The Built-in Bias of Technology

Steven H. VanderLeest
Department of Engineering, Calvin College, Grand Rapids, MI

1. Abstract

It is widely, though not universally, accepted in the literature that technology is non-neutral, i.e., it is partial to certain uses. However, this understanding is not widespread amongst engineering students, and the perception of neutrality can have perilous societal consequences. Some preliminary work has identified pedagogical approaches to instilling better understanding of non-neutrality in the classroom. This paper continues that line of thought. Starting with Kranzberg’s assertion that “technology is neither good nor bad; nor is it neutral” I explore the cultural appropriateness of technology as a sign of inherent bias. This leads to a brief examination of mental models of technology as an approach to understanding how the unintended consequences of a technology may not be as haphazard as first appears. With these concepts as background, I analyze the means by which bias is built into each stage of the design process. Carl Mitcham has suggested that a fruitful area of further investigation would be to examine how this bias can be identified in the structures of the technological products themselves. I conclude with some initial ideas on approaches to such an analysis, e.g., using reverse engineering analysis to translate form back to intended function.

2. Introduction

Technology is not neutral. It has an intrinsic bias that is built into it from the original inception of a particular problem, throughout the entire design process, all the way to the implementation, use, and disposal of a product. Technology is obviously biased towards at least one use – the use intended by the designer. It is designed to perform its intended function. However, other biases may be at work. Even though the designer did not consciously build in other functions, the technological product often can be used in ways the designer did not intend. Consider the lawn mower that is used to trim hedges, the car that is used to pull skateboarders, the shoe that is used to hammer in a nail, or the aircraft that is used as a weapon. Humans are ingenious tool makers but also imaginative tool users (sometimes devilishly so). Even when the maker did not envision a use, the user might boldly conceive of a new use. The possible uses of a tool are not limitless however. The structure of the tool lends itself to some uses better than others. Because the hammer pounds nails better than the shoe, we will use the hammer, given the choice. The set of likely uses, or even the wider set of viable uses provides a map of the bias that is built into the technological product.

The structure of the technology includes physical aspects (such as the shape of the hammer handle) but may also include virtual aspects (such as the layout of a computer graphical interface as a metaphor of a desktop). A technological product, especially a tool, is often considered well designed when form follows function, when the structure intuitively suggests the use. Any
particular structure can lead to multiple possible functions, and a product made of multiple structures will have a multitude of possible functions – some intended by the designer, some not. Werner Rammer recognized one facet of this bias in that “social concepts and practices are consciously and non-consciously incorporated in the machine and inscribed in the programs.”

Several philosophers of technology have recognized that technology extends human capabilities in some respects, while limiting them in others. This extension and limitation is an outgrowth of the bias built into the technology. Don Ihde called this effect amplification and reduction (such as a microscope that magnifies one area while simultaneously reducing your field of vision). Marshall McLuhan, thinking especially about communication technology (the media), called it extensions and amputations. Egbert Schuurman considered it a reduction: “When the computer dominates education, student users may adapt themselves uncritically to the computer’s reduced worldview. Whatever does not make it to the computer screen is then simply not considered by the student. Computer programs contain only information that has to do with measurements, sizes, and statistics” Similarly, Pacey sees the technology itself as a distraction that limits our view: “One problem seems to be that the use of such equipment [visual processing, computers, etc.] is sometimes associated with a narrowing of the focus of what one looks at. The equipment itself is part of the fascination, so one tends to look only at what it reveals.”

While a few philosophers deny the bias of technology completely, buying into “guns don’t kill people – people kill people”, Rammert (following some of John Dewey’s thought) is a bit more subtle: “Technology has no existence and function outside of its use. It is what I would like to call the use-relations that create both the handled object as a tool and the manipulating gesture as technical practice.” But this does not sufficiently recognize the role of the technology in the use. A technology biases the user towards those actions for which it is particularly well-suited, and dissuades the user from actions for which it is poorly suited. For example, guns and knives don’t kill people by themselves, but people without guns and knives are less likely to kill than those with them.

3. Teaching about Bias in Technology

It is widely (though not universally) accepted in the literature that technology is non-neutral, i.e., it is partial, or biased, to certain uses. Many scholars have argued effectively against the neutrality of technology. Allchin claims “it is disingenuous nowadays to claim [that technology is neutral]...the very artifacts bias who can and cannot use them.” Ibo van de Poel enumerates three reasons that the instrumental view (technology as a “means” alone, and thus neutral) is unsatisfactory: 1) formulation of goals and choice of technological means are not independent, 2) choice of means for a given end is not neutral, 3) technologies realize more than their intended goal. Borgmann also considers the “means” viewpoint and claims “...it is an equivocation to speak indifferently of tools in a modern and in a pretechnological setting. A means in a traditional culture is never mere but always and inextricably woven into a context of ends.” Don Ihde makes the case that “this non-neutral, transformative power of humans enhanced by technologies is an essential feature of the human-technology relations.”

Unfortunately, this understanding of the non-neutrality of technology is not widespread amongst engineering students. Students are often inclined to believe the neutrality premise and the perception of neutrality can have perilous societal consequences. If the product is neutral, then
logically, the designer has no ethical responsibility for how the product is used or abused. Under this view, the engineers, manufacturers, and distributors need not worry about the social impact of their products. If the product is neutral, then no user can rightly complain about an unsafe product, and no seller would be liable for any damage or harm caused by use of their product. Thus it is essential that students understand the non-neutrality concept and its implications. Some preliminary work has identified pedagogical approaches to instilling better understanding of non-neutrality in the classroom, including use of descriptive yet simpler terminology (such as “bias” rather than “non-neutral”), relating bias to the responsibility of manufacturers to produce safe products, use of real world problems with no single right answer, and examination of technology as an embodiment of human will.

In a typical class one finds that some students quickly buy into the concept of bias, while others cling tenaciously to non-neutrality. The latter claim that obviously technology cannot be held morally responsible for the action a human user took with it. But that is not quite the claim I am making. I have found it useful to let this argument develop during class discussion so that I can point out that something can be biased without having agency. That is, technology can bias the user (who does have agency) towards certain actions and away from other actions. Yes, the gun does not by itself kill people, but a person with a gun is more likely to kill than without it. A useful demonstration I have tried is to bring in a wooden board with two nails partially nailed into it, along with a sack containing two tools. I invite two students to come to the front to complete the job of pounding the nails in. The first student is one who claims technology can be biased, and so is allowed to choose which of the two tools in the sack they will use for the job. The second student is one who claims technology is neutral, so they must use whichever tool is left in the sack. When they look inside the sack they find a hammer and a screwdriver. As the hammer-wielding student quickly finishes one nail while the screwdriver-wielder makes little headway, it becomes obvious that the tools are not neutral with respect to particular uses. I then make the further point that to the person holding a hammer, everything starts to look like a nail, i.e., because a technology is biased towards certain uses, we tend toward uses aligned with the biases of the technology. Discussions like this set the stage for presentation of more structured methods, such as design norms, that make identification and analysis of technology bias an intrinsic part of the design process.

4. Bias is in the eye of the beholder

Mel Kranzberg has asserted that “technology is neither good nor bad; nor is it neutral” His point is that the same technology can be good for some users in some contexts and bad for other users in the same or other contexts. He uses the example of a chemical pesticide that is banned in one context because it is found to be a carcinogen, but nevertheless is used widely in another context because it controls malaria-carrying mosquitoes. Kranzberg’s claim, in essence, is that a particular technology cannot be judged good or bad in a universal or global sense. The value judgments we make about a technological product are context dependent: “...technical developments frequently have environmental, social, and human consequences that go far beyond the immediate purposes of the technical devices and practices themselves, and the same technology can have quite different results when introduced into different contexts or under different circumstances.” Like Kranzberg, I claim that technology is not neutral, but rather than disavowing any overall judgment regarding its goodness, I observe that technological products have a variety of good and bad effects on all stakeholders. These effects are related not
Design requires making choices between alternative solutions. This selection is made on the basis of prioritized design criteria (some explicit, some perhaps implicit). That prioritization represents the bias of the designer as influenced by the customer, the employer, internal and governmental standards, the local design environment, and a variety of social forces. Recognizing these influences is not easy – akin to asking a fish to recognize the influence of the water. Just as the fish only recognizes the environment of the water once removed from that environment, our students need similar “cultural leverage.” A liberal education provides such leverage by figuratively placing the student in worlds long ago or far away and exploring the differences to our present social context. Recognizing differences in social context is a first step in identifying how one’s own social context influences one’s decisions.

Design norms are moral guidelines for how designs ought to be, one of them being “cultural appropriateness.” Determining whether a design is culturally appropriate requires making a value judgment on whether a design fits the context in which it is placed, i.e., evaluating the good and bad effects on the stakeholders. This is obviously a rather difficult task, requiring the designer to envision how the technology will be used (and possibly abused) as influenced by the whole milieu of cultural and social forces. Understanding these forces on the macro (e.g., economic, political) and micro (e.g., psychological) scale requires a strong liberal education.

If all technological products are in some way a form of communication between the designer, manufacturer, and user, if the artifacts embody the thinking, imply the will, or capture some of the soul of the creator, and further, if the user interprets that technology according to their own social context, then perhaps in this postmodern age we dare not hope for a technology that is appropriate. Indeed, Werner Rammert expounds “…it is rather difficult if not impossible either to reduce an artifact to one general function or to interpret an artifact’s particular meaning. Should we look for the inventor’s vision or should we review the engineering and marketing plans of the producer or should we observe and ask the users of technology?” Can the postmodernist deconstruct the meaning of technology as easily as deconstructing a text? Does the designer have any hope for communicating the intent of the technology to the user? Certainly! There is not only hope, but expectation of understanding. Most designs are used in the way they were intended by most users. In fact, when a product is not quite right – it does not perform the intended function as well as the user desires – the user often intuitively grasps the intended function. Indeed, most of us have criticized technology at one time or another as missing the mark, implying that we intuitively understood the mark that was missed. Rammert seems to despair that because there are multiple interpretations of functionality, one cannot hope to find consensus on a single intended function. But consensus is not necessary. A technological product can have many functions, some of which were intended by the designer, but possibly others that were invented by the user. We can find value (good or bad) in each of these uses, as a function of both the user (in context) as well as the technology (which is biased towards some uses more than others).

Engineering easily and practically answers a deconstructionist view of technology because our products generally perform as advertised. The proof is in the pudding. At least some intent must be effectively communicated from user to designer in specification and from designer to user in
design and production, because users actually use the products for the envisioned purpose. But engineers can also learn from the deconstructionist approach, in that the varied contexts and varied users within which a technology is placed produce a complex web of value relationships, which may sometimes be conflicting. Just as trade-offs are required amongst technical design criteria, so too are trade-offs required in balancing soft design criteria (i.e., design norms such as cultural appropriateness or stewardship). Just as design decisions must be reached despite incomplete and uncertain technical knowledge, so too must they be reached despite incomplete and uncertain understanding of the value and impact of the technology in multiple contexts. Society also examines these trade-offs, e.g., the risk/benefit analysis of a technology in a product liability legal case. In these cases, the jury judges whether the designer, manufacturer and/or distributor of a product could reasonably foresee the negative consequences of their technology.

If technology indeed is biased towards certain uses, and users may sometimes use the product in ways the designer did not intend (ways that may be dangerous), then should we actively prevent users from using a technology in a way the designer did not intend? One can imagine just such a technical solution, such as a handgun that checks for an authorized fingerprint in order to disengage the safety, or more simply, a medicine bottle with a child-proof cap. Nevertheless, it is impossible to stop all abuses. Furthermore, most problems cannot be solved with purely technical solutions, but require a broader approach. For example, safe operation of a vehicle requires not only technical solutions such as traffic lights, but social solutions such as driver’s training, traffic laws along with enforcement of them, and so forth.

5. Mental models and Technology Bias

One way to understand how well communication of intent and functionality occurs between the designer and user of technology is the concept of mental models. Mental models are representations (usually visual in nature) that people use internally to understand technology – constructed and adjusted as the user gains experience with the technology. The models are both descriptive and predictive in nature. For example, the designer can better understand user requirements by discovering the user’s mental model of the product\textsuperscript{16}. During development the design team may have a team mental model of the technology or some members may share various mental models of certain aspects of the design\textsuperscript{17}. Mental models as representations of technology can help us perceive both the designer’s and the user’s understanding of the function of a technology\textsuperscript{18, 19}.

The concept of mental models helps us see why sometimes the intended functionality of a technology does not come to full fruition and sometimes unintended consequences arise. In fact, mental models can help us see why the unintended consequences are unforeseen – because the mental model has some “blind spots” and thus the unintended consequences are not as haphazard as they may first appear. During the design process, a proposed alternative solution is by nature incomplete. When evaluating this alternative for suitability, different stakeholders may have differing mental models of the alternative. Thatcher and Greyling identify several ways in which mental models are constructed (and adjusted): “Individuals form their mental models of systems through specific training instruction, through interaction with a system and through watching others interact with a system.”\textsuperscript{20} Since the opportunity for instruction regarding a proposed solution or interaction with it are quite limited during the design stage (afforded primarily through simulation models, analytical models, or prototypes), the designers and anyone else with
input to the design decisions will have correspondingly limited mental models. Mental models of existing technologies that are incomplete lead to mistakes: “mental models of experienced users are richer, more abstract and more complete than less experienced users….Experienced users were likely to commit fewer errors, the mistakes were less serious, they engaged in a wider range of problem-solving strategies, they were less likely to suffer from ‘cognitive lock-up’, and were more efficient users.” A mental model dictates what functions we can envision for a particular technology, perhaps leading a user to new uses that the designers did not predict with their own mental model or perhaps even disguising potential dangers of a particular use.

By the very nature of the design process, designers are inexperienced with proposed alternative solutions (since they have not been actually implemented yet) and thus will have incomplete mental models that may lead to misunderstandings of the alternative and perhaps imbalanced decisions when choosing between the alternatives. For example, mental models are influenced by past experience – designers often assume analogy to some known entity, understanding the new in terms of the old. This lends inertia to current understandings and usage, and may undervalue the novel approach which is less understood but may actually be a better choice.

One partial remedy for poor mental models during the design process is the use of diverse design teams that discuss models early in the process so that better models can be formed by the team. Divergent thinking can help fill in the gaps and blind spots of the team’s mental model of the technology. Pacey gets at this idea of adjusting the model through interaction: “…although ideas may arise in all sorts of ways that may be described as intuitive or participatory, there is always an obligation to translate them into more rigorous, often mathematical formulations, so that others may understand and check them, and explore their precise implications.”

6. Bias in the Design Process

Let us now explore the means by which bias may be introduced during each stage of the design process. Engineering design projects typically begin with a problem specification phase. The constraints, requirements, and specifications of the design are elicited from the customer, from internal standards, external standards, and so forth. Bias in the technological product can start already at this stage. Sometimes the problem is inappropriately stated in terms that tend to emphasize one particular solution. Designers are more likely to catch requirements they understand and more likely to miss or misunderstand requirements that are in unfamiliar areas. This kind of inertia may give unfair advantage to existing approaches and discourage use of more innovative approaches that may actually solve the problem in a better way. The designer may not even be solving the right problem. For example, the designed solution may provide short-term relief for a problem but fail to address the deeper issues that would lead to a more permanent solution.

Once the problem has been specified, the engineering team identifies potential solutions by brainstorming about possible alternative approaches to solving the problem. However, the designer’s experience with proven alternatives may give these potential solutions more weight than they deserve. “Heuristics, precisely because they are empirical regularities that are useful rather than water-tight, are always associated with biases…. There is a bias towards selecting candidates from the solution space close to existing solutions.” In fact, design reuse is a common principle for reducing design costs. However, reuse without careful, critical evaluation
is precarious, in that designers may miss better solutions simply because they are innovative. This may be a good argument for not only diverse design teams, but also for broad technical education that provides sufficiently broad experience to help the designer identify multiple ways to solve a problem.

In order to choose between alternative solutions that were identified in the previous phase of the design, typically a set of design criteria are identified (such as “low cost”, “low weight”, “durability”, and so forth). Each alternative is then evaluated against these criteria in order to determine which solution offers the best trade-off amongst the criteria. Normally these criteria are evaluated from the point of view of the customer and the designer. However, other stakeholders must be recognized. Stakeholders are all those that will be affected by the product, e.g., people living near a proposed power plant location are not paying for the design of the plant, but will certainly be affected by it and are thus stakeholders in the design. How are the objectives of non-paying stakeholders recognized? In some ways they become aggregated into the objectives through government regulations, the desire of the customer to avoid legal liability issues, or the customer’s desire to build goodwill. However, the engineer must take care to explicitly recognize the stakeholder’s issues where critical health and safety concerns are involved. It is not enough to say “I was just doing my job.” – society does not recognize this as a sufficient defense. Even where legal responsibility may end, moral and ethical responsibilities may require the engineer to look at the objectives of other stakeholders. A further danger in omitting objectives is noted by Talbot: “every invention, from television to nuclear power, tends to incarnate the will (conscious or unconscious) of its employer. And if that will is less than fully conscious, the invention yields us more than we wield it.”

Design criteria often push a design in conflicting directions. In order to evaluate a design against multiple criteria, the criteria are prioritized or weighted. The prioritization process forces the engineer to explicitly identify how various criteria will be valued. One simple test to help recognize when objectives are missing is to include an “unreasonable” alternative solution and then objectively evaluate that alternative solution to be sure it does not win the decision. If it does, then some important objective has probably been omitted. For example, in solving a personal transportation problem, one might choose an army tank over an automobile if safety is the only decision criterion, but such a result points out that some important objectives were left out, such as high gas mileage or ease of parking.

After alternative solutions have been brainstormed and some preliminary analysis performed on each, and after design criteria have been identified and prioritized, a design decision must be made. Each alternative solution is evaluated for how well it meets each design criteria. This is another judgment call on the part of the engineer. One technique, the design matrix, is to list alternative solutions in the rows of a table and list design criteria as columns. Each alternative is numerically scored against each criteria, scaled by the criterion weights. The scores are then totaled to see which alternative best meets the criteria. The difficulty with non-technical criteria is that they are not easily quantified (and such quantification may require non-linear scales). Ranking design alternatives against multiple weighted criteria is an important area of design research and beyond the scope of this paper, except for two relevant points. First, some have argued that Arrow’s Impossibility Theorem indicates that there can be no consistent method for making a choice amongst alternatives with multiple criteria. Scott and Antonsson make the case that Arrow’s theorem does not extend from the arena of social choice into the arena of the
engineering design decision making. A second result of interest, from Busby and Lloyd, finds that for engineers making design decisions, “…the perceived merits of a design, in the real world, can be unrelated to the content of the design. Instead, they can arise from association with other events.” That is, non-technical biases can influence technical decisions.

Even the implementation stage of design can introduce bias. Manufacture and assembly with the goal of mass production may subtly alter product. For example, over time a manufacturer may attempt to minimize material costs on a product and reduce certain dimensions or perhaps replace certain materials in a product. This may lead to product failures or it may even lead to different usage of the product (perhaps an item was originally too heavy to be carried in a purse or backpack but over time, the weight is reduced enough that new usages become possible, usages that were not envisioned by the designer).

7. Identifying Bias through Structure

Carl Mitcham has suggested that a fruitful area of further investigation when discussing the non-neutrality of technology would be to examine how bias might be identified in the structures of the technological products themselves. While there is substantial literature on the social impact of a variety of technological products, there has been scant work attempting to tie bias directly to the actual structure of the technology. I suggest two possible approaches here.

One method of getting at the connection of structure to bias would be a comparison study: find two alternative technologies that purport to solve the same problem and compare the side-effects of each, with the hopes that the structures unique to each solution can be related to the side-effects unique to each. This approach is challenging, since differing solutions to a problem are often adopted within different cultural contexts (making it difficult to distinguish between contextual differences and technological differences) and some biases are not in the physical structures of the technology so much as in the surrounding social structures, such as advertising, training, and other cultural conditioning. So for example, can we point to fixed classroom seating (compared to flexible seating) as a cause of lecture styles of teaching (compared to more interactive styles), or is the lecture actually the cause of fixed seating? Can we point to the microwave oven as the cause of families spending less time together, or is the more hurried pace of the modern family producing the driving need for the microwave?

A second possible approach for getting at the connection between bias and structure is reverse engineering, which attempts to trace form back to function by analyzing an existing product as a black box to see how it transforms flows (energy, material, signals). Analysis of this transformation results in a model of the function of the device. Reverse engineering might include interviews with users to help refine our model of the function of the device. However a single technology can have many possible functions. For example, most users may use an oven to heat food, but a few might (unwisely) use it to heat the room. Thus we cannot assume a single function, but must look at how closely a product fits other functions. When considering the problem of inserting a nail into the wall to hang a picture, we might identify several objects at hand that have sufficient mass, a hard surface, and are easily grasped so that we can swing them and pound the nail with some momentum. While a hammer might be the ideal choice, a rock or a rifle butt might be selected in its absence. Humans are rather good at identifying makeshift tools to perform tasks (perhaps we all are engineers at heart, i.e., we are homo faber), but
unfortunately, we are not so good at identifying consequences beyond the function we desire to fulfill. The rifle butt may well pound the nail, but it may also discharge – causing holes where we do not desire them!

As a small case study, consider a structure common in many technological products – the push or toggle button used for the human-machine interface. The button has only two states (on or off) so that it is easily connected to an internal microprocessor. The button is also easily activated by a human finger. However, a button that is electrically connected so that a small physical push with the finger produces some action elsewhere in the machine (and often disproportionately more powerful) may not be the best interface. In fact, it is biased. The ease of initiating an action which may be irrevocable – such as the trigger on a hand gun or clicking the “send” button in an email program – can lead to poorly thought out yet morally significant decisions. Can we reverse engineer such a simple aspect of technology? We can infer some of the design specifications. For instance, an input force leads to a binary electrical output. This explains part of the problem: the output is not a linearly proportional function of the input. If we think of technology as an extension or amplification of human abilities, then compare a tool such as a carpenter’s hammer to the button on a nail gun. The simple hammer allows some proportional feedback so that the user literally feels how well a nail is struck or whether the blow glanced off. When the button on a nail gun is depressed, there is some tactile feedback in the form of vibration that the action has been taken, but the user cannot tell by feeling alone whether the nail was embedded properly, at the correct angle and to the proper depth. Similarly, the decision to strike a blow with the hammer gives the user some pause, because one feels the weight of the hammer and has direct tactile feedback from the results. In the case of the nail gun, it is very easy to push the button without thinking it through. Of course this generalization is not universally true, since one can still mistakenly hammer one’s thumb, and one probably takes more care with a nail gun because it is more powerful. In addition, a nail gun will typically have a safety mechanism to prevent inadvertent use. (But of course habitual use may lead one to inadvertently circumvent the safety.) Albert Borgmann has noted this effect that remote activation of technology has on human behavior, as summarized by Crouch: “…the tools of ‘careful power’ – the practices that enabled us to come alongside creation’s own inherent power… have been replaced by the tools of ‘regardless power,’ the buttons, knobs, and switches that require no skill and give us the impression that the world simply awaits our command.”

Understanding how biases can be built into technology and recognizing some of these biases in existing products is important not only for engineers as the designers of those technologies, but also for users of those technologies. It is important that engineering educators convey these ideas to their students, with the result, hopefully, that future designs better address the full societal context into which they are introduced and that the designers better understand their ethical responsibilities with respect to their designs.


5. Rammert, p. 31.


15. Rammert, p. 31.


20. Thatcher and Greyling, p. 300.


Proceedings of the 2004 American Society for Engineering Education Annual Conference & Exposition
Copyright © 2004, American Society for Engineering Education
26 Busby and Lloyd, p. 164


STEVEN H. VANDERLEEST is a Professor of Engineering at Calvin College. He has an M.S.E.E. from Michigan Tech. U. (1992) and Ph.D. from the University of Illinois at Urbana-Champaign (1995). He was recently program director for a FIPSE grant, “Building Information Technology Fluency into a Liberal Arts Core Curriculum.” His research interests include responsible technology and reliable systems via multi-version programming.