Notes on Design and Science in the HCI Community

Christoph Bartneck

The human-computer interaction (HCI) community is diverse. Academics and practitioners from science, engineering, and design contribute to its lively development, but communication and cooperation between the different groups is often challenging. Designers struggle to apply the results of scientific studies to their design problems. At times, open conflicts between the different groups emerge, in particular between scientists and designers, since they have the least common ground.[1]

The Computer Human Interaction (CHI) Conference of the Association for Computing Machinery (ACM), which is the largest and arguably one of the most important conferences in the field, is organized through the Special Interest Group Computer Human Interaction (SIGCHI). At the 2005 SIGCHI membership meeting, discussion of the CHI 2006 conference ignited a shouting match between academics and practitioners.[2] This outbreak of emotion illustrates the tension between the different groups, and it can be explained by taking a closer look at their values, and at the barriers that separate them. Snow[3] was the first to talk about such barriers, even though he focused on only two cultures: the scientific and the literary intellectuals. While his political ideas have become somewhat obsolete with the decline of the USSR, his vision for the benefits of cooperating experts still holds:

> The clashing point of two subjects, two disciplines, two cultures—of two galaxies, so far as that goes—ought to produce creative chances. In the history of mental activity that has been where some of the break-throughs came.[4]

More recently, John Carroll[5] suggested that "An integrated and effective HCI can be a turning point in both disciplines and, perhaps, in human history." He also acknowledged several rifts that run through the HCI community. Cooperation remains difficult. At the CHI 2006 conference itself, this conflict was evident in the "Design: Creative and Historical Perspectives" session. Paul Dourish took the role of defending the science of ethnography against its degradation to a service for designers.[6] Next, Tracee Verring Wolf and Jennifer Rode defended creative design against the criticism of scientists by referring to design rigor that is as critical as scientific rigor.[7] Both groups felt the need to defend themselves, which showed that they had the feeling of being under attack. Stuart Feldman, president

1 Christoph Bartneck and M. Rauterberg, "HCI Reality—An Unreal Tournament," *International Journal of Human Computer Studies* 65:8 (2007): 737–43.

2 Jonathan Arnowitz and Elizabeth Dykstra-Erickson, "Chi and the Practitioner Dilemma," *Interactions* 12:4 (2005): 5–9.

3 C. P. Snow, *The Two Cultures: And a Second Look* (Cambridge: Cambridge University Press, 2nd ed., 1964).

4 Ibid.

5 J. M. Carroll, "Human-Computer Interaction: Psychology as a Science of Design," *Annual Review of Psychology* 48:1 (1997): 61–83.

6 Paul Dourish, "Implications for Design." Paper presented at the SIGCHI conference on Human Factors in Computing Systems, Montreal, Quebec, Canada, 2006.

7 Tracee Vetting Wolf, Jennifer A. Rode, Jeremy Sussman, and Wendy A. Kellogg, "Dispelling 'Design' as the Black Art of Chi." Paper presented at the SIGCHI conference on Human Factors in Computing Systems (CHI), Montreal, Quebec, Canada, 2006.

of the ACM, wrote another chapter in this conflict. In his opening speech at the CHI 2007 conference, he made an astonishing statement about the HCI community:

> It is also wonderful to have a group that is absolutely adherent to the classic scientific method. Not a description, I am afraid, of all the fields in computing.

However, it is obvious that the methods used by the HCI community are as diverse as its members. So, by emphasizing the classical scientific method above all other methods, Feldman was expressing the ACM's expectation of what methods the HCI community should use. This preference for the scientific method also manifests itself in the division of the CHI proceedings into "main conference proceedings" and "extended abstracts." The "main" proceedings are considered to be of higher quality, as the name already suggests, and they include a high proportion of scientific studies. These papers usually receive more presentation time at the conference. Nonscientific studies, such as experience reports and case studies, are more often found in the extended abstracts. This division also is reflected in Citeseer's estimates: the impact of the main proceedings is 1.61; while his estimate for the CHI extended abstracts is 0.51 (as of May 2003). Furthermore, the main proceedings use the "archival format," while the extended abstracts do not. The omission of the term "archival" from the format of the extended abstracts suggests that these publications are not important enough to be archived. However, both types of publications are being stored in the ACM digital library, which turns this distinction into a symbolic gesture. At the risk of oversimplification, it can be observed that scientific studies are more highly regarded, and thus published in the archival main proceedings; while nonscientific studies are less highly regarded, and are published only in the non-archival, extended abstracts. But why would the designers bother about this division? Their main focus is on improving society directly through the invention of artifacts, and not through writing papers. Chalmers gives us a hint:

> Science is highly esteemed. Apparently, it is a widely held belief that there is something special about science and its methods. The naming of some claim or line of reasoning or piece of research "scientific" is done in a way that is intended to imply some kind of merit or special kind of reliability.[8]

Being able to use such a powerful term to describe one's own activities is very attractive. Declaring that something has been designed does not carry as much value as declaring that something is scientifically proven. References to science also are frequently used to advertise products. For years, the Odol mouth refresher produced by GlaxoSmithKline claimed on its label: "According to the current state of scientific knowledge it can be shown that Odol is particularly

8 A. F. Chalmers, *What Is This Thing Called Science?* (Indianapolis: Hackett, 3rd ed., 1999).

to be recommended for complete mouth care." It follows that the denial of the status "scientific" to a study is a negative value judgment. Engineers, in particular, are very sensitive to the view that technology is hierarchically subordinate to science, serving only to deduce the implications of scientific discoveries, and to give them practical application. This sensitivity becomes clear in provocative statements by engineers such as Vincenti:

> Airplanes are not designed by science, but by art, in spite of some pretence and humbug to the contrary.[9]

Even though science is highly esteemed, Chalmers[10] argued that "There is no general account of science and scientific method to be had that applies to all sciences at all historical stages in their development." Cross, Naughton, and Walker[11] even suggested that the confusing epistemology of science may be unable to function as a blueprint for the epistemology of design. Levy[12] then suggested that transformations within the epistemology of science should be seen as active growth and development, and that they should be considered as providing an opportunity for design to participate in its ongoing improvement. As a matter of fact, any person can, in principle, contribute to the growth of science. It is an old rule of logic that the competence of a speaker has no relevance to the truth of what he says. The world's biggest fool can say the sun is shining, but that doesn't make it dark outside.[13] Designers and engineers can discover new knowledge without applying the classic scientific method or becoming a scientist. The more important question is how valuable this new knowledge is, and how efficient their methods are in finding it. Therefore, in the first part of this paper, I would like to discuss criteria that serve to assess the quality of knowledge. If design wants to make a contribution to science, then its insights must be judged against these criteria. By comparing the quality criteria of science with those of traditional design, the similarities and differences of the respective communities will become apparent. This comparison also may provide insights into the direction in which design methods have to evolve to become more scientific. This study does not attempt to discuss nonscientific knowledge that designers create for other designers. Many design books, for example, provide hands-on and relevant knowledge for the design practitioner. This knowledge does not attempt to be scientific, and it will not help to define a design science. It is, therefore, not in the scope of this study.

This comparison of quality criteria does also not imply that design *should* use the classical scientific method. Cross provided an excellent historical review of the developments in the various design methodologies.[14] He attested to a healthy growth in the field during the 1980s. The design community may continue to define its own method to turn itself into design science, as was attempted at the CHI 2007 workshops on "Converging on a Science of Design through the Synthesis of Design Methodologies,"[15] and on "Exploring

9 Walter G. Vincenti, *What Engineers Know and How They Know It: Analytical Studies from Aeronautical History*, Johns Hopkins Studies in the History of Technology (Baltimore: Johns Hopkins University Press, 1990).

10 A. F. Chalmers, *What Is This Thing Called Science?* (Indianapolis: Hackett, 3rd ed., 1999).

11 Nigel Cross, John Naughton, and David Walker, "Design Method and Scientific Method," *Design Studies* 2:4 (1981): 195–201.

12 R. Levy, "Science, Technology and Design," *Design Studies* 6:2 (1985): 66–72.

13 Robert M. Pirsig, *Zen and the Art of Motorcycle Maintenance: An Inquiry into Values* (New York: Morrow, 1974).

14 Nigel Cross, "Science and Design Methodology: A Review," *Research in Engineering Design* 5:2 (1993): 63–69.

15 Gerhard Fischer, Elisa Giaccardi, Yunwen Ye, Chris DiGiano, and Kumiyo Nakakoji, "Converging on a Science of Design through the Synthesis of Design Methodologies." Paper presented at the CHI '07 Extended Abstracts on Human Factors in Computing Systems, San Jose, CA, 2007.

Design as a Research Activity.[16] In the second part of this paper, I will challenge this goal of defining a design science, and try to outline a possible solution.

Before diving into these topics, it appears necessary to clarify the terminology of this paper. The different interpretations of the word "research" alone account for considerable friction between designers and scientists. Scientists can barely resist pointing out that designers' research does not provide reliable and valid knowledge. It follows that design decisions made on this basis are also in doubt.

First, we need to distinguish between the verb "research" and the noun "research." When designers, in particular practitioners, do research they predominantly collect relevant information. For scientists, "to research" describes the activity of conducting science, and the noun "research" is used as a synonym for "science." Since there is no verb form of "science," it appears necessary to continue to use the verb "research" for it. It follows that the activities of designers to collect information must be labeled with a different term, and "to explore" appears a good choice. A design science project that does not use the classical scientific method can then be described as "an exploration." Having clarified this important term, we may now proceed to discuss the quality criteria. The scientific reader may be familiar with them, and hence there is a danger of preaching to the converted. However, the comparison with related criteria in design may still be enlightening. Many of the concepts discussed are still under discussion. The meaning of truth, for example, has been disputed for more than 2000 years, and one can easily get lost in the labyrinth of arguments. The intention of this study is to attempt an overview, and the interested reader may further indulge in the specific topics by consulting the references provided. The overview might also help designers to better judge the quality of scientific studies they intend to utilize.

Quality Criteria for Science and Design

The generalizability of scientific knowledge is one of the most important criteria. It describes the degree to which general statements can be derived from a particular statement. The more general statements that can be derived, the better the particular statement. Newton's law of gravity was not only able to describe the behavior of the apple that inspired him, but also all other apples, fruits, organic materials, and inorganic materials. Even the motion of the stars could be described by it. His law, therefore, is of high value. If, on the other hand, a statement depends on the individual researcher, then its generalizability is low. If I state: "Bugs are awful." this may hold true only for people who share my paranoia about small creatures with many legs. Thus, objectivity is a good method for increasing the generalizability of a statement. Generalizability also is related to the repeatability of an experiment. If the results of an experiment are objective, meaning

16 Matthew R. Peters, Steven Haynes, David J. Saab, Helena Mentis, and Abigail Durrant, "Exploring Design as a Research Activity." Paper presented at the CHI '07 Extended Abstracts on Human Factors in Computing Systems, San Jose, CA, 2007.

that they are not dependent on the experimenter, then others should be able to repeat the experiment with exactly the same results. But, of course, not every repeatable experiment is automatically highly generalizable. If an experiment is conducted with only undergrad psychology students as participants, it may easily be repeatable; but the results may not be generizable to senior citizens.

Designers know a similar concept: "universality." It describes the degree to which general problems can be solved by a particular solution. The more universal a solution is the better. A hammer, for example, is more universal than a pair of horseshoe pliers, and hence more valuable. However, there often is a tradeoff between effectiveness and universality. Specific solutions usually work better than general solutions at the price of having to create a solution for each problem. The challenge is to find the right balance between universality and effectiveness. Science, on the other hand, strives towards the highest level of generalizability.

The knowledge that designers typically create in their design projects, suffers from its lack of generalizability. The solutions found for a given problem are limited to the scope of that problem, and cannot be applied easily, if at all, to different problems. Also, the solutions are dependent on the individual designer. A different designer might have come up with a different solution.

Falsifiability is another important criterion that is known to both scientists and designers. Originally proposed by Karl Popper in 2002, "falsifiability" describes the property of statements that they must admit of logical and empirical counterexamples. The latter refers to the condition that it must be possible, at least in principle, to make an observation that would show the statement to be wrong, even if that observation is not actually made. The statement "all swans are white" is in principle falsifiable by observing a black swan. The higher the number of logical and empirical counterexamples that a statement withstands, the higher is its value.

The use of falsifiability in design is similar. A solution must admit the existence of logical and empirical counterexamples. If, for example, a certain device is intended to continuously increase one's karma, then its function is impossible to falsify. Such a device could not be considered a design. Falsifiability plays a less important role in design in comparison with science, since it often deals with concrete and well-defined problems. The effects of a solution usually are easy to observe, and this criterion overlaps with the criterion of effectiveness that will be discussed later.

Truth is a key criterion in science, and it also plays an important role in design. However, its definition is still an object of philosophical discussion, and so multiple definitions of truth exist at this point in time. An in-depth discussion about the underlying philosophical issues is beyond the scope of this paper, but the interested reader may consult Kirkham,[17] who offers a survey of all the major philosophical theories of truth. The acknowledged Wikipedia[18] alone

17 Richard L. Kirkham, *Theories of Truth: A Critical Introduction* (Cambridge, MA: MIT Press, 1992).

18 Jim Giles, "Internet Encyclopaedias Go Head to Head," *Nature* 438:7070 (2005): 900–1.

lists many theories of truth including correspondence, coherence, constructivist, consensus, pragmatic, performative, semantic, and Kripke's theory. The correspondence and coherence theories probably are the most acknowledged, so this study focuses on them. In the coherence theory, truth is primarily a property of a whole system of statements. The truth of a single statement can be derived only from its accordance with all the other statements. If a new statement contradicts an existing statement, then both statements need to be reconsidered. In the previously used example of swans, one of the statements must be false. Either not all swans are white or the particular swan is not black. The equivalent concept in design is known as "compatibility." If a new component is introduced to an existing system, then it should not prevent any existing component from operating correctly. For example, the installation of new software on a computer can lead to incompatibilities in which previous functions cease to operate.

The correspondence theory of truth deals with the relationship between statements and reality. If theories correspond to observations in reality, they are considered to be true. This direction in the relationship between truth and reality usually is attributed to science. The other direction can be attributed to design. If an artifact corresponds to theory, then it is considered true. Our understanding of the physical world makes it difficult to invent artifacts that could not be explained fully by existing theories of physics. Many attempts have been made to invent a perpetual motion machine, and patents have even been filed, but no working model has ever been built. The United States Patent and Trademark Office (USPTO) refuses to grant patents for perpetual motion machines without a working model:

> With the exception of cases involving perpetual motion, a model is not ordinarily required by the Office to demonstrate the operability of a device. - 608.03 Models, Exhibits, Specimens [R-3].

However, solutions often have been used without full theoretical understanding. The Bayer Company patented aspirin as early as 1899, and has successfully marketed it ever since. Its pain-relieving effect was not understood until 1971. In 1982, John Robert Vane received the Nobel Prize in the Physiology of Medicine for this discovery.

Another important quality criterion for scientific knowledge is novelty. Rediscovering Newton's Law has little value. But newness in itself is not sufficient. A novel scientific theory does not only need to be different from existing theories, but it also has to explain more than existing theories. Galileo's theories extended Aristotle's; Newton's Law extended Galileo's theories; and Einstein's extended Newton's. In design, the same principle is known as "innovation." Novelty, in its pure "newness" definition, is even a requirement for patents. Moreover, new artifacts are expected to work not only differently, but also better. Modern PCs are currently even powerful enough to completely simulate older computers, for example, simulating the Commodore 64 using the VICE emulator. Modern PCs can do everything that older ones can, and more.

The criterion of parsimony, also known as "Occam's razor," is the preference for the least complex statement to explain a fact. A good example can be found in the field of astronomy. The Copernican model is said to have been chosen over the Ptolemaic due to its greater simplicity. The Ptolemaic model, in order to explain the apparent retrograde motion of Mercury relative to Venus, posited the existence of epicycles within the orbit of Mercury. The Copernican model (as expanded by Kepler) was able to account for this motion by displacing the Earth from the center of the solar system and replacing it with the Sun as the orbital focus of planetary motions, while simultaneously replacing the circular orbits of the Ptolemaic model with elliptical ones. In addition, the Copernican model excluded any mention of the crystalline spheres that the planets were thought to be embedded in according to the Ptolemaic model. At a single stroke, the Copernican model reduced the complexity of astronomy by a factor of two.

In design, simplicity plays a similar role. Simplicity is the preference for the least complex solution to achieve a given goal. Just twenty years ago, the only way to print a photo required a complete photochemical process that involved various toxic chemicals and sophisticated machines. These days, everybody can print his own pictures with cheap inkjet printers.

Finally, the scientific criteria of accuracy, precision, and efficiency are discussed, together with their counterparts in design: effectiveness, reliability, and efficiency. Accuracy refers to the degree to which a statement or theory predicts the facts it is intended to predict, while precision refers to the degree to which a statement or theory predicts the exact same facts. The analogy of bullets shot at a target is useful to explain the difference between these two related concepts and, at the same time, to show the similarity between design and science criteria.

In this analogy, a gun firing at a target (design) parallels a theory predicting observations (science). The "effectiveness" of the gun describes the closeness of the bullets to the center of the target (see Figure 1). Bullets that strike closer to the center are considered more effective. The parallel is that the closer the observations concur with the predictions of the theory, the more accurate the theory.

Figure 1
High effectiveness but low reliability (left); high reliability but low effectiveness (middle); and high reliability and high effectiveness (right).

To continue the analogy, the reliability of the gun refers to the spread of the bullets. The closer together the bullets strike, the higher the reliability (see Figure 1, middle). In science, the closer the observations are to each other, the more precise is the theory. The bullets do not necessarily need to be close to the center for this. The bullets (or observations) can be reliable (precise) without being effective (accurate). However, for bullets (and observations) to be perfectly effective (accurate), they also need to be reliable (precise) (see Figure 1).

For science, efficiency refers to the resources expended in relation to the precision and accuracy of the observations predicted, and for design, efficiency refers to the resources expended in relation to the effectiveness and reliability of the goals achieved.

So far, only those quality criteria of design that have a direct relation to the quality criteria of science have been discussed. Of course, design also has criteria that are of less relevance to science. Conformity to social customs, popularity, ego satisfaction, reputation, pleasure, and commercial success are examples. It is difficult to define general design criteria, since each design can be judged only in its specific context of use. The Hummer sport utility vehicle (SUV), for example, is a car that is not intended to be environmental friendly, so it should not be judged by the fuel consumption criterion. The Hummer SUVs are not designed for driving fuel-efficiently from points A to B.

Conclusions on Quality

At first sight, there appears to be a considerable overlap in the quality criteria for design and science. Pirsig[19] attested that they are just two different complementary ways of looking at the same thing. At the most immediate level (dynamic quality), they have never been separate. Pirsig provides a description of dynamic quality as being "the continuing stimulus which our environment puts upon us to create the world in which we live."[20] The sense of quality that guides a scientist in selecting a certain hypothesis for further investigation is the same sense for quality that helps designers to choose one solution over another. Both disciplines are creative: designers create primarily artifacts, and scientists primarily knowledge. This similarity may mislead people into believing that design already is a science. In the next section, I will discuss the possibility of a design science. For now, I would like to characterize the different quality criteria using the Metaphysics of Quality (MOQ) framework.[21] Of course, it is impossible to provide an exhaustive description of his work within the limited space available, but the interested reader may consider Anthony McWatt's thesis as a starting point for delving deeper into this philosophical idea.[22]

Besides dynamic quality, Pirsig distinguishes four static quality patterns (see Figure 2) that have evolved over time, and that are ordered in a hierarchy from the inorganic (lowest coherence) to the intellectual (highest coherence). The highest pattern contains

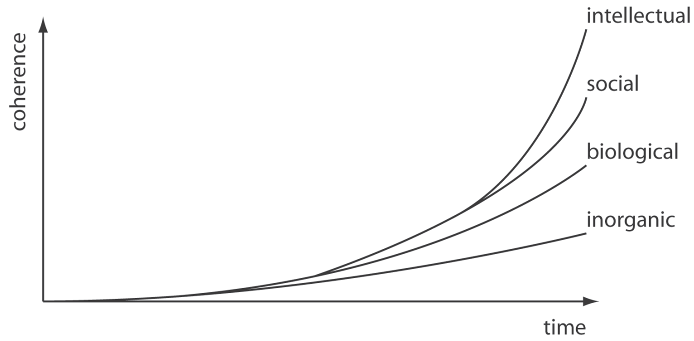
19 Robert M. Pirsig, "Subjects, Objects, Data & Values." Paper presented at the Einstein Meets Magritte Symposium, Vrije University, Brussels, 1995.

20 Robert M. Pirsig, *Zen and the Art of Motorcycle Maintenance: An Inquiry into Values* (New York: Morrow, 1974).

21 Robert M. Pirsig, *Lila: An Inquiry into Morals* (New York: Bantam Books, 1991).

22 Anthony McWatt, "An Introduction to Robert Pirsig's Metaphysics of Quality." PhD thesis, University of Liverpool, 2003.

Figure 2
The evolutionary order of static quality patterns.



intellectual patterns such as theology, science, and philosophy. This pattern is confined to the skilled manipulation of abstract symbols that have no specific corresponding experience, and behave according to rules of their own.[23] Social patterns include such institutions as family, church, and government. They are the patterns of culture that the anthropologist and sociologist study.

The hierarchical structure of the patterns suggests that the intellectual pattern takes precedence over social, biological, and inorganic matters. During the Renaissance, science rejected religious dogmas, social prejudices, and biological emotions to set its own higher intellectual patterns such as truth. When science is mixed with social patterns, such as religion, it can quite correctly be argued that these patterns corrupt science. The times in which the church could dictate truth are hopefully over, and even Galileo was rehabilitated by Pope John Paul II in 1992. However, the recent upsurge of creationism is alarming, and we can only hope that the intellectual value of truth will continue to prevail over religious beliefs, as it did in the Dover, Pennsylvania, case.[24] In this court case, the attempt of creationists to introduce their religious ideas into the biology science class of a Dover public high school was successfully blocked. In a similar fashion, social patterns are superior to biological patterns. The value of a stable family overrides the value of spreading one's genes as much as possible through adultery.

While there is a hierarchy of quality patterns, they are still dependent on each other. Every intellectual pattern also is a social and biological pattern, but not every social pattern is an intellectual one. The ideas of science come from scientists who work in a community and need food to survive. However, not every community produces scientific knowledge, and not every animal forms communities.

The quality criteria of design (universality, falsifiability, compatibility, correspondence to theory, novelty, simplicity, reliability, effectiveness, and efficiency) operate predominately at the social level. Designers, in particular practitioners, create artifacts to transform the world into a desired state.[25] Their results are essential to society, but remain subordinate to the intellectual level at which science operates. Notice that "intellectual" does not refer

23 Robert M. Pirsig, "Letter to Paul Turner Dated September 23rd, 2003."

24 Judge Jones, *Tammy Kitzmiller, et al. v. Dover Area School District, et al.* Case 4:04-cv-02688-JEJ (2005): 1–139.

25 Herbert A. Simon, *The Sciences of the Artificial* (Cambridge, MA: MIT Press, 3rd ed., 1996).

to the social title of being an intellectual, but to the quality pattern described above. Of course, designers can be intellectuals. With respect to dynamic quality, design and science are similar. But in the framework of static quality patterns, they differ. If design wants to contribute to the growth of scientific knowledge, then it will primarily have to improve the generalizability of its results. Most of all, to guarantee objectivity, its results need to become independent of the designer. Pitt claimed[26] that such a method would lead to knowledge that is "far more reliable, secure, and trustworthy than scientific knowledge."

So far, I have considered design in its classic form in which it does not qualify as a science. Still, it does have the potential to contribute to the growth of scientific knowledge. Next, I would like to discuss the challenges faced by design when it attempts to become a science.

The Challenge of a Design Science

Science consists of a method for observing reality and abstracting it into models that are then used to explain and predict reality. Newton's law of gravity, for example, explains why an apple hit Isaac Newton, and it also helps us to predict the position of the planets in the future. The various sciences claim certain parts of reality as their phenomena under investigation.

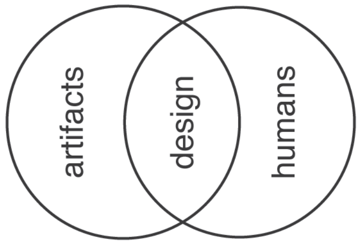
The methods of science are to some degree universal, and often are referred to as "scientific method." The scientific method is a body of techniques for investigating phenomena and acquiring new knowledge, as well as for correcting and integrating existing knowledge. It is based on gathering observable, empirical, measurable evidence, subject to the principles of reasoning. Chalmers[27] provides an impartial discussion of the scientific method, and this is probably what Stuart Feldman had in mind when he referred to the classic scientific method. However, a methodology in itself can never constitute a science. Let us take the example of the dissection method. Biologists may use dissection to analyze animals, but butchers also use it to cut steaks. The method is the same, but one results in scientific knowledge, the other in a delicious meal. Moreover, in the same way that biology is not a science of how biologists work, design science cannot be a science of how designers work. This conceptual limitation cannot be overcome even by converging on a specific design method. Again, a method does not constitute a science, and design methodologies cannot be the phenomena of design science.

The sciences distinguish themselves not through their methods, but through the phenomena they investigate. Biology, for example, is the science of living organisms. What a design science is primarily missing is a phenomenon. This demarcation problem becomes clearer when we consider that the prime objective of design lies at the intersection between artifacts and users (see Figure 3).

26 Joseph C. Pitt, "What Engineers Know" *In Techne* 5:3 (2001): 17–30.

27 A. F. Chalmers, *What Is This Thing Called Science?* (Indianapolis, IN: Hackett, 3rd ed., 1999).

Figure 3
The phenomenon of design.



Designers contribute to the creation of artifacts that interact with humans.

Everything there is to know about the artifact (Figure 3) is available from its manufacturer. Its dimensions, material properties, and functions are known. The artifacts, therefore, are not good phenomena for investigation. Also the creation of new materials and operational principles already has been claimed by engineering and physics. Engineers also discuss rational design methodology that relies heavily on mathematics.[28] Interestingly, these rational design methodologies are not often used in the area of design, even though they have one fundamental characteristic that brings them closer to science: the results produced through these methods are objective. This means that the results are independent of the designer who applies them. This independence is a major step forward in the direction of generalizability. When we take a look at the body of scientific knowledge, we also see that it is engineers who have attempted to create a consistent and logical body of knowledge for design solutions.[29]

On the other hand (Figure 3), understanding human beings is the prime objective of medicine, anthropology, and psychology. Design science would have difficulty in competing. Even "design methodology" or, to be more general, "solving human problems," already has been treated as a phenomenon investigated by psychologists.[30]

As we can see, both artifacts and humans have been claimed as phenomena by physics, engineering, psychology, and medicine. The definition of a design phenomenon is possibly the most urgent step in the development of a design science. The arena of design science is filled with actors from many different disciplines, and one may then ask why the designers in the HCI community are so keen on turning design into a science?

It is a noble goal to create good and reliable design that improves society, but this cannot be achieved by using the scientific method, nor can the claim of a design science be a good response to the criticisms of scientists. Not everything has to be scientific, and designers are playing an important role in the creation of artifacts. They should be proud of the role they already play in the HCI community.

If we attempt to turn design into a science, then we face the demarcation problem that will be difficult to overcome unless

28 Walter G. Vincenti, *What Engineers Know and How They Know It: Analytical Studies from Aeronautical History*, Johns Hopkins Studies in the History of Technology (Baltimore, MD: Johns Hopkins University Press, 1990). Herbert A. Simon, *The Sciences of the Artificial.* Christopher Alexander, *Notes on the Synthesis of Form* (Cambridge, MA: Harvard University Press, 1964).

29 Walter G. Vincenti, *What Engineers Know and How They Know It: Analytical Studies from Aeronautical History.* Vladimir Hubka and W. E. Eder, *Design Science: Introduction to Needs, Scope and Organization of Engineering Design Knowledge* (Berlin and New York: Springer, 1996).

30 J. Dorfman, V. A. Shames, and J. F. Kihlstrom, "Intuition, Incubation, and Insight: Implicit Cognition in Problem Solving" in *Implicit Cognition*, G. Underwood, ed. (Oxford: Oxford University Press, 1996). G. Feist, "The Affective Consequences of Artistic and Scientific Problem Solving," *Cognition and Emotion* 8 (1994): 489–502.

Figure 4
Ontology of Pirsig's Metaphysics of Quality

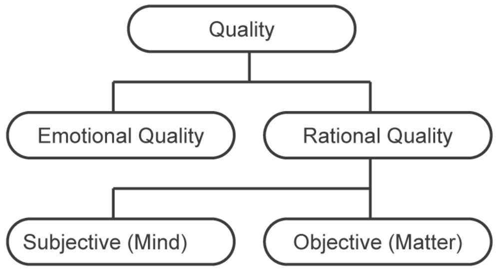
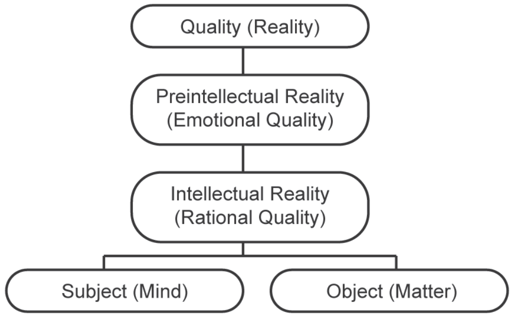


Figure 5
The process of quality.



we resolve the subject-object dichotomy that was highlighted by Descartes.[31] In this duality, objective knowledge is superior to subjective knowledge, and together they constitute an antagonistic relationship that constantly generates dichotomies: mind and matter, science and art, and feeling and reason.[32] The Metaphysics of Quality has the potential to overcome this dichotomy. It reduces this duality to a secondary role, and places quality alone at the top (see Figure 4) as "the parent, the source of all subjects and objects."[33]

Even though quality itself cannot be defined, its existence can be proved. Pirsig provides a pragmatic proof by subtracting it from the description of our world, and showing that a world without quality would be dysfunctional.[34] To understand this statement, one has to detail the process of quality (see Figure 5). Our environment presents us with a quality stimulus, which we sense pre-intellectually before we intellectualize it, thereby dividing it into subjects and objects.

The pre-intellectual sensitivity to quality can be compared to Kant's a priori pure cognition of time and space. It can even be argued that quality might fulfill Kant's requirements of necessity and universality, and thus may be considered a third form of a priori pure cognition. Pirsig describes the sense of quality this way:

> This sense [for quality] is not just something you are born with, although you are born with it. It is also something you can develop. It is not just "intuition" not just unexplain-

31 R. Descartes, *Principia Philosophiae* (Amsterdam: Danielem Elzevirium, 1644).

32 George Basalla, "Man and Machine," *Science* 187 (1975): 248–50.

33 Robert M. Pirsig, *Zen and the Art of Motorcycle Maintenance: An Inquiry into Values*.

34 Ibid., 193.

> able "skill" or "talent." It is the direct result of contact with basic reality, Quality, which dualistic reason has in the past tended to conceal.[35]

Dynamic quality, and its division into classical and romantic quality, is of most importance, since it is the point connecting science and design:

> What relates science to the arts [design] is that science explores the Conceptually Unknown [dynamic quality] in order to develop a theory that will cover measurable patterns emerging from the unknown. The arts [design] explore the Conceptually Unknown [dynamic quality] in other ways to create patterns such as music, literature, painting, that reveal the Dynamic Quality that produces them. (Square brackets were added by the author)[36]

The MOQ's view of the world as shown in Figure 4 is now able to overcome the subject-object dichotomy and the difficulties it creates for defining the phenomena of design science. Design science should focus on what is inherent to both subjects and objects: quality. Design science is the science of quality.

Discussion

Science has established several criteria for assessing the quality of the knowledge it produces. Some of these criteria overlap or relate to criteria that are used in design. Design methods are not yet optimized for the creation of scientific knowledge, and therefore they generally produce knowledge that is of lesser scientific quality. Their weakest area is generalizability, since the knowledge produced is often based on the individual designer. Currently, designers who want to work as scientists often have to become either engineers and work within a rational problem-solving framework,[37] or they can choose to become psychologists. John Carroll suggested that psychology could be considered a science of design.[38] Since designers often lack training in these disciplines, they have a natural disadvantage. It would be preferable if they could become scientists without becoming bad engineers or bad psychologists. To create a design science, we first need to define its phenomena. This can be achieved by overcoming the subject-object dichotomy. The Metaphysics of Quality has the potential to bridge the gap between subjects and objects, and by doing so it also defines the phenomena of design science: quality. The formal recognition of quality will then also have a direct influence on science:

> I think that it will be found that a formal acknowledgement of the role of Quality in the scientific process does not destroy the empirical vision at all. It expands it, strengthens it and brings it far closer to actual scientific practice… By returning our attention to Quality it is hoped that we can get technological work out of the non-caring subject-object

35 Ibid., 225.

36 Robert M. Pirsig, "Subjects, Objects, Data & Values."

37 Herbert A. Simon, *The Sciences of the Artificial*.

38 J. M. Carroll, "Human-Computer Interaction: Psychology as a Science of Design," *Annual Review of Psychology* 48:1 (1997): 61–83.

> dualism and back into craftsmanlike self-involved reality again, which will reveal to us the facts we need when we are stuck.[39]

But would knowledge about quality be generalizable? A common criticism of Pirsig's Metaphysics of Quality is the question of how we could possibly disagree on quality when quality is supposed to be universal. How is it possible that we have difficulties agreeing on which is the better poem? Pirsig considers that our previous experiences influence our perception of quality:

> The names, the shapes and forms we give quality depend only partly on the quality. They also depend partly on the a priori images we have accumulated in our memory. We constantly seek to find, in the quality event, analogues to our previous experiences. If we didn't, we'd be unable to act. We build up our language in terms of these analogues… The reason people see Quality differently is because people come to it with different sets of analogues.[40]

Pirsig speculated that, if two people had identical a priori analogues, they would see quality identically every time. This still would not explain why listening to a new record over and over again can change the experience from being exciting to being boring. In his second book,[41] Pirsig modifies his previous division of quality (classical and romantic) to static and dynamic quality, in which "dynamic quality" represents both romantic and classical quality. The main advantage of this division is that it prevents the perception that quality consists of two fundamentally different types: classical and romantic. Furthermore, dynamic quality can more easily include mystical experiences which, according to Pirsig, are not well described by romantic quality. The division between dynamic and static quality can then easily explain the change of experience after listening to a record numerous times. It changes from a dynamic to a static quality pattern. This new division also provides a different way of looking at the disagreement on which poem is better. While dynamic quality is the same for everyone, and therefore generalizable, static quality patterns depend on the individual's prior experiences. When judging a poem, we use both dynamic quality and static quality, which results in some uniformity between individuals; but not complete uniformity.

What is not yet clear is the degree to which the classical scientific method can be used to investigate quality because it is based on the assumption that subjects and objects are not connected. The investigator is not supposed to influence the phenomenon under investigation. Quality is the source of subjects and objects, so a method that presupposes that they are not connected may be unsuitable for investigating it.

The main challenge for a design science is the definition of a method that may be capable of investigating quality, but at least

39 Robert M. Pirsig, *Zen and the Art of Motorcycle Maintenance: An Inquiry into Values*. (p. 224)

40 Ibid.

41 Robert M. Pirsig, *Lila: An Inquiry into Morals*.

some first attempts have been made. Nakashima et al.[42] attempted to make implicit quality knowledge of executing a task explicit in order to measure if this knowledge would help a second user improve his performance. Zimmerman et al.[43] emphasizes the benefit of creating artifacts for HCI research, and also propose four criteria for the evaluation of its success (process, invention, relevance, and extensibility).

Compared to this methodological challenge for design science, the popular separation of design and science into activities that transform the world into a desired state (design) versus activities that attempt to understand the world (science) can easily be overcome. The prime reason for the existence of science is the assumption that a state of knowing is better than a state of ignorance. By definition, this turns all scientific activities into design activities. It is unlikely that many scientists would feel comfortable with this classification of their work, but the fundamental argument remains.

Until considerable progress has been made in defining a suitable epistemology for design science, we shall have to take small steps forward using current methods and policies. Design has to acknowledge that the knowledge it produces is, from a scientific perspective, not very generalizable, and thus of lesser value. However, from a design perspective, it may very well be valuable since its concreteness makes it easy to use in everyday practice. Schön[44] even argued that "wicked design problems" may resist formalization, and hence cannot be approached with a pure scientific approach. The concrete design knowledge may then be the best approximation to scientific knowledge.

Design also has to acknowledge that its focus is on the social, and not on the intellectual, level. Scientists, on the other hand, need to acknowledge that the highly general knowledge they produce often is too abstract to improve society. It requires a skilled designer to translate this knowledge into a specific context of use.

The hierarchy levels of static quality patterns may even justify the division of the CHI proceedings into sections. However, it would be wise to follow Confucius' recommendation to "rectify the names." Labeling only one section "archival" when both sections will be stored in the ACM Digital Library is confusing. Also, the labels "main proceedings" and "extended abstracts" are ambiguous. Pirsig's quality patterns appear suitable for defining the sections, but the terms "intellectual" and "social" carry different meanings in the various sub-communities, and may cause misunderstandings. Maybe the sections could be called "Discovery" and "Invention." The latter would collect contributions that are aimed at improving society, and that operate on Pirsig's social level. The discovery section would gather contributions that present scientific insights, and operate at the intellectual level. Whatever principle is used to divide the proceedings, it should be made explicit.

The use of "best paper" awards is another ranking method.

42 Hideyuki Nakashima, Masaki Suwa, and Haruyuki Fujii, "Endo-System View as a Method for Constructive Science." Paper presented at the 5th International Conference of the Cognitive Science, Vancouver, BC, Canada, 2006.

43 John Zimmerman, Jodi Forlizzi, and Shelley Evenson, "Research through Design as a Method for Interaction Design Research in HCI." Paper presented at the SIGCHI Conference on Human Factors in Computing Systems (CHI), San Jose, CA, 2007.

44 Donald A. Schön, *The Reflective Practitioner: How Professionals Think in Action* (Aldershot, UK: Arena, 1991).

Excellence should be rewarded. However, rankings should not be used to discriminate between communities. Excellence can be found in design papers as well as in scientific papers. The factors that influence paper rankings should be made explicit. This would require the agreement of the community on the factors used. The CHI community is diverse, and it may be difficult to reach agreement. But nothing worthwhile is ever easy. As long as no shared quality criteria are defined for the community as a whole, it will remain a trans–disciplinary rather than a multidisciplinary community. We will continue to tolerate each other rather than contributing to each other's success. The sub-communities of design, education, engineering, management, research, and usability will coexist, but future shouting matches cannot be excluded.

Conclusion

The comparison of the quality criteria used in design and science hopefully creates a better mutual understanding between these two communities, and a guide towards a design science. The classical scientific method does not appear to be suitable for the investigation of the phenomenon of design science: quality. While this study is not able to propose a new investigation method, it at least pointed out one of its requirements. Based on the Metaphysics of Quality, it has been concluded that the method of design science must overcome the subject-object dichotomy. This ambitious goal may not be achieved quickly, and thus we also need to focus on the small steps to improve the HCI community. The observations and suggestions we made for the CHI conference may create a better and more equal framework in which the different sub-communities inspire each other.