mind to the new technology."[293] The steady accumulation of stable innovation itself serves as the basic condition for further innovation, which then begins to select for particular cognitive adaptations in particular ecological niches. The effect intensifies through apprenticeship learning, engineering of the social learning environment, and further cognitive adaptation over time. But the material cultural basis is what appears to permit the cycle to intensify to the degree that it has in the hominin lineages.

§5.3. Creative Innovation

I said earlier that despite her bold attempt to break with the standard approach to thinking about the relation between society and its technologies, Preston cannot seem to avoid engaging with the old theme of TD. Although she does not invoke the language of TD or of autonomous technology in a direct way, it is clear that a perceived threat like TD does trouble Preston at a fundamental level of her discussion. She is quite sensitive to how her account of proper function "problematizes" individual agency, its structure, and the customary autonomy of human agents in casting them as products of material cultural propagation. She recognizes that the idea I have been calling TD's autonomy thesis is one that is plausibly and perhaps justifiably attributed to her theory.

To preemptively respond, I believe, to the kind of reading I have been defending, Preston also develops an account of the possibility of creative innovation in material culture. In this section I want to examine how she thinks about innovation, and the role it plays in her overall theory of material culture. In §5.4 I argue that a reading of her account on the basis of the autonomy these can accommodate this aspect of it.

[293] Sterelny (2012), §5.2.2

§5.3.1 System Function and Creative "Leeway"

Preston recognizes that her sociogeneric stance "problematizes individual innovation and autonomy by emphasizing the extent to which individual agents are products of their culture rather than suigeneric producers of it."[294] As I argued in the previous section, the purposes and intentions of agents get cast as products of the propagation of material cultural lineages through traditions of proper function.

Preston proposes to solve this problem by developing an account of how creative innovation is not only possible but essential to processes of reproduction. Her solution is to invoke the system-functional side of her pluralist view of function. Creative innovation occurs through system-functional departures from the standard proper-functional uses in which agents are trained. Individuals can exercise some "leeway" in how they utilize items of material culture even when their training in the use of these items follows tradition. Small departures from the proper functions of items accumulate over time, leading to larger shifts and new proper functions. "Material culture is misunderstood if it is viewed only in terms of proper-functional constraints on action and on the purposes people acquire…the phenomena of system-functional use show that material culture is even more fundamentally an opening up of possibilities for creative action."[295] Both sociogenerism/proper function and the possibility of system-functional creative departures are fundamental to explaining the propagation of material culture over time.

How do system functions accomplish this leeway? "System functions occur where an item of material culture is used for a purpose other than the one corresponding to its

[294] Preston (2013), 208.
[295] Preston (2013), 220.

proper function."[296] Agents are not utterly determined by those uses they learn as the proper functions of items; they can defy their training. Being trained to use a toothbrush for cleaning one's teeth does not mean one is also unable to use it to scrub nooks and crannies around the bathroom (after its life as a dental tool is done). Preston is pointing to a basic insight here about the relation between proper function and system function in the domain of material culture: "system functions are dependent on standard purposes, and thus ultimately on agents defined in part by configurations of such purposes…system functions depend on existing items of material culture that can be turned to account for purposes other than the ones corresponding to their proper functions."[297] Pluralism about function makes even more sense when we recognize how complementary these are.

The sort of innovation on offer here is one that eschews radical novelty in favour of incremental, cumulative shifts and variations. Most innovations are not significant but rather come as minor adjustments and subtle departures from proper functional usages. Her example of the electric guitar as a long-term, socially diffuse achievement brings out the key role of system-functional uses of various components such as microphones and amplifiers.[298] Building on her earlier discussion of improvisation, Preston argues that individuals can subvert their material cultural heritage through such innovative departures from proper function. This happens in the following way.

System functions are specified in virtue of the role some structure performs in a containing system or operational context. ("Structure" just seems to be a material affordance whose effects drive proper-functional reproduction of some type of item.) It follows that reproducing such a structure requires that "its situation in an appropriate em-

---

[296] Preston (2013), 207.
[297] Preston (2013), 207.
[298] Preston (2013), 217.

bedding system" also be reproduced, complete with its relations to and interactions with other components.[299] This stands in contrast to the way in which proper-functional structures get reproduced. Structures are reproduced in virtue of their proper functions when certain historical conditions are met, i.e. when an item belongs to a certain reproductive heritage. Preston reminds us that the distinction between proper function and system function rests on the fact that "what an item of material culture *can* be used for is not a completely reliable guide to what it *is* (or was) properly used for."[300] In neither case is function entirely dependent on structure.

For example, pipe cleaners were reproduced for many years in virtue of their function to draw coarse fibres through a narrow aperture, removing debris from the stem of a pipe. They were made with a flexible structure so as to allow for use with different types of pipes and smoking implements. Cleaning pipes was their proper function because it was this function that drove their reproduction.

Gradually, however, children began to use pipe cleaners in art projects, capitalizing on the flexibility of the stems and texture of the pile to facilitate a wide variety of other crafts. This new function—a system function—obviously caught on, and today pipe cleaners are available in many textures, sizes, and colours that have no added benefit in the original proper function of cleaning pipes. Going by sheer numbers, the craft function is arguably a new proper function of pipe cleaners since it now likely accounts for more tokens than cleaning does. (Are there more kindergartners than pipe smokers?) And this new proper function won out over other common ones, e.g. bundling or catching drips on

---

[299] Preston (2013), 195.
[300] Preston (2013), 196.

bottlenecks. Furthermore, it was accomplished by children and the adults who supplied them with craft materials.

In the reproduction of proper-functional items, there must be a tradition of "making" or tradecraft corresponding to the function in question. This in turn corresponds to the types and purposes of individuals generated by the material culture itself in perpetuating these lineages. And it is interaction with the material culture, and not just with other human agents, which drives this generative activity.[301] The questions about power, social regulation of individual behaviour in and through material culture, and agent autonomy which came up earlier thus receive a very strong reply from Preston. The worry "that the interaction of the individual with her material culture—an interaction that pervades action in general—is constitutive of the individual's purposes and overall character as an individual,"[302] and that thorough material cultural socialization constrains action, need not be utterly stifling. These constraints can, in fact, enable individuals to break away from convention and pursue a novel utilization that may or may not succeed.[303]

How has Preston managed to break the impasse between material cultural constraints and genuine innovation? Via her notion of creative leeway: "the leeway individuals have with regard to social constraints…is typically a matter of simple deviation from social norms rather than outright, politically charged resistance to them," even though "ordinary items of material culture sometimes mediate exercises of political and ethical power."[304] Artifacts still have politics, but those politics are not immutable. They can shift and respond to changing usage over time.

[301] Preston (2013), 204.
[302] Preston (2013), 209.
[303] Preston (2013), 212.
[304] Preston (2013), 210.

What makes system-functional leeway possible, according to Preston, is "the multiple realizability of function and the multiple utilizability of structure…Function and structure are mutually constraining, but also mutually underdetermining."[305] No proper function coincides exactly with all and only the full range of possible uses/affordances of an item. Coat hangers can hold mufflers up, and pipe cleaners can make very sturdy stick figures in children's collages. There will always be further uses of an item. Nor is system function wide open to things beyond the basic range of possibility enabled by the constraints drawn from proper functions. "Leeway" is a good term here, for it truly captures the idea of creative license that remains contained within boundaries that, while not explicitly specified, are nonetheless real.

§5.3.2 Incremental Deviations

The mechanisms by which system-functional utilizations are invented takes us back to the ubiquity of improvisation. Deviant use is also an essential part of our developing psychology, Preston claims. Citing psychological research on play and its importance for learning to problem-solve using creative solutions, she argues that system-functional departures are involved in the capacity for pretense. "Non-proper use is learned by children simultaneously with the learning of proper-functional use."[306] The deviant utilization of items of material culture reflects a natural creative inclination. During their training in the culture, people learn both what the customary proper functions of an item are, along

[305] Preston (2013), 211.
[306] Preston (2013), 214.

with common subversions of it, and in particular they learn methods of subversion if their training emphasizes creative engagement.[307]

Quite naturally, improvisatory strategies are among the central means by which system-functional departures occur. Whereas proper-functional use practices are implemented according to habit and regimen, system-functional ones are found through improvisation and, in particular, rely on the strategies of *appropriate-and-extend*, *proliferate-and-select*, and *turn-taking*. Appropriate-and-extend, for example, can involve taking items as they are, or modifying them to suit a new purpose. A pipe cleaner can be extended to use in crafts in its original form, or the fibers can be adjusted, it can be made coarser or fuzzier, its colour can be changed, etc. Some new uses do require modification, such as cutting holes in a pop bottle to use it as a bird feeder. Proliferate-and-select is perhaps the strategy most closely related to play in that children are observed to use items for many different purposes. Adults will also use items in many different ways before settling on any one particular usage. Turn-taking, of course, is thoroughly collaborative.

Preston is arguing that the constraints *themselves* are resources which enable individual departures to manifest as innovations.[308] Individuals exploit the mutual underdetermination of function and structure, finding holes and gaps with which to work against the confines established by proper functional lineages, yet always starting out from within them. Items of material culture would not seem to acquire system functions at all unless such novel uses were departing from an established lineage[309] System functions are discovered when new uses come about that do not coincide with proper function. On this

[307] Preston (2013), 215.
[308] Preston (2013), 212.
[309] Preston (2013), 207.

view, creativity occurs by quite gradual deviations.[310] For example, the invention of wristwatches first required a long process in which the efficiency of clockwork was refined, incorporating new mechanisms once scaled designs became available. Historically, the "invention" of the wristwatch occurred over many decades. Other examples are similar: the modern digital computer is arguably an invention centuries in the making, incorporating a good deal of the history of mathematics, physics, and electrical engineering. Boyd and Richerson give the interesting example of maritime magnetic compasses:

> First, Chinese geomancers noticed the peculiar tendency of small magnetite objects to orient in the earth's magnetic field, an effect that they used for purposes of divination. Then, Chinese mariners learned that magnetized needles could be floated on water to indicate direction at sea. Next, over several centuries Chinese seamen developed a dry compass mounted on a vertical pin-bearing, like a modern toy compass. Europeans acquired this type of compass in the late medieval period. European seamen then developed the fixed card compass that allowed a helmsman to steer an accurate course by aligning the bow mark with the appropriate compass point. Compass makers later learned to adjust iron balls near the compass to zero out the magnetic influence from the ship and to gimbal the compass and fill it with liquid to damp the motion imparted to the card by the roll and pitch of the ship. Even such a relatively simple tool was the product of at least seven or eight innovations separated in time by centuries and in space by the breadth of Eurasia.[311]

Both proper functions and system functions have tradeoffs. Proper functions give us items with reliable and efficient functions, but with a cost of socially enforced constraint on patterns of use. An agent is trained (apprenticed, acculturated) to propagate a given set of proper functions in step with the inherited lineages of a culture. Imagine asking for scissors to cut your steak at a restaurant, or refusing to wear any clothing in public when the weather is warm. Departures originate as system-functional uses, but these too exact a cost: wholesale repudiation of a given proper function is rare. Deviation from proper function is in general quite gradual. Preston thinks it can lead to genuine, widespread social change, but this can in turn also establish new lineages of proper functions that constrain anew the agency and purposes of individuals. There can be a shifting, cy-

[310] Preston (2013), 210.
[311] Boyd and Richerson (2005), 424.

clic interplay over time between proper and system functions, and this is an essential feature not only of Preston's pluralism about function but of her ambivalence on the question of individual agency versus the autonomy of material culture.

Preston emphasizes this interplay as it occurs during the training and acquisition phase of agent socialization and throughout individual development. She appeals to empirical work on play and creativity to support a claim that neither adherence to nor deviation from proper function is "primary" or "basic" but, rather, each is a flip side of the same coin, i.e. aspects of the same skill. Individuals learn how to depart from standard uses of items of material culture even as they are learning to adhere to them.[312] It is the relatively minor departures which lay the groundwork for major cultural changes in the forms of political resistance, design innovation, etc. Improvisation is a major catalyst for such change. Preston does not claim outright that social change *just is* improvised action and system-functional deviation from proper functional use. Her claim is that these are the basic phenomena which tend to generate it: "it is at this ground-floor level of interaction with material culture [i.e. system-functional deviation] that the basic activity patterns necessary for acts of political or social resistance on the one hand, or for participation in significant cultural innovation on the other hand, are established."[313]

§5.4 Responding to Preston's Stance on Innovation

Preston's stance on innovation is quite nimble and poses a genuinely compelling response to the TD troubles raised by her sociogenerism. To close this chapter I consider whether her response is enough. I argue that TD's autonomy thesis can accommodate the

[312] Preston (2013), 215.
[313] Preston (2013), 217.

sort of account of creative innovation we find in Preston. It does so because the capacity for improvisation does not necessarily confer credit for innovation on the improvising agent, and even when it does we have no reason to think this capacity is not also enabled by the agent's training in material culture.

§5.4.1 The Ambivalence of Preston's Stance

Preston's account straddles a fundamental ambivalence which, to her credit, she does not downplay but rather faces head on. The ambivalence stems from her socio-generic stance, on the one hand, and the role she ascribes to innovation, on the other. It amounts to her walking a difficult tightwire between claiming that the competencies of individual agency are formed by material culture and social training while also saying that those same competencies can generate new ideas. The most direct confrontation she gives with the account's ambivalence is probably in this passage:

> Our account of action and material culture will be of little comfort to [those who are committed to the Enlightenment ideal of suigeneric individualism]. The opposing socio-generic view we have argued for is much more in consonance with the anti-Enlightenment tendencies of hermeneutics and post-modernism. In particular, it depicts the relationship between individual and society as ambivalent…the individual is formed in and through her interaction with material culture, such that her goals, purposes, motivational structure, and so on, are constituted in part by this interaction. In this sense the individual is not autonomous, but rather heteronomous—regulated to her core by the goals, purposes and motives of others, especially as they are embodied in the proper functions of the material culture she inherits. On the other hand, this same material culture is the springboard for individual and group departures from proper-functional usage in the pursuit of local purposes and goals. But, these departures are more in the nature of spin-offs than of unique inventions from scratch…as far as material culture goes, individuals are definitely not just instruments in a fully orchestrated game. They have a say in the orchestration, and sometimes that say eventuates in significant cultural innovation. But any notion of the individual abstracting her self completely from her material culture, even temporarily, as the Enlightenment ideal seems to require, must be abandoned.[314]

So individual agents "have a say in the orchestration," and sometimes this can lead to a tipping point for significant cultural change. We individuals are more than "just

[314] Preston (2013), 221.

instruments in a fully orchestrated game." Preston is directly challenging a position like TD's autonomy thesis here, despite not identifying it in relation to known TD proponents. She wants to deny that the material and technological conditions in which individuals acquire cognitive and other forms of agency are so influential that they eclipse genuine creative novelty. Does she succeed in making her case?

One point to consider is that even if the training regimen of our material culture does result in "ambivalently skilled adults,"[315] i.e. adults who are competent both in following established proper functions and in departing from them, it is mostly true for most people most of the time that they do not make many such departures. People tend to stick to the customary uses of things; they are rather conservative about function. I think Preston agrees that departures are far less common among adults than among children, but this point could stand some more emphasis. What she is saying about system-functional departures is probably quite true in a general sense, but there is strong reason from categorization psychology to think that people are less liberal about function than she makes them out to be. Studies show that adults tend to categorize artifact functions on the basis of the demonstrated or inferred intentions of designers and other users, rather than on affordance (i.e., what an item could plausibly be used for given its shape, etc.).[316] Connecting system-functional deviation as she does with the ubiquity of improvisation serves, I think, to make it seem more common than it is.

Then there is the question of the purposes with which individuals manage to set out on their own, beyond the confines of their proper functional heritage. Preston says that people make the ideas for their departures through improvisation. This means that

---

[315] Preston (2013), 219.
[316] Cf. studies on artifact categorization cited in §1.2.

individuals don't have to be looking for novel utilizations, but they have to be capable of generating conditions in which they can be stumbled upon, and of recognizing them when they do occur. Recall that Preston thinks that while improvised action does not involve planful intentions, it is still intentional. We might say it's characterized by a form of tacit intention, but then the origin of these tacit intentions is unclear. Why shouldn't we say the deviant uses are formed on the basis of volitions, desires, etc. which derive from the cultural standards by which one was trained? Preston really does think "it is as correct to say that the purposes we have depend on the functions of items of our material culture, and are externally imposed on us, as to say that the functions of items of material culture depend on our purposes, and are externally imposed on them."[317] Individuals do not just come up with their purposes on their own, but rather they inherit them through e.g. apprenticeship learning. (She even cites Marx with approval on this point.) Most of what agents want to do with material culture is learned in training with it.

If I want to suggest that Preston over-emphasizes improvisation here, I ought to present an alternative source of departure by which to explain creative innovation. And here I could appeal to the affordances of material cultural items themselves. It is true that gains are had in the efficiency of technical and social affairs through the establishment of proper functional lineages. But sometimes the conditions for new efficiencies emerge. An engineer might realize that a frustrating piece of code is perfectly well-written in the high-level language but poorly suited to the machine code into which it gets translated. This little nudge might prompt her to reassess the high-level language. System-functional departures might not be novel innovations after all, but rather random gains that happen to catch on once they are exhibited.

[317] Preston (2013), 205.

Perhaps it is true that agents must provide the impetus, in the form of improvised utilization or just plain dumb discovery, for taking advantage of previously unnoticed system functions. I'm still not sure this is an appropriate assignment of credit. Is it too simple to wonder where agents get their capacity to improvise? Preston seems to think improvisation is our most fundamental capacity as agents.[318] This strikes me as a rather large claim, and I do not see what argument is given for it. It seems to follow from the fact that plan theories are inadequate, and so there *must* be a more basic, non-intentional capacity at work. But this rationale is almost entirely negative. Improvisation is a concept we use to designate what is essentially random behaviour that does not quite rise to the status of an action. It is true that we can marshal this undirected playfulness and it can serve as a valuable resource. But it is also true that many discoveries come from sheer accident. And if the capacity for playfulness is something engendered by material culture, as most agent capacities and purposes seem to be, then I do not think we should so easily accept that fruits of undirected action lie beyond the scope of the autonomy thesis.

§5.4.2 Autonomy and the Material Conditions for Cognitive Agency

In this section I argue that Preston's account of innovation is compatible with TD's autonomy thesis. The autonomy thesis says that the technological conditions of material culture shape human cognitive agency. Preston's theory of creativity is not enough to defeat this claim.

Preston argues for an incremental notion of innovation, and we saw that this is supported to a fair degree by anthropology and especially the apprenticeship model of

[318] Preston (2013), 43.

social learning. However, we can ask whether Preston makes more of the occurrence of system-functional departures than is warranted. By her own account, they happen largely through improvisation and are probably quite rare. Innovations are not the norm; most apprentices and even adult users never contribute major changes to proper functional lineages. The nature of innovation on this account is incremental, depending on the resources of improvisation and collaboration. These resources, by her own argument, provide no intentional innovation other than in the loose sense that agents can recognize benefits of novel utilizations and then capitalize on them.

Preston's position here casts the creativity of action as "specific to the functional aspects of material culture."[319] Hence the available functional aspects, combined with strategies gleaned from cognitive adaptations and social training, already bind the possibilities for creativity in a given community. Material cultural traditions in China or indigenous Australia just would not have produced any such item as a toaster. Even if they are capable of making items similar to the early American prototypes (e.g., small metal cages held up to flames), other required resources are not in place for later innovations in the lineage. The problem for would-be toaster inventors in the Chinese tradition is not simply that they can imagine a plan for the design of such a device, but are unable to realize it due to lack of material. The reason such devices do not appear in this context is that they are in fact quite difficult to imagine from there.

It is instructive to recall here the sort of response Dennett makes concerning freedom and caricatures of determinism, e.g. genetic determinism. "A proper human self is the largely unwitting creation of an interpersonal design process in which we encourage small children to become communicators and, in particular, to join our practice of asking

[319] Preston (2013), 223.

for and giving reasons, and then reasoning about what to do and why. For this to work, you have to start with the right raw materials."[320] Technology and material culture are also impersonal design forces that constrain human agency. But the autonomy thesis need not threaten human agency any more than the truth of the corresponding thesis about genes has threatened it. Harnessing these constraints is only possible once we have begun to appreciate that they are there.

Dennett's target, of course, is the agency involved in moral responsibility, but the point carries across. Technologies have clearly shaped human cognitive agency and social organization in crucial respects over time. But the gradual accumulation of material culture has also enabled sophisticated forms of action. The point does not even have to extend to the sort of innovation Preston argues for. She wants to undercut the idea that material conditions constrain our agency by casting these same constraints as a springboard for innovation. But this overstates the occurrence of such novelty. Most system-functional departures do not even require an appeal to improvisational techniques or resources. They can be explained in terms of multiple utilizability, of existing affordances at last being utilized by some agent. Preston goes further than is really required to defeat the autonomy thesis, and hence I submit that she misses the target.

§5.4.3 Autonomy and Agency

Here is where things stand. TD's inexorability thesis seems like a tractable claim, depending how flexible one is regarding cultural heritability. TD's autonomy thesis, as a

[320] Dennett (2003), 273.

claim about the material conditions of human cognitive agency, meets a subtle and compelling response in Preston's account of creative innovation.

Finally, we should face the following objection. Isn't it the case that either TD is a straw view, or that its tenable content is so mundane as to be unworthy of discussion? On the one hand there is no genuine threat that individual agency is determined by the material conditions of technology any more than it is by genetics, physical laws, or social organization. Preston shows that even if we ascribe stark efficacy to the material conditions in which our agent capacities are forged, those same capacities can alter the conditions for later generations of agents. But if we ignore the nominal associations of the doctrine's name (i.e., technological *determinism*), then all we are left with is a claim about the impact (presumably causal) that "technology" has on "society" broadly construed. And since there is of course some impact, the problem is idle. Hence either TD is a straw view or the content that can be ascribed to its central claims is trivial.

As I argued in the introduction, I do not think the effects of technology are always obvious, nor is it entirely clear how to investigate them. And I think that if we take away nothing else from the profound ambivalence of Preston's stance, we should accept that the material conditions of human life are a matter of deep interest and not simply an obvious truth that can be remarked upon and then put aside. Whether or not technology has an impact on agency and social organization is not in dispute; clearly it does. The extent of this impact, and how much room it leaves for what is familiar to us as agency, is not so clear. Are we left wondering about agency, about whether it has any true purchase?

If Preston's picture is correct, then perhaps human agency has always been to some extent an illusion. On the other hand, perhaps Andy Clark is right when he claims that, in the face of technology's onward march, "the kind of control that we, both as individuals and as society, look likely to retain is precisely the kind we always had: no more, no less. Effective control is often a matter of well-placed tweaks and nudges, of gentle forces applied to systems with their own rich intrinsic capabilities and dynamics. The fear of 'loss of control,' as we cede more and more to a supporting web of technological innovations is simply misplaced."[321] Dennett strongly echoes this sentiment: there is no problem to be reconciled between our frank sense of agency and the clear constraints from which it springs. Some find this cold comfort. In the next chapter I try to show just what can be so warm about it.

[321] Clark (2003), 175.

A Design Model for Cognitive Engineering

In this final chapter I confront some of the practical consequences of a technogenic account of cognitive agency. In §6.1 I motivate the normative dimension by examining the commitment to a narrow locus of control shared by most extended theories of cognition. In §6.2 I examine the recent appeal some of these theorists have made to responsibilist theories of knowledge to preserve their commitment. This gives me an opportunity to explore factitious intellectual virtue as a way to defend these sorts of appeals. And in §6.3 I argue that factitious virtue has several benefits as a normative design model for the practice of cognitive engineering.

Up to this point I have been arguing that the two basic theses of TD are in fact more plausible than the reputation of this doctrine would suggest. Technologies can plausibly propagate via the evolution of material culture. Technologies have also plausibly played a critical role in shaping the cognitive agency of modern humans. Whether this role is best characterized as constraining or enabling is perhaps a matter of perspective, but if it does come down to perspective then this has practical consequences for how we think about cognitive agency and also for how we design the technologies with which we work and live. In this final chapter I switch gears to discuss one of the most critical areas of design: cognitive engineering. Having arrived at a place where we can stipulate with relative safety that technologies do impact individuals in the respects indicated by the autonomy thesis, I want to look at the consequences this has for design. To do so, I consider Alfano's theory of factitious intellectual virtue as a model for cognitive engineering.

TD doubtless has its greatest purchase in colloquial fears about the threats of contemporary technologies. We live in the midst of unprecedented innovation both in scale and variety. Computers perform feats of calculation only theorized in past eras. Communication technologies connect individuals anywhere on the planet virtually instantaneously. Automated entities carry out financial and strategic decision-making, while robots interact with consumers in many facets of commercial life. Persuasive technologies nudge people into prescribed behaviours, from controlled intersections and self-driving cars programmed to sacrifice drivers to save pedestrians, to blatant psychological manipulation by machine interfaces and social media algorithms. Biotechnology is expected to explode any lingering dichotomy between "natural" and "artificial" systems, and the integration of social and technological design is already rather advanced.

Putting things so starkly can have a dizzying effect, and surely there is some overstatement here, but we are confronted by the basic fact that so far, nobody seems to have figured out how to think very well about technologies and their impacts. If the impacts are as great as a proponent of TD is inclined to think, then this is a scandal; even if the impacts fall well short of this bar, there is still much fundamental research to be done. In this chapter I address the practical side of our engagement with these impacts through design practices. I argue that for the practice of cognitive engineering, factitious intellectual virtue serves as a good model for the design, support, and enhancement of cognitive agency. Technological aids and enhancements are already all around us. Philosophers of mind have begun to notice their constitutive role in cognition. But it is only when the task of responding to an idea like TD is taken up that we can confront questions about what kinds of thinking entities we want to be and work to modify the enabling conditions.

In §6.1 I motivate the engineering task raised by TD's autonomy thesis in terms of an issue of the scope of cognitive agency noted by defenders and opponents of extended cognition in philosophy of mind. The commitment in this literature to a narrow locus of control for cognitive agents is probably ultimately a normative assumption rather than a well-founded description. But in responding to objections recent proponents have seen fit to appeal to responsibilist theories of knowledge. In §6.2 I examine this appeal and present Alfano's (2013) theory of factitious intellectual virtue as a promising response to the challenges it faces. Factitious virtue is conceived as a form of cognitive technology, and in §6.3 I argue that it should be adopted as a general model for cognitive engineering on the basis of its pragmatic benefits.

§6.1 The Scope of Cognitive Agency

Questions like "what is cognition?" and "how does it work?" are among the most contested cross-disciplinary issues on the open agenda of the sciences. Many writers have expressed doubts that there even is any substantive phenomenon available for empirical study, given how vague, problematic, and unscientific the proposed definitions have been.[322] Provoked by studies in several fields including cognitive science, work analysis, and psychology, Clark and Chalmers (1998) sparked one of the most intense confrontations with these issues ever observed in philosophy of mind. The ensuing fallout has been both revealing and instructive about the fundamental commitments of our conceptions of agency and of how it appears to manifest in relation to an environment.

From the standpoint of a concern with the design principles of cognitive engineering, the most critical intellectual commitment revealed by these exchanges has been the commitment to a narrow locus of control for cognitive agents. It has rarely been explicitly acknowledged that there even is such a commitment, nor is there any clear sense as to its status as e.g. a normative versus methodological commitment. Is there good empirical evidence that cognition hangs on such a locus? Or is the idea that it does a normative one,

[322] Cf. Serrano et al. (2014) for a recent survey.

and perhaps arbitrary and in need of defending? In this section I suggest that the parties to the debate about extended cognition go to great lengths to preserve a narrow locus of control without actually justifying this conception. Adopting a more deflationary conception can avoid this commitment as well as show that more familiar worries about discovering a "mark of the cognitive" are less troublesome than they have seemed.

Anyone who has ever lost a scrap of paper with an important idea scrawled on it can attest to the basic appeal of extended cognition. Thoughts originating in a cranium need not exclusively reside there. Nor is the exercise of cognitive control limited to those activities of which we are explicitly conscious at a given time. We perform many cognitive tasks without even being aware of them, such as when we drive while holding a conversation, or check messages while walking and talking. We also perform some tasks with the aid of environmental resources. Clark (2003, 2008) has described many of the studies and psychological experiments in support of this.

Clark and Chalmers proposed the parity principle to help identify when an external resource is contributing to a cognitive performance. According to their principle, any substrate, be it neural or otherwise, which plays some determinate functional role in realizing a cognitive process is *ipso facto* a part of that process. Adams and Aizawa have argued by counterexample against the parity principle, e.g. "How does a pencil know that 2 + 2 = 4?"[323] They suggest that proponents of extended cognition commit the now famous "coupling/constitution fallacy": mistaking tools *of* cognition, such as writing implements, *for* cognition or cognitive processing itself. Clark responds that it is as absurd to suggest that a bare neuron thinks as it is to suggest that a pencil does.

[323] Adams and Aizawa (2010).

> The appeal to coupling is not intended to make any external object cognitive (insofar as this notion is even intelligible). Rather, it is intended to make some object, which in and of itself is not usefully (perhaps not even intelligibly) thought of *as either cognitive or noncognitive*, into a *proper part of some cognitive routine*. It is intended, that is to say, to ensure that the putative part is poised to play the kind of role that *itself* ensures its status as part of the *agent's* cognitive routines.[324]

Clark's response cites a variety of ethological and psychological evidence in support of the parity principle's underlying functionalism. External resources like artifacts and other features of an agent's environment are causally coupled but remain cognitively neutral. The functional role a resource plays in an agent's cognitive routine, and nothing special about the resource itself, determines its cognitive contribution. (This insistence on cognitive neutrality is interesting and somewhat resembles what we find in a moral context, *cf*. Kroes and Verbeek: "Anybody who thinks that technical artefacts are morally good, bad or neutral erroneously takes technical artefacts as objects of moral evaluation instead of acts with or related to these artefacts."[325]) Clark suggests that taking external resources as either cognitive or non-cognitive is not even intelligible apart from the putative roles they perform. It is not even appropriate to treat them as cognitively neutral since this is also a sort of cognitive status, which they simply cannot hold unless in virtue of such a role. This response clearly affirms the agent as a locus of control.

Cognitive agency is in fact *narrowly* realized by extended systems. Proponents of extended cognition have largely been content to keep cognitive extension as a metaphor that really just describes functional instantiation. The agent as locus of control never actually seems to seep across its bodily boundaries. Cognitive agency on the locus of control conception has remained firmly intracranial. What keeps it there is a concern that an unchecked extension of the cognitive beyond cranial borders cannot be curtailed without

[324] Clark (2008), 87.
[325] Kroes and Verbeek (2013), 3.

some satisfactory answer about what counts as a cognitive process.[326] If we say that cell phones and autopilots perform cognitive tasks, why not pencils and bicycles too?

Wilson most forcefully articulates this "*narrow subject/extended systems*" view.[327] Using an analogy of extended digestion, he argues that body-bound agency with an extended reach makes intuitive sense. But his analogy argues from an alleged fact about an organic process—digestion—to a conclusion about a process—cognition—that is active in a very different way and is said to involve a locus of control. There is no reason to appeal to this notion of a locus of control except to account for agency, something usually defined in terms of self-control. Digestion, by contrast, is a rather passive or automatic process, like breathing or sleeping, and would not normally be described in this way. Does it suggest that partially chewed food fed to infants and baby birds would also count as a case of extended digestion?

Perhaps sensing the weakness of the digestion analogy in a context concerned with agency, Wilson makes a sudden appeal to action:

> Does action stop at the skin? The intricacies of action theory to one side, there are good reasons to think not, including the idea that a skin-bound view of action would leave us with an impoverished view of what agents do, one that confuses action with mere movement. Those holding the commonsense view that agents do things that *extend into the world*—like making a cup of coffee, driving a car, or writing a letter—seem to feel no compunction in appealing to a regular bodily-bound agent for such *extended actions*. Here a *narrow agents, extended actions* view is a natural default; the reasons one has for accepting extended actions do not challenge the status quo about the boundaries of agents themselves.[328]

So according to Wilson, narrow agency sits as a locus of control over an extended cognitive system. Extended cognition need not imply extended agency. But just why should we think that a narrow locus of control is the most reasonable default? Of course it seems

326 Allen-Hermanson (2013), 793.
327 Wilson (2004), 142.
328 Wilson (2014), 26.

plain that pencils lack agency. However, it also sometimes seems like unruly mobs have minds of their own, or that complex technical machinery is often able to fix performance glitches all on its own. What compels even these most radical proponents of extended cognition to insist on a narrow locus of control? Wilson sees the dispute between Clark and Adams and Aizawa as trading partly on a mereological equivocation.[329] Clark argues the pencil is a putative part of an agent's long division, Adams and Aizawa accuse Clark of making the pencil part of the agent, and Wilson claims they are conflating mereological and functional roles. But all parties appear to agree that the *agent* as locus of cognitive control must not be located outside the brain.

Why is there such an axiomatic commitment to a narrow locus? Perhaps partly because it is difficult to imagine what a wide locus of control would be like. If it were easier to imagine, objections like bloat and the coupling/constitution fallacy would fail to get as much traction as they have. Recently, proponents of extended cognition have doubled down on the narrow locus by appealing to responsibilist theories of knowledge. The thoroughly narrow notion of cognitive character is cited as a promising avenue of reply to these objections of promiscuous agency.

§6.2 Appealing to Responsibilism

The appeal to responsibilist theories of knowledge is an interesting development in the study of cognitive agency. In this section I point out that any such appeal faces the same challenges as have been put to these theories. In particular, the situationist challenge to cognitive character is a serious problem. But the response to this challenge found in

[329] Wilson (2014), 26-27.

Alfano (2013) also suggests an intriguing approach to cognitive engineering. His notion of factitious intellectual virtue will stand, I argue in §6.3, as an effective model for the design of cognitive technology. One of its chief benefits is that it provides room to move away from the axiom of a narrow locus of control.

§6.2.1 Extended Cognition and Epistemic Responsibility

Why is an appeal to epistemic responsibilism by proponents of extended cognition significant for considering TD or its autonomy thesis? The appeal reflects a strong desire to preserve a "narrow locus of control" conception of agency, even in a context where the driving thought is supposedly that cognitive processing is wide. If the study of cognitive agency and its material conditions suggests that technologies play both a formative and a constitutive role in that agency (i.e., the autonomy thesis), then this could mean that the proper scope of the locus of control is wider than typically thought. It could also mean that the very presumption of a locus of control is erroneous. Hence an appeal to responsibilist theories in the philosophical study of cognition appears prima facie to be a step backwards next to conclusions suggested by the study of material culture.

Take, for example, Roberts (2012). Roberts wants to avoid the problem of cognitive bloat. He appeals to the responsibility agents can take for exercising their own cognitive faculties. This gets around both the constitution problem and the promiscuity of cognition. By invoking the skilful exercise of a cognitive faculty, such as belief-formation in reading the newspaper or basic calculation when preparing one's tax return, no worry about the constitutive role either of the printed word or of the calculating device arises because the agent herself initiates and achieves the cognitive performance. At the same time, this response keeps external resources like artifacts in their roles as merely putative parts of a cognitive system for which a subject personally takes ownership. “Such resources [as artifacts] may assist the subject in the course of intelligent behaviour, but when they do not fall under her responsibility they do not extend her mind.”[330]

Now in Roberts’s case, the notion of responsibility he invokes is not intended to be as strong as the robust character traits of responsibilist intellectual virtue.[331] His aim is “to provide an account, couched in personal-level language, of what it is for a cognitive activity to be under the responsibility of an agent.”[332] He focuses on the exercise of cognitive faculties rather than on stronger forms of character. He argues that “norm-countenancing is what marks the difference between the (mere) reliable production of a belief of a particular sort, and the successful exercise of a cognitive faculty.”[333] To countenance a norm is to enact some bit of procedural knowledge—i.e., a plan. This plan must be seen to reflect the cognitive norm governing success in the performance of the particular cognitive task.

For example, an agent is getting ready to file her taxes. She installs a program on her computer to take her through the process of preparing an accurate return. If she uses the tax software to prepare her tax return based on planful procedural knowledge she has herself acquired, e.g. in her training in an accounting course, then she can be said to be countenancing the relevant norms (e.g., claiming appropriate expenses and declaring taxable incomes). She is using the tax program simply to expedite a procedure she understands on the basis of past exposure to successful practice in this domain. If, however, she

[330] Roberts (2012), 143.
[331] Roberts (2012), 142.
[332] Roberts (2012), 134.
[333] Roberts (2012), 138.

is someone with no such training, who relies entirely on the software as an external resource to guide her through the complex process of preparing her tax return, then she is not countenancing the relevant norms. Hence she is not responsible for the cognitive performance of filing an accurate tax return. Furthermore, the external resource does not become "transparent" to the second agent because she has not engaged with it in a way that counts as an achievement on her part. She does not manage to conscript the resource into an exercise of her cognitive faculties; rather she merely offloads the task onto the software. Hence she does not deserve epistemic credit for preparing her tax return.

Now the agent with training in accounting may not be countenancing all the intellectual norms involved in the use of the tax software. For instance, she may not countenance mathematical and computational norms that have been used to design the software or the machine that runs it. Roberts argues that at this level of responsibility, only norms of testimony, by which a subject has reason to trust the mechanics of a device, need to be countenanced.[334] So the first agent need not countenance norms of computing even though the second agent is expected to countenance norms of accounting. The reason is that success in a given intellectual domain ought not to rely purely on testimony, even though success there might involve auxiliary norms countenanced via testimony.

This appeal to responsibilist accounts of knowledge to prevent unwanted slippage of cognitive processing and attribution of agency is fascinating from a standpoint of the study of the technological material cultural conditions of cognitive agency. It appears to be motivated by an axiomatic commitment to a narrow locus of control. Since the review of such studies in the previous chapters suggested, in line with TD's autonomy thesis, that the locus of control is in all likelihood much wider than has been presumed, this de-

[334] Roberts (2012), 142.

velopment in the defense of extended cognition is somewhat perturbing. However, the appeal these theorists have made is not without its own challenges. And it is in the response to these challenges that I think a way forward for the development of design principles for the practice of cognitive engineering comes to light.

§6.2.2 Factitious Intellectual Virtue

The notion of factitious intellectual virtue is a response to an objection faced by responsibilist theories of knowledge and cognitive character: the empirically mounted situationist challenge to responsibilist intellectual virtue. Alfano (2013) responds to this challenge by presenting evidence that intellectual (and moral) virtues are attributes/traits/habits generated through (and hence responsive to) mechanisms of socialization and psychological development. Alfano argues that we can treat these mechanisms as forms of cognitive (and moral) technology with which to bolster desired cognitive (and moral) attributes. Let me explain how factitious virtue works in response to the situationist challenge before examining its merits as a model for cognitive engineering (in §6.3).

The situationist challenge attacks the possibility of cognitive character on empirical grounds. Both reliabilist and responsibilist virtue epistemology face a version of this challenge. Responsibilist knowledge, for example, "is true belief acquired through the exercise of such virtues as flexibility and creativity,"[335] yet the acquisition of belief can be shown to be sensitive to epistemically irrelevant factors. For instance, mood affects intellectual curiosity, flexibility, and creativity according to modified studies on two psy-

[335] Alfano (2013), 122.

chology staples known as the candle task and the remote associates test.[336] The candle task measures flexibility and creativity: subjects presented with a matchbox, a candle, and a thumbtack are asked to arrange them so that the candle does not drip when lit; being able to solve the problem depends on how the items are presented. In the remote associates test, subjects who are creative, flexible thinkers are able to generate companion words that create common phrases when added to words on a list. One 1987 study alone found that altering the mood of participants before administering the candle task or remote associates test, by showing them a comedy film or giving them candy, significantly improved the rate of successful performance by over 65% in each case.[337]

Faced with evidence of this strength, responsibilists must therefore either restrict their trait attributions to quite narrow traits, which are not cognitively admirable, or reject traits outright, which is tantamount to skepticism about responsibilist knowledge. Most persuasively against responsibilism, Alfano spells out the situationist challenge to intellectual courage. Here such classic social psychological results as the Asch effect and the Milgram experiments cast doubt on the claim that people actually possess this trait. "Rather than being *intellectually-courageous* or even *intellectually-courageous-to-speak-in-the-face-of-social-disapproval*, most people are at best *intellectually-courageous-to-speak-unless-faced-with-unanimous-dissent-of-at-least-three-other-people*."[338] The kinds of traits we associate with cognitive agency, especially intellectual courage, appear simply not to hold up under empirical scrutiny.

Consulting related literatures from educational and social psychology, Alfano develops a response to meet the situationist challenge that proposes how to generate what

[336] Alfano (2013), 120-121.
[337] Alfano (2013), 121.
[338] Alfano (2013), 135.

he calls *factitious virtue*: “For there to be factitious responsibilist virtue, attributions of intellectual motivations [must] function as self-fulfilling prophecies.”[339] The idea here is that public attributions of desirable character traits generate expectations that, like a publically enforced placebo, tend to produce target traits. He characterizes factitious virtue as a “technology” because the idea is to utilize social conditioning as an external aid. The results are striking: huge increases in IQ scores are observed when students are praised and treated by teachers as hard-working, determined, etc.

Alfano’s basic strategy here is to co-opt the situationist challenge itself and show how it already includes a path of response by which cognitive virtues can be enhanced or even generated virtually on demand. Much like how Dennett thinks that the bogey of determinism actually reveals causal mechanisms that enable agency and ultimately help us to design new and better ways to be morally responsible agents, factitious cognitive virtue also shows a new way forward. The result, for both ethics and epistemology, is a “holistic calibration” of normative theory, empirical psychology, and technological method.[340]

For cognitive agency, the aim is simply to utilize features of social situations to encourage people to act in greater accordance with desired intellectual attributes.[341] It may be a fiction at the time of the initial attribution, but people do tend to conform to these attributions with prolonged exposure. Alfano proposes that there are likely two distinct psychological mechanisms that explain how factitious virtue is possible: self-concept and social expectation. “People are averse to cognitive dissonance—acting contrary to their images of themselves—so if someone begins to think of himself as open-

[339] Alfano (2013), 159.
[340] Cf. Alfano (2013), 203.
[341] Alfano (2013), 158.

minded because he has been labeled thus, he will act open-mindedly to maintain his self-concept. In addition, people are typically averse to disappointing others' expectations, so when someone is publicly labeled conscientious, she will be more inclined to act the part."[342] Would-be social and cognitive engineers can exploit each of these mechanisms with relative ease.

For generating factitious responsibilist intellectual virtues, the two key conditions are publicity and plausibility of the attributions. The publicity condition is required because people simply don't like to disappoint others, and this is a powerful motivating impulse that can mitigate situational influences. The plausibility condition is required because the attribution cannot be too extravagant or incongruous with the subject's existing self-concept. It must be an attribution the subject thinks him-/herself capable of actually achieving. If these two conditions are met, the evidence suggests that the subject will come to conform to the attribution.[343]

Alfano closes with some enticing comments about the social nature of intellectual virtue and the respects in which its achievement can depend on one's environment, culture, and other such factors. Sticking to a traditional conception of virtue as a robust global trait is still an option, of course, even if it is susceptible to the situationist objection. In that case, factitious virtue can at least provide a band-aid. But he suggests that there might be an opportunity here to radically rethink both moral and cognitive virtue and to revise normative theory: "we could reconceptualize virtue as a triadic relation among an agent, a social milieu, and an environment, which would then make the social expectations (and the potential for common knowledge of those expectations) that under-

[342] Alfano (2013), 160.
[343] Alfano (2013), 171.

lie factitious virtue part of the metaphysics of virtue, rather than a situational influence that induces virtue."[344] Perhaps cognitive character traits can and ought to be pursued through external aids and, in particular, technologies. The design of the technological environment would then be part of the proper cultivation of cognitive virtue, in addition to its social and other aspects.[345] This makes factitious intellectual virtue an ideal model, in my view, for the practice of cognitive engineering.

§6.3 Factitious Intellectual Virtue as a Model for Cognitive Engineering

In this section I argue that factitious intellectual virtue could serve as a model for more direct applications of technologies in cognitive engineering. The worry behind TD—especially the autonomy thesis—is that the impacts of technologies are a threat to how people live their lives as agents. This worry runs from the moral domain right through to the cognitive domain. People do not want self-driving cars to decide whether their lives ought to be sacrificed in a collision any more than they want circuits and electrodes implanted in their brains by the state for tracking or thought control. These are outlandish examples, to be sure, but the technologies are real enough. Not only moral decisions but also rather sophisticated cognitive and epistemic ones are being off-loaded onto engineered features of the material environment.

Many such automated decision support systems are justified on the grounds that their design actually helps individuals to be better agents. Persuasive technologies are perhaps the most prevalent example. Persuasive technologies are items or interactive features of environments that compel users to perform desired actions. Their designs rely on

[344] Alfano (2013), 183.
[345] Alfano (2013), 178.

behavioural psychology to elicit certain behaviours in users. People can buy alarm-equipped locks for their refrigerators to prevent overeating, or install breathalyzers on their cars to prevent drunk driving. The speed bump is a simple example with a moral twist.[346] Other examples include: power cords that shrivel up like a writhing worm when the devices they power draw electricity for too long, to encourage more modest energy consumption; "glowcaps"—pill containers that light up to remind patients when it is time to take their medications; software that integrates with your bank's system to prevent you from exceeding a set limit on purchases, to help users save. In computing, examples abound: users of Google devices regularly receive automated alerts to remind them of appointments or news topics, to send messages on their behalf, etc. Most robots that are currently available incorporate at least some persuasive technologies, either to guide interactions with human technicians or to facilitate the actual service they provide.

Such forms of design require more than only ethical attention. There are important questions about how we want to interact with our technological environments, e.g. what kinds of things we want them to do for us and to us. How much help do we want making day-to-day decisions? A lot? A little? People today can feasibly set up an entirely automated lifestyle, with transactions to receive funds and use them to pay bills, with food and other services delivered according to a computer-generated schedule, and with mobility arrangements made on demand or in advance. This is not a point about the capacity to do things today that were not possible in earlier periods; one has always been able to use money to lighten one's load. But the extent to which this is now possible, and the sheer range of kinds of things that one can off-load onto technologies, has made it clear that the outsourced nature of cognitive performance was not fully appreciated in previous times.

[346] Verbeek (2011); Kroes and Verbeek (2013), 20.

Hence the assessment of TD raises a philosophical question about the design of agency: if human cognitive agency is radically open to design and redesign, what kinds of agents do we want to be, and what principles should guide this engineering?

§6.3.1 Cognitive Engineering

To give some orientation for thinking about design in cognitive engineering, perhaps it will be prudent to briefly examine a recent argument about the practical implications of this work and of the competing conceptions of cognitive agency that are available to philosophers. Blomberg (2011) argues that what he calls an expansive but deflated conception of cognition makes the most sense from an engineering standpoint. A decision to restrict the class of cognitive systems could have important ramifications for design: some philosophers and cognitive scientists have argued that cognition also occurs in larger systems of which individuals are only parts.[347] If endorsed, a restriction on cognitive engineering could be harmful if the bounds of cognition turn out to be more fluid than was thought.

Blomberg considers three competing frameworks in cognitive engineering: distributed cognition; joint cognitive systems/cognitive systems engineering; and cognitive work analysis. He compares them with respect to their practical implications for cognitive engineering. The first is just the familiar idea of extended cognition construed as a design approach. Blomberg describes it in terms of its focus on socio-technical systems that can sometimes generate goals not shared by individual human agents within the system. This approach also stresses the task of collecting ethnographic data to inform design decisions

[347] Blomberg (2011), 86.

about cognitive routines.[348] As a design framework, distributed cognition is perhaps even more radical than its formulation as a theory of mind, since there can be a tendency to shift the locus of agency depending on the goals of a cognitive system.

The other two frameworks are even more radical. "The emphasis in the joint cognitive systems approach is on understanding human–machine *coagency* rather than human–machine interaction," which has been the traditional focus of cognitive engineers.[349] Rather than treating humans and technologies as functionally distinct, emphasis is placed on the functional integration and unity of whole systems. This means that the information processing picture of human cognition one finds lurking behind most philosophical theories of mind, including those of extended cognition, falls away in favour of a more abstract view about the nature of activity as a way to cope with or reduce complexity.[350] A problem about how to pound a nail using a hammer might be expressed not in terms of how to realize the goal "driving a nail" by formulating a plan to use the tool, but instead of how the arm/hammer hybrid will drive any loose nails in its vicinity. Cognitive problems are seen not as how to realize a goal using available resources of information processing patched together functionally, but rather as questions about how to simplify the features of an environment. A loose nail is a complexity waiting to be reduced.

Blomberg argues that there are practical ramifications for cognitive engineering that differ between these frameworks. He focuses principally on function allocation: how should distinct functions be divided or shared between human operators and automated counterparts? The question itself reveals a certain bias, he suggests. Whether the system is viewed as a functional coupling with a designated locus of control or as a composite

[348] Blomberg (2011), 91.
[349] Blomberg (2011), 93.
[350] Hollnagel and Woods (2006), 21.

co-agent greatly impacts how task success is measured. On the more integrated sort of framework, cognitive functions are not necessarily performed using interchangeable components. One could feasibly drive a nail using a screwdriver, by hitting it hard enough, but it could injure the hand that is coupled to it, and it would be a far less efficient solution to the problem than the hammer. Human drivers do not always perform as well when placed in an unfamiliar vehicle, even if all the controls operate in the same basic way. The second and third frameworks emphasize the potential risks in a mere functional coupling view like distributed/extended cognition, as opposed to a more integrated stance, and all three oppose the more common internalist view that all that really matters for cognitive performance is contained with organic brains.

The view of cognition Blomberg advocates he calls expansive and deflationary because, on the one hand, it is not internalist and, on the other, he does not think there is any ultimate criterion by which to distinguish cognitive from noncognitive processes apart from primarily pragmatic considerations. The question of whether a system or aspect of a system "is cognitive" can be treated as a question of design aims. Are organisms cognitive systems? Are computers? Or large socio-technical systems? If treating them as such can yield better methods for design and manipulation, then the label is justifiable. This is not unlike a cognitive corollary of Dennett's pragmatic rationale for his intentional stance.[351] And it does not rule out the possibility that competing notions of cognition, e.g. a more strongly internalist conception with a clearly delineated locus of control versus the composite co-agential conception, would be appropriate depending on different design aims in different circumstances.[352]

[351] Dennett (1987).
[352] Blomberg (2011), 99.

§6.3.2 Design Lessons from Ergonomics

Now that we have a brief orientation, let's talk a little about the design challenges faced by cognitive designers. The theoretical and empirical study of cognitive engineering in general is known as ergonomics. Two major sets of results of interest to philosophers investigating cognitive design are the studies on automation and on decision support systems. Research on automation reveals surprising forms of risk associated with such systems as automated aviation and Advanced Driver Automated Systems (ADAS), which are among the most heavily automated hybrid systems in current application. Other studies on these topics examine automated safety systems at nuclear facilities and energy production sites, operational drones, robotics, space and deep-sea exploration, etc.

One result simply confirms the natural suspicion that overreliance on automation has adverse effects. An alarming number of plane crashes in the past decade have been attributed to degraded failure response by pilots in situations where an automated system failed or was automatically deactivated, only to have the pilot who reasserted control of the aircraft fail to execute a correct action.[353] A similar result has been found in studies of self-driving cars.[354] In general, researchers have tended to conclude that since the returns on automation diminish in a hybrid system, there are two broad design strategies. Either designers can pursue maximal automation (e.g., remote and drones) or they can withhold automating system features so as to keep human contributors more engaged rather than less. The key is to avoid designs that produce hypovigilance, which can lead to catastrophic failure. Automation bias is, however, an ever-present danger in design.

[353] Hancock et al. (2013), 10.
[354] Vanderhaegen (2012).

Even more pertinent results are found in research on decision support systems, which studies the design and use of cognitive aids for a variety of judgement tasks. One aim is to introduce debiasing methods into existing decision-making strategies. Another is to improve the efficiency of existing heuristics. Lehto et al. (2012) outline several approaches along these lines. "Genetic" algorithms use replication and rapid testing to eliminate large areas of a search space for a given problem, thus mimicking evolutionary processes. Another approach uses fuzzy logic to design decision supports that better handle uncertainty, especially on budgeting tasks. A third approach is the use of software learning agents, such as high-frequency traders on the stock market or gate assignment programs in airports. These systems make real-time decisions faster than human controllers or even interactive systems can respond to complex changing information. Some such systems use established parameters while others are more flexible, being programmed to adjust their own parameters in response to other automated systems with which they compete. The "flash crash" of the commodities markets on May 6, 2010 was a direct result of such behaviour in high-frequency traders.

One effect reported in this research on decision support systems is that the adoption of a support system tends to encourage users to adapt their own strategy selection process to the range of available aids in an attempt to reduce cognitive load.[355] In other words, human users seek to minimize their own cognitive effort even at the expense of the quality of the decision process. "When designing [decision support systems], effort minimization should be given considerable attention, as it can drive the choice of decision strategy, which in turn influences the decision accuracy."[356] A result in Todd and

[355] Lehto et al. (2012), 225.
[356] Lehto et al. (2012), 225.

Benbasat (1999) suggested that in general, people prefer not to make choices even if efficiency gains are available. They compared the effect of incentives versus cognitive effort in different strategies and found that even incentivized strategies for performing a task were only adopted when there was greater support (requiring less effort). We might infer that people are predisposed to rely on forms of procedural knowledge where it is available, and to utilize decision support systems that embody procedures when possible, since they do so even when it means they end up using a less efficient method.

These results indicate some basic points to consider for cognitive design. We can make recommendations like “don’t overdo automation” and “to increase adoption, get people to follow an available procedure,” but these maxims can be helpful or harmful depending on the needs of the situated task. For example, in software engineering a checklist for code review can help developers remember what tests to run on a new piece of software, which can expose bugs more effectively. Professional developers can be expected to know what tests to run, but often they will skip ones they deem too routine. Using a checklist and encouraging reliance on it can improve the cognitive task of code development. In trauma resuscitation work, following a prescribed routine often makes the difference between life and death of a patient.

Both automation bias and its opposite, procedural underreliance, grow from the same root and generate the same kinds of cognitive errors. Placing too much trust in automation makes human operators prone to trust their own judgement less and also to fail to respond when the automated component fails, whereas placing too little can increase the incidence of human error in system performance. But of course it is also not true that there is an ideal balance to be struck. Following the line suggested by Blomberg, we might wonder whether the design principle at issue here is making the system *either* human-centred *or* machine-centred. Most automation design has adopted the former as an entrenched design principle according to which final decision authority must always rest with human operators, even at the expense of machine performance. For example, letting humans override automated driving systems is deemed appropriate even though it will almost certainly produce outcomes in which pedestrians are killed who might otherwise have been saved had the automated driving system been in control of collision avoidance.

Inagaki (2008), for one, cautions against accepting the principle of human-centred design, arguing that "it is not wise to assume, without careful analyses, that human operators must be maintained as the final authority at all times and for every occasion."[357] He claims that a more adaptive approach to automation design is superior because it is a more flexible principle for guiding hybrid system dynamics overall. What this adaptive principle means for the design of cognitive agency is, I suggest, an affirmation of Blomberg's more pragmatic approach. Rather than viewing cognitive engineering and the design of heavily automated supports as a problem about how to shift back and forth between the human locus of control and a machine locus, we can see control itself as a system function that is likewise subject to automated rationing.

§6.3.3 Factitious Virtue and Cognitive Engineering

The overriding consideration in favour of factitious virtue as a design model in cognitive engineering is that it gives us exactly this kind of flexible approach to the locus of control. As we saw above, the appeal to responsibilism to preserve a narrow locus

---

[357] Inagaki (2008), 165.

faces a situationist challenge. Factitious virtue offers a solution: we can co-opt situationist impacts on cognitive character to make design changes in the social and material environment that generate desired cognitive traits. A quite pragmatic conception of cognitive agency is the main result of this model.

It is possible to design cognitive technologies in many ways. Their design can reflect many attitudes and values about how decisions should be made in a particular set of circumstances, and how people ought to be encouraged to think about challenges they face or problems they must solve. This is particularly so when considering how best to guide the design of our own cognitive scaffolding. It might strike us as obvious that a design framework should be formulated and applied so as to benefit human beings as much as possible in hybrid and technologically mediated circumstances. For example, most of us would want our work environment, if it were heavily machine-laden, to also be designed for our own comfort and ease of use, even at the expense of the functionality of some machines. But we can easily imagine scenarios where such priorities are not held, e.g. in which human workers are slaves and their comfort unimportant to designers. Such malicious designers might place a higher value on machine functionality at the expense of worker comfort, or on overall hybrid cost-effectiveness. (We would of course abhor this design principle on moral grounds, but this does not mean it is a poor design. The stacked design of industrial chicken coops is morally abhorrent yet still a clever layout.)

Most work in cognitive engineering frames these design issues in human-centred terms, and of course this is quite natural. So we speak of technological *aids* and task *support* systems, always keeping in mind that the human agent is the presumed locus of control in a cognitive work environment. But there are clearly many circumstances where it can make sense to identify a larger system as a locus of cognitive control, or perhaps where it makes sense not to identify a locus at all. When a social organization must make a decision about an action to take, such as a government, corporation, or other institution, it is often individuals who decide on the basis of conferred authority. But it is also reasonable to think that such entities could utilize automated decision systems to replace the need to appoint a potentially unreliable human proxy. And we can also imagine fully distributed automated decision systems, with no true locus of control, designed perhaps on the model of peer-to-peer networks.

Factitious virtue lets us take our cues for what counts as a cognitive trait partly from our own design aims and from the environment itself. What makes it attractive as a design model for cognitive engineering is how flexible the criteria for cognition are. They can be dictated by the situation and designated aims, and once the criteria are known they can be enacted. Cognitive engineers must face the problem of reinforcing a locus of control because this is what other design principles, chief among them human-centred design, tend to demand. But these demands can bog systems down. Having a locus of control is extremely useful in many situations, but this does not mean that we should expect it to be useful in all situations or cognitive environments. The challenge is to come up with a model or principle for design in cognitive engineering that lets the nature of a situation partly determine what counts as a cognitive solution.

An objection to this application of factitious virtue might be that the theory itself is a flawed conception of virtue. Indeed Alfano responds to this criticism in his defense of the theory, first in connection with the charge that factitious moral virtue is a “noble lie” and not true virtue. Initially at least, it involves a deception. A parallel worry exists for intellectual virtue in that it also appears rooted in dishonest representation of oneself in the eyes of others.[358] Applying factitious virtue to bolster cognitive agency is like giving a placebo to someone sick: it can make them feel and even perform better, but the effect is largely illusory. The real problem of how to underwrite genuine agency remains and it cannot be solved by appeal to false virtue. Virtue just is the sort of thing that requires real features of character. It must not be facilitated through the aid of technological design. And it must inhere in an agent who possesses an authentic locus of control.

My response is just that this point reflects a conception of virtue which I, like Alfano, have no qualms about defying. This is one way of conceiving character, but it leaves itself open to situationist objections and does not even seem to take the challenge seriously. For another example of a more flexible model, take Morton's view: intellectually virtuous agents are often required to compete against their own training. On this view it is the appropriate call of a trait or application of a capacity that is praiseworthy, rather than its possession: "the characteristic we praise is not so much possession of general capacity C as application of C when it is appropriate…Note how situationist worries and an emphasis on simultaneous sensitivity to the environment and to the state and aims of the agent coincide. Both considerations push us towards taking intellectual virtues as features of the way capacities are mobilized in particular circumstances in the service of particular aims."[359]

The cases in which this comes through most clearly are those of paradoxical virtues. Situations sometimes call for contradictions to be tolerated, or for evidence to be ignored, or for bias to be embraced. The idea is that paradoxical virtues can require that

[358] Alfano (2013), 176 (ch. 7 §4).
[359] Morton (2012), 65.

we compete against our own cognitive training, and hence that which particular virtue is called for on an occasion is not a matter of some higher excellence but rather of the profile of a particular person faced with a particular problem. “Virtues come into their own under the following circumstances. There is a type of situation, to which a person can react in a number of ways. Some of these ways are good in some respect, and a person can have a capacity to react well, without much deliberation. The reaction is best, for this person in this situation, when it is produced by this capacity rather than in some other way. Virtues are among such capacities.”[360]

On this model, virtues already are and always have been aids of character. Virtues are “are capacities that aid our possibly misguided functioning.”[361] Being able to call on the appropriate ones in the appropriate situation is what agency amounts to. And this is exactly what cognitive engineering can achieve when it follows the model of factitious intellectual virtue.

[360] Morton (2012), 63.
[361] Morton (2012), 66.

References

Adams, Frederick R. and Kenneth Aizawa (2010). "Defending the Bounds of Cognition," in Menary, Richard (ed.), *The Extended Mind*. Cambridge, MA: Bradford Books/MIT Press.

Alfano, Mark (2013). *Character as Moral Fiction*. Cambridge, UK: CUP.

Allen-Hermanson, Sean (2013). "Superdupersizing the Mind: Extended Cognition and the Persistence of Cognitive Bloat" in *Philosophical Studies* 164, 791-806.

Bimber, Bruce (1990). "Karl Marx and the Three Faces of Technological Determinism," *Social Studies of Science*, vol. 20, no. 2 (May), 333-351.

Bimber, Bruce (1994). "The Three Faces of Technological Determinism," in M. Roe Smith & L. Marx (eds.), *Does technology drive history?* Cambridge, MA: MIT Press, 79–100.

Blackmore, Susan (1999). *The Meme Machine*. Oxford: Oxford University Press.

Blomberg, Olle (2011). "Conceptions of Cognition for Cognitive Engineering," *The International Journal of Aviation Psychology*, 21 (1), 85–104.

Bratman, Michael E. (1987). *Intention, Plans, and Practical Reason*. Cambridge: Harvard University Press.

Bratman, Michael E. (2007). *Structures of Agency: Essays*. Oxford, UK: OUP.

Bratman, Michael E. (2014). *Shared Agency: A Planning Theory of Acting Together*. Oxford, UK: OUP.

Baker, Lynne Rudder (2004). "The Ontology of Artifacts" in *Philosophical Explorations* vol. 7, no. 2, 99-111.

Baker, Lynne Rudder (2008). "The Shrinking Difference Between Artifacts and Natural Objects" in *American Philosophical Association Newsletter*, vol. 7, no. 2 (Spring), 2-4.

Bloom, Paul (1996). "Intention, History, and Artifact Concepts" in *Cognition* vol. 60, 1-29.

Boyd, Robert and Peter J. Richerson (1985). *Culture and the Evolutionary Process*. Chicago, IL: University of Chicago Press.

Boyd, Robert and Peter J. Richerson (2005). *The Origin and Evolution of Culture*. City of New York, NY: Oxford University Press.

Carrara, Massimiliano and Pieter E. Vermaas (2009). "The Fine-Grained Metaphysics of Artifactual and Biological Functional Kinds" in *Synthese* vol. 169, 125-143.

Clark, Andy (2003). *Natural-Born Cyborgs: Minds, Technologies, and the Future of Human Intelligence*. Oxford, UK: OUP.

Clark, Andy (2008). *Supersizing the Mind: Embodiment, Action, and Cognitive Extension*. Oxford, UK: OUP.

Clark, Andy and David Chalmers (1998). "The Extended Mind," *Analysis* 58 (1): 7-19.

Coeckelbergh, Mark (2013). *Human Being @ Risk: Enhancement, Technology, and the Evaluation of Vulnerability Transformations* in Kroes, Peter and Peter-Paul Verbeek (eds.), *Philosophy of Engineering and Technology* volume 12. Heidelberg; New York; London: Springer Dordrecht.

Costall, A., and O. Dreier (2006). *Doing Things with Things: The Design and Use of Everyday Objects*. Aldershot: Ashgate.

Cummins, Robert (1975). "Functional Analysis" in *The Journal of Philosophy* Vol. 72, No. 20 (Nov. 20), 741-765.

Cummins, Robert (1983). *The Nature of Psychological Explanation*. Cambridge, MA: MIT Press.

Cummins, Robert (2002). "Neo-Teleology" in Ariew, André, Robert Cummins and Mark Perlman (eds.), *Functions: New Essays in the Philosophy of Psychology and Biology*. Oxford, UK: OUP, 157-172.

Cummins, Robert (2010). *The World in the Head*. Oxford: OUP.

Davies, Stephen (2012). *The Artful Species: Aesthetics, Art, and Evolution*. Oxford, UK: Oxford University Press.

Dawkins, Richard (1976). *The Selfish Gene*. Oxford, UK: Oxford University Press.

Dennett, Daniel (1987). *The Intentional Stance*. Bradford Books.

Dennett, Daniel (1990). "The Interpretation of Texts, People, and Other Artifacts" in *Philosophy and Phenomenological Research*, Vol. 50, 177-194.

Dennett, Daniel C. (1995). *Darwin's Dangerous Idea: Evolution and the Meanings of Life*. City of New York, NY: Simon & Schuster.

Dennett, Daniel C. (2003). *Freedom Evolves*. City of New York: Penguin.

Devitt, Michael (1984/1991). *Realism & Truth*. Oxford, UK: Blackwell.

Dipert, Randall R. (1993). *Artifacts, Artworks, and Agency*. Philadelphia, PA: Temple University Press.

Dipert, Randall R. (1995). “Some Issues in the Theory of Artifacts: Defining ‘Artifact’ and Related Notions” in *The Monist* vol. 78, no. 2 (April), 119-135.

Dittmar, Anke and Peter Forbrig (2013). Introduction to Special Issue “Cognitive Ergonomics for Designing Collaborative Activities,” *Cognition, Technology, and Work*, vol. 15, 359-362.

Elder, Crawford (2007). “On the Place of Artifacts in Ontology” in Margolis, Eric and Stephen Laurence (eds.) *Creations of the Mind: Theories of Artifacts and Their Representation*, 33–51. Oxford: Oxford University Press.

Ereshefsky, Marc (2004). “Bridging the Gap between Human Kinds and Biological Kinds” in *Philosophy of Science* vol. 71, 912-921.

Franssen, Maarten. (2006). “The Normativity of Artefacts” in *Studies in History and Philosophy of Science* Part A vol. 37, no. 1, 42–57.

Franssen, Maarten, Peter Kroes, T. A. C. Reydon, and Pieter E. Vermaas, eds., (2014). *Artefact Kinds: Ontology and the Human-made World*. Synthese Library, vol. 365. Berlin: Springer.

Godfrey-Smith, Peter (2012). “Darwinism and Cultural Change.” *Phil. Trans. R. Soc. B* 367, 2160–2170.

Grunwald, Armin and Matthias Achternbosch (2013). “Technology Assessment and Approaches to Early Engagement,” in Doorn, Neelke, Daan Schuurbiers, Ibo van de Poel, and Michael E. Gorman (eds.), *Philosophy of Engineering and Technology, Vol. 16: Early Engagement and New Technologies: Opening Up the Laboratory*. Heidelberg; New York; London: Springer Dordrecht, 15-36.

Hancock et al. (2013). “Human-Automation Interaction Research: Past, Present, and Future” in *Ergonomics in Design: The Quarterly of Human Factors Applications* vol. 21, no. 9 (April), 9-14. DOI: 10.1177/1064804613477099. <http://erg.sagepub.com/content/21/2/9>

Hempel, Carl, “The Logic of Functional Analysis” in *Aspects of Scientific Explanation*, City of New York: Free Press, 1965.

Henrich, Joseph and Robert Boyd (2002). “On Modeling Cognition and Culture: Why Cultural Evolution Does Not Require Replication of Representations” in *Journal* *of Cognition and Culture* vol. 2 (2), 87-112.

Hilpinen, Risto (1992). “Artifacts and Works of Art” in *Theoria* vol. 58, 58–82.

Hilpinen, Risto (1993). “Authors and Artifacts” in *Proceedings of the Aristotelian Society* vol. 93, 155–178.

Hilpinen, Risto (1995). “Belief Systems as Artifacts” in *The Monist* vol. 78, 136-155.

Hofstadter, Douglas R. (1995), *Fluid Concepts and Creative Analogies*. City of New York: Basic Books (Harper Collins).

Hofstadter, Douglas R. (2004). “Moore’s Law, Artificial Evolution, and the Fate of Humanity” in Booker, Lashon et al. (eds.), *Perspectives on Adaptation in Natural and Artificial Systems*. City of New York, NY: OUP. Pages 163-196.

Hollnagel, Erik (2003/2008). *Handbook of Cognitive Task Design*. London: Lawrence Erlbaum Associates, Inc.

Houkes, Wybo (2006). “Knowledge of Artefact Functions” in *Studies in History and Philosophy of Science* Part A vol. 37, no. 1, 102-113.

Houkes, Wybo and Pieter Vermaas (2004). “Actions versus Functions: A Plea for an Alternative Metaphysics of Artifacts” in *The Monist*, vol. 87, no. 1 (Jan.), 52-71.

Houkes, Wybo and Pieter E. Vermaas (2009). “Contemporary Engineering and the Metaphysics of Artefacts: Beyond the Artisan Model” in *The Monist* vol. 92(3): 403–419.

Houkes, Wybo and Pieter E. Vermaas (2010). *Technical Functions: On the Use and Design of Artefacts*. Vol. 1 of *Philosophy of Engineering and Technology*. Dordrecht: Springer.

Houkes, Wybo and Pieter E. Vermaas (2013). “Pluralism on Artefact Categories: A Philosophical Defence” in *Rev. Phil. Psych.* vol. 4, no. 3 (Sept.), 543-557.

Houkes,Wybo and Pieter E. Vermaas (2014). “On What Is Made: Instruments, Products and Natural Kinds of Artefacts” in Franssen, Maarten and Peter Kroes (eds.), *Artefact Kinds: Ontology and the Human-made World.*

Houkes, Wybo, Pieter E. Vermaas, K. Dorst and M. J. de Vries (2002). “Design and Use as Plans: An Action-Theoretical Account” in *Design Studies* vol. 23, 303-320.

Hughes, Jesse L. (2009). “An Artifact Is to Use: An Introduction to Instrumental Functions” in *Synthese* Vol. 168, 179-199.

Hughes, Jesse L., Peter Kroes and S. D. Zwart (2007). "A Semantics for Means-end Relations" in *Synthese* vol. 158, 207–231.

Illies, Christian and Anthonie Meijers (2009). "Artefacts Without Agency" in *The Monist* vol. 92, no. 3, 420-440.

Inagaki, Toshiyuki (2008). "Adaptive Automation: Sharing and Trading of Control," Chapter 8 in Hollnagel, Erik (ed.), *Handbook of Cognitive Task Design*. London: Lawrence Erlbaum Associates, Inc., 147-170.

Jablonka, Eva and M. Lamb (1995). *Epigenetic Inheritance and Evolution: The Lamarckian Dimension*. Oxford, UK: Oxford University Press.

Jablonka, Eva (2000). "Lamarckian Inheritance Systems in Biology: A Source of Metaphors and Models in Technological Evolution" in John Ziman (ed.), *Technological Innovation as an Evolutionary Process*. Cambridge, UK: Cambridge University Press, 27-40.

Jablonka, Eva and Marion J. Lamb (2005). *Evolution in Four Dimensions: Genetic, Epigenetic, Behavioral, and Symbolic Variation in the History of Life*. Cambridge, MA and London, UK: The MIT Press (A Bradford Book).

Kaptelinin, Victor and Bonnie A. Nardi (2006). *Acting with Technology: Activity Theory and Interaction Design*. Cambridge, MA: MIT Press.

Kirchhoff, Michael (2014). "Extended cognition & constitution: Re-evaluating the constitutive claim of extended cognition" in *Philosophical Psychology* vol. 27, no. 2, pp. 258-283.

Kornblith, Hilary (1980). "Referring to Artifacts" in *The Philosophical Review* vol. 89, no. 1 (Jan.), 109-114.

Kornblith, Hilary (2007). "How to Refer to Artifacts" in Margolis, Eric and Stephen Laurence (eds.), *Creations of the Mind: Theories of Artifacts and Their Representation*. Oxford, UK: OUP, 138-149.

Kroes, Peter (2010). "Engineering and the Dual Nature of Technical Artefacts," *Cambridge Journal of Economics*, vol. 34, no. 1, 51–62.

Kroes, Peter (2012). *Technical Artefacts: Creations of Mind and Matter, A Philosophy of Engineering Design*. Dordrecht: Springer.

Kroes, Peter and Anthonie Meijers, eds., (2006). "The dual nature of technical artifacts" (introduction) in *Studies in History and Philosophy of Science* Part A vol. 37, no. 1, 1-4.

Kroes, Peter and Peter-Paul Verbeek (2013). "Introduction: The Moral Status of Technical Artefacts," in Kroes, Peter and Peter-Paul Verbeek (eds.), *Philosophy of Engineering and Technology, Vol. 17: The Moral Status of Technical Artefacts*. Heidelberg; New York; London: Springer Dordrecht, 1-10.

Kroes, Peter, Maarten Franssen and L. L. Bucciarelli (2009). "Rationality in Engineering Design" in Meijers, Anthonie (ed.), *Philosophy of Technology and the Engineering Sciences.* Elsevier.

Krohs, Ulrich (2009). "Functions as Based on a Concept of General Design," *Synthese* vol. 166, 69-89.

Krohs, Ulrich and Peter Kroes (2009). "Philosophical Perspectives on Organismic and Artifactual Functions," Ulrich Krohs and Peter Kroes (eds.), *Functions in Biological and Artificial Worlds: Comparative Philosophical Perspectives*. MIT Press.

Lehto, Mark R., Fiona Fui-Hoon Nah, and Ji Soo Yi (2012). "Decision-Making Models, Decision Support, and Problem Solving" in Salvendy, Gavriel (ed.), *Handbook of Human Factors and Ergonomics* 4th Edition. Hoboken, NJ: Wiley, pp. 192-242.

Lewens, Tim (2002). "Darwinnovation!" [Review of Ziman, 2000.] *Stud. Hist. Phil. Sci.* 33: 199–207.

Lewens, Tim (2004). *Organisms and Artifacts: Design in Nature and Elsewhere*. Cambridge, MA: MIT Press.

Lewens, Tim (2009). "Seven Types of Adaptationism," *Biol. Philos.* 24: 161–182.

Lewens, Tim (2012). "Cultural Evolution," in Kincaid, Harold (ed.), *The Oxford Handbook of Philosophy of Social Science*. Oxford, UK: OUP.

Lewens, Tim (2014). [Review of Kim Sterelny (2012), *The Evolved Apprentice*.] *Brit. J. Phil. Sci.* 65, 185–189.

MacIntyre, Hector (2013). "Category Cognition and Dennett's Design Stance" in *Review of Philosophy and Psychology* vol. 4, no. 3 (Sept.), 483-495.

Marconi, Diego (2013). "Pencils Have a Point: Against General Externalism About Artifactual Words" in *Rev. Phil. Psych.* vol. 4, no. 3 (Sept.), 497-513.

Margolis, Eric and Stephen Laurence, eds., (2007). *Creations of the Mind: Theories of Artifacts and Their Representation*. Oxford: Oxford University Press.

Marzke, Mary W. (2013). "Tool making, hand morphology and fossil hominids," *Phil. Trans. R. Soc. B* 368: 1-8.

Meijers, Anthonie (2001). "The Relational Ontology of Technical Artifacts" in Peter Kroes and Anthonie Meijers (eds.), *The Empirical Turn in the Philosophy of Technology*. Amsterdam: Elsevier.

Mercader, Julio, et al. (2002). "Excavation of Chimpanzee Stone Tool Site in the African Rainforest" in *Science* vol. 296, 1452–1455.

Merlin, Peter W., Gregg A. Bendrick and Dwight A. Holland (2011). *Breaking the mishap chain: human factors lessons learned from aerospace accidents and incidents in research, flight test, and development*. Washington, DC: NASA Aeronautics Book Series.

Millikan, Ruth Garrett (1984). *Language, Thought, and Other Biological Categories: New Foundations for Realism*. Cambridge, MA: MIT Press.

Millikan, Ruth Garrett (1989a). "In Defense of Proper Functions" in *Philosophy of Science*, vol. 56, no. 2 (June), 288-302.

Millikan, Ruth Garrett (1989b). "An Ambiguity in the Notion of Function" in *Biology and Philosophy* Vol. 4, 172-176.

Millikan, Ruth Garrett (1999). "Wings, Spoons, Pills, and Quills: A Pluralist Theory of Function" in *The Journal of Philosophy*, Vol. 96, No. 4 (April), 191-206.

Millikan, Ruth Garrett (2002). "Biofunctions: Two Paradigms" in Ariew, André, Robert Cummins and Mark Perlman (eds.), *Functions: New Essays in the Philosophy of Psychology and Biology*. Oxford, UK: OUP, 113-143

Mitcham, Carl (2009). "A Philosophical Inadequacy of Engineering," *The Monist*, vol. 92, no. 3, 339-356.

Mitcham, Carl (2013). "Agency in Humans and in Artifacts: A Contested Discourse," in Kroes, Peter and Peter-Paul Verbeek (eds.), *Philosophy of Engineering and Technology, Vol. 17: The Moral Status of Technical Artefacts*. Heidelberg; New York; London: Springer Dordrecht, 11-30.

Mokyr, Joel (2000). "Evolutionary Phenomena in Technological Change" in John Ziman (ed.), *Technological Innovation as an Evolutionary Process*. Cambridge, UK: Cambridge University Press, 52-65.

Morton, Adam (2006). "Finding the Corkscrew" in *Studies in History and Philosophy of Science* Part A vol. 37, no. 1, 114-117.

Morton, Adam (2012). *Bounded Thinking: Intellectual Virtues for Limited Agents*. Oxford, UK: OUP.

Nelson, J. A. (1982). “Schwartz on Reference” in *Southern Journal of Philosophy* vol. 20, 359–365.

Perlman, Mark (2004). “The Modern Philosophical Resurrection of Teleology” in *The Monist*, vol. 87, no. 1 (Jan.), 3-51.

Pollock, J. (1987). “Epistemic Norms” in *Synthese* vol. 71, no. 1, 61–95.

Pollock, J. (1988). “My Brother, the Machine” in *Noûs* vol. 22, no. 2, 173–211.

Preston, Beth (1998). “Why Is a Wing Like a Spoon? A Pluralist Theory of Function” in *The Journal of Philosophy*, Vol. 95, No. 5 (May), page 220 (215-254).

Preston, Beth (2013). *A Philosophy of Material Culture: Action, Function, and Mind*. City of New York, NY: Routledge, page 30.

Putnam, Hilary (1975). “The Meaning of ‘Meaning’” in Hilary Putnam, *Philosophical Papers* vol. 2. Cambridge: Cambridge University Press, 215–271.

Putman, Daniel (1982). “Natural Kinds and Human Artifacts” in *Mind* vol. 91, 418-419.

Richerson, Peter J. and Robert Boyd (2005). *Not by Genes Alone: How Culture Transformed Human Evolution*. Chicago: University of Chicago Press.

Roberts, Tom (2012). “You Do the Maths: Rules, Extension, and Cognitive Responsibility” in *Philosophical Explorations* vol. 15, no. 2 (June), 133-145.

Rovira, Ericka and Raja Parasuraman (2010). “Transitioning to Future Air Traffic Management: Effects of Imperfect Automation on Controller Attention and Performance” in *Human Factors* vol. 52, no. 3 (June), 411–425.

Schwartz, Stephen P. (1978). “Putnam on Artifacts” in *The Philosophical Review* Vol. 87, No. 4 (Oct.), 566-574.

Schwartz, Stephen P. (1980). “Natural Kinds and Nominal Kinds” in *Mind* vol. 89, 182-195.

Schwartz, Stephen P. (1983). “Reply to Kornblith and Nelson” in *Southern Journal of Philosophy* vol. 21, 475–479.

Searle, John R. (1995). *The Construction of Social Reality*. New York, NY: The Free Press.

Searle, John R. (2010). *Making the Social World: The Structure of Human Civilization*. Oxford, UK: OUP.

Serrano, J. Ignacio, M. Dolores del Castillo, and Manuel Carretero (2014). “Cognitive? Science?” *Found. Sci.* 19: 115–131.

Shannon, Claude Elwood (1937). *A Symbolic Analysis of Relay and Switching Circuits*. Cambridge, MA: MSc. thesis available from MIT Library.

Soavi, Marzia (2009). “Realism and Artifact Kinds” in Ulrich Krohs and Peter Kroes (eds.), *Functions in Biological and Artificial Worlds: Comparative Philosophical Perspectives*. Cambridge, MA: MIT Press, 185-202.

Sperber, Dan (1996). *Explaining Culture: A Naturalistic Approach*. Oxford, UK: Blackwell.

Sperber, Dan (2007). “Seedless Grapes: Nature and Culture” in Margolis, Eric and St phen Laurence (eds.), *Creations of the Mind: Theories of Artifacts and Their Representation*. Oxford, UK: OUP, 124-137.

St Clair, James J. H. and Christian Rutz (2013). “New Caledonian crows attend to multiple functional properties of complex tools,” *Phil. Trans. R. Soc. B* 368: 1-9.

Sterelny, Kim (2006a). “The Evolution and Evolvability of Culture” in *Mind & Language* vol. 21, no. 2 (April), 137-165.

Sterelny, Kim (2006b). “Memes Revisited” in *Brit. J. Phil. Sci.* vol. 57, 145-165.

Sterelny, Kim (2012). The Evolved Apprentice: How Evolution Made Humans Unique. Cambridge, MA: MIT Press (A Bradford Book).

Thomasson, Amie L. (2003). “Realism and Human Kinds” in *Philosophy and Phenomenological Research* Vol. LXVII, No. 3 (Nov.), 580-609.

Thomasson, Amie L. (2007). “Artifacts and Human Concepts” in Eric Margolis and Stephen Laurence (eds.), *Creations of the Mind: Theories of Artifacts and Their Representation*. Oxford and New York: Oxford University Press, 52–73.

Vaesen, Krist (2006). “How Norms in Technology Ought to Be Interpreted” in *Techné* vol. 10, no. 1, 117-133.

Vaesen, Krist (2011). “The Functional Bias in the Dual Nature of Technical Artefacts Program” in *Studies in History and Philosophy of Science*, vol. 42, 190-197.

Vaesen, Krist and Melissa van Amerongen (2008). “Optimality vs. Intent: Limitations of Dennett’s Artifact Hermeneutics” in *Philosophical Psychology*, vol. 21, no. 6 (Dec.), 779-797.

Vanderhaegen, F. (2012). “Cooperation and Learning to Increase the Autonomy of ADAS” in *Cognition, Technology, and Work* vol. 14, 61-69.

Verbeek, Peter-Paul (2011). *Moralizing Technology: Understanding and Designing the Morality of Things*. Chicago, IL: University of Chicago Press.

Vermaas, Pieter E. (2009). “Artefacts in Analytic Metaphysics” in *Techné* vol. 13, no. 2, 74-81.

Vermaas, Pieter E. and Wybo Houkes (2003). “Ascribing Functions to Technical Artefacts: A Challenge to Etiological Accounts of Functions” in *Brit. J. Phil. Sci.* 54, 261-289.

Vermaas, Pieter E. and Wybo Houkes (2006a). “Use Plans and Artefact Functions: An Intentionalist Approach to Artefacts and Their Use” in Costall, A. and O. Dreier (eds.), *Doing Things with Things: The Design and Use of Everyday Objects*. Aldershot: Ashgate, 29-48.

Vermaas, Pieter E. and Wybo Houkes (2006b). “Technical Functions: A Drawbridge between the Intentional and Structural Natures of Technical Artefacts” in *Studies In History and Philosophy of Science* Part A, vol. 37, no. 1, 5–18.

Wilson, Robert A. (2004). *Boundaries of the Mind: The Individual in the Fragile Sciences: Cognition*. Cambridge University Press.

Wilson, Robert A. (2014). “Ten Questions Concerning Extended Cognition” in *Philosophical Psychology* vol. 27, no. 1, 19-33.

Winner, Langdon (1977). *Autonomous Technology: Technics-out-of-Control as a Theme in Political Thought*. Cambridge, MA: MIT Press, page 15.

Winner, Langdon (1986). *The Whale and the Reactor: A Search for Limits in an Age of High Technology*. Chicago, IL: University of Chicago Press.

Woods, David D. and Erik Hollnagel (2006). *Joint Cognitive Systems: Patterns in Cognitive Systems Engineering*. London, UK: Taylor & Francis.