DESIGN SCIENCE

A Module for the Master of Industrial Design

Christoph Bartneck
Eindhoven University of Technology, Department of Industrial Design
Room HG 2.51, Den Dolech 2. 5600 MB Eindhoven, Netherlands

And what is good, Phædrus,
And what is not good...
Need we ask anyone to tell us these things?

Understanding the status of design as a form of research is both important and problematic. The National Science Foundation's Science of Design program, which highlights the role of design in the development of interactive systems, and the First International Conference on Design Science Research in Information Systems and Technology both point to the need for a community concerned with articulating different visions of design science and design research.

This module will shed some light on the underlying concepts of design and science and their inherent conflicts. The students will reflect on their role as a designer and the methodologies they follow. In addition, the quality criteria that are used in design and science to evaluate the value of the produced knowledge and artifacts will be considered. This documents compiles relevant literature for the Design Science module.

Literature included:

Santa Fe High School student will unquestionably perceive the inevitable pregame prayer as stamped with her school's seal of approval.").

We arrive at this conclusion by initially considering the plain language of the disclaimer, paragraph by paragraph. The first paragraph reads as follows:

> The Pennsylvania Academic Standards require students to learn about Darwin's Theory of Evolution and eventually to take a standardized test of which evolution is a part.

P-124. The evidence in this case reveals that Defendants do not mandate a similar pronouncement about *any other aspect of the biology curriculum or the curriculum for any other course*, despite the fact that state standards directly address numerous other topics covered in the biology curriculum and the students' other classes, and despite the fact that standardized tests cover such other topics as well. Notably, the unrefuted testimony of Plaintiffs' science education expert Dr. Alters, the only such expert to testify in the case sub judice explains, and the testimony of Drs. Miller and Padian confirms, the message this paragraph communicates to ninth grade biology students is that:

> [W]e have to teach this stuff[.] The other stuff we're just going to teach you, but now this one we have to say the Pennsylvania academic standards require[] students to . . . eventually take a test. We'd rather not do it, but Pennsylvania academic standards . . . require students to do this.

Trial Tr. vol. 14, Alters Test., 110-11, Oct. 12, 2005.

Stated another way, the first paragraph of the disclaimer directly addresses and disavows evolutionary theory by telling students that they have to learn about evolutionary theory because it is required by "Pennsylvania Academic Standards" and it will be tested; however, no similar disclaimer prefacing instruction is conducted regarding any other portion of the biology curriculum nor any other course's curriculum.

The second paragraph of the disclaimer reads as follows:

> Because Darwin's Theory is a theory, it continues to be tested as new evidence is discovered. The Theory is not a fact. Gaps in the Theory exist for which there is no evidence. A theory is defined as a well-tested explanation that unifies a broad range of observations.

P-124. This paragraph singles out evolution from the rest of the science curriculum and informs students that evolution, unlike anything else that they are learning, is "just a theory," which plays on the "colloquial or popular understanding of the term ['theory'] and suggest[ing] to the informed, reasonable observer that evolution is only a highly questionable 'opinion' or a 'hunch.'" Selman, 390 F. Supp. 2d at 1310; 14:110-12 (Alters); 1:92 (Miller). Immediately after students are told that "Darwin's Theory" is a theory and that it continues to be tested, they are told that "gaps" exist within evolutionary theory without any indication that other scientific theories might suffer the same supposed weakness. As Dr. Alters explained this paragraph is both misleading and creates misconceptions in students about evolutionary theory by misrepresenting the scientific status of evolution and by telling students that they should regard it as singularly unreliable, or on shaky ground. (14:117 (Alters)). Additionally and as pointed out by Plaintiffs, it is indeed telling that even defense expert Professor Fuller agreed with this conclusion by stating that in his own expert opinion the disclaimer is misleading. (Fuller Dep. 110-11, June 21, 2005). Dr. Padian bluntly and effectively stated that in confusing students about science generally and evolution in particular, the disclaimer makes students "stupid." (Trial Tr. vol. 17, Padian Test., 48-52, Oct. 14, 2005).

In summary, the second paragraph of the disclaimer undermines students' education in evolutionary theory and sets the groundwork for presenting students with the District's favored religious alternative.

Paragraph three of the disclaimer proceeds to present this alternative and reads as follows:

> Intelligent Design is an explanation of the origin of life that differs from Darwin's view. The reference book, Of Pandas and People, is available for students who might be interested in gaining an understanding of what Intelligent Design actually involves.

P-124. Students are therefore provided information that contrasts ID with "Darwin's *view*" and are directed to consult Pandas as though it were a scientific text that provided a scientific account of, and empirical scientific evidence for, ID. The theory or "view" of evolution, which has been discredited by the District in the student's eyes, is contrasted with an alternative "explanation," as opposed to a "theory," that can be offered without qualification or cautionary note. The alternative "explanation" thus receives markedly different treatment from evolutionary "theory." In other words, the disclaimer relies upon the very same "contrived dualism" that the court in McLean recognized to be a creationist tactic that has "no scientific factual basis or legitimate educational purpose." McLean, 529 F. Supp. at 1266.[6]

---

[6] The McLean court explained that:

> The approach to teaching 'creation science' and 'evolution science' . . . is identical to the two-model approach espoused by the Institute for Creation Research and is taken almost verbatim from ICR writings. It is an extension of Fundamentalists' view that one must either accept the literal interpretation of Genesis or else believe in the godless system of evolution.
>
> The two model approach of creationists is simply a ***contrived dualism*** which has no scientific factual basis or legitimate educational purpose. It assumes only two explanations for the origins of life and existence of man, plants and animals: it was either the work of a creator or it was not. Application of these two models, according to creationists, and the defendants, dictates that all scientific evidence which fails to support the theory of evolution is necessarily scientific evidence in support of creationism and is, therefore, creation science 'evidence[.]'

529 F. Supp. at 1266 (footnote omitted)(emphasis added).

The overwhelming evidence at trial established that ID is a religious view, a mere re-labeling of creationism, and not a scientific theory. As the Fifth Circuit Court of Appeals held in Freiler, an educator's "reading of a disclaimer that not only disavows endorsement of educational materials but also juxtaposes that disavowal with an urging to contemplate alternative religious concepts implies School Board approval of religious principles." Freiler, 185 F.3d at 348.

In the fourth and final paragraph of the disclaimer, students are informed of the following:

> With respect to any theory, students are encouraged to keep an open mind. The school leaves the discussion of the Origins of Life to individual students and their families. As a Standards-driven district, class instruction focuses upon preparing students to achieve proficiency on Standards-based assessments.

P-124.

Plaintiffs accurately submit that the disclaimer mimics the one that the Fifth Circuit struck down as unconstitutional in Freiler in two key aspects. First, while encouraging students to keep an open mind and explore alternatives to evolution, it offers no scientific alternative; instead, the only alternative offered is an inherently religious one, namely, ID. Freiler, 185 F.3d at 344-47 (disclaimer urging students to "exercise critical thinking and gather all information possible and closely examine each alternative toward forming an opinion" referenced "Biblical version of Creation" as the only alternative theory, thus "encourag[ing] students to read and meditate upon religion in general and the "Biblical version of Creation" in particular.) Whether a student accepts the Board's invitation to explore Pandas, and reads a creationist text, or follows the Board's other suggestion and discusses "Origins of Life" with family members, that objective student can reasonably infer that the District's favored view is a religious one, and that the District is accordingly sponsoring a form of religion. Second, by directing students to their families to learn about the "Origins of Life," the paragraph performs the exact same function as did the Freiler disclaimer: It "reminds school children that they can rightly maintain beliefs taught by their parents on the subject of the origin of life," thereby stifling the critical thinking that the class's study of evolutionary theory might otherwise prompt, to protect a religious view from what the Board considers to be a threat. Id. at 345 (because disclaimer effectively told students "that evolution as taught in the classroom need not affect what they already know," it sent a message that was "contrary to an intent to encourage critical thinking, which requires that students approach new concepts with an open mind and willingness to alter and shift existing viewpoints").

A thorough review of the disclaimer's plain language therefore conveys a strong message of religious endorsement to an objective Dover ninth grade student.

The classroom presentation of the disclaimer provides further evidence that it conveys a message of religious endorsement. It is important to initially note that as a result of the teachers' refusal to read the disclaimer, school administrators were forced to make special appearances in the science classrooms to deliver it. No evidence was presented by any witness that the Dover students are presented with a disclaimer of any type in any other topic in the curriculum. An objective student observer would accordingly be observant of the fact that the message contained in the disclaimer is special and carries special weight. In addition, the objective student would understand that the administrators are reading the statement because the biology teachers refused to do so on the ground that they are legally and ethically barred from misrepresenting a religious belief as science, as will be discussed below. (Trial Tr. vol. 25, Nilsen Test., 56-57, Oct. 21, 2005; Trial Tr. vol. 35, Baksa Test., 38, Nov. 2, 2005). This would provide the students with an additional reason to conclude that the District is advocating a religious view in biology class.

Second, the administrators made the remarkable and awkward statement, as part of the disclaimer, that "there will be no other discussion of the issue and your teachers will not answer questions on the issue." (P-124). Dr. Alters explained that a reasonable student observer would conclude that ID is a kind of "secret science that students apparently can't discuss with their science teacher" which he indicated is pedagogically "about as bad as I could possibly think of." (14:125-27 (Alters)). Unlike anything else in the curriculum, students are under the impression that the topic to which they are introduced in the disclaimer, ID, is so sensitive that the students and their teachers are completely barred from asking questions about it or discussing it.[7]

---

[7] Throughout the trial and in various submissions to the Court, Defendants vigorously argue that the reading of the statement is not "teaching" ID but instead is merely "making students aware of it." In fact, one consistency among the Dover School Board members' testimony, which was marked by selective memories and outright lies under oath, as will be discussed in more detail below, is that they did not think they needed to be knowledgeable about ID because it was not being taught to the students. We disagree.

Dr. Alters, the District's own science teachers, and Plaintiffs Christy Rehm and Steven Stough, who are themselves teachers, all made it abundantly clear by their testimony that an educator reading the disclaimer is engaged in teaching, even if it is colossally bad teaching. See, e.g., Trial Tr. vol. 6, C. Rehm Test., 77, Sept. 28, 2005; Trial Tr. vol. 15, Stough Test., 139-40, Oct. 12, 2005. Dr. Alters rejected Dover's explanation that its curriculum change and the statement implementing it are not teaching. The disclaimer is a "mini-lecture" providing substantive misconceptions about the nature of science, evolution, and ID which "facilitates learning." (14:120-23, 15:57-59 (Alters)). In addition, superintendent Nilsen agrees that students "learn" from the statement, regardless of whether it gets labeled as "teaching." (26:39 (Nilsen)).

Finally, even assuming arguendo that Defendants are correct that reading the statement is not "teaching" per se, we are in agreement with Plaintiffs that Defendants' argument is a red herring because the Establishment Clause forbids not just "teaching" religion, but any governmental action that endorses or has the primary purpose or effect of advancing religion. The constitutional violation in Epperson consisted not of teaching a religious concept but of forbidding the teaching of a secular one, evolution, for religious reasons. Epperson, 393 U.S. at 103. In addition, the violation in Santa Fe was school sponsorship of prayer at an extracurricular activity, 530 U.S. at 307-09, and the violation in Selman was embellishing students' biology textbooks with a warning sticker disclaiming evolution. 390 F. Supp. 2d at 1312.

A third important issue concerning the classroom presentation of the disclaimer is the "opt out" feature. Students who do not wish to be exposed to the disclaimer and students whose parents do not care to have them exposed it, must "opt out" to avoid the unwanted religious message. Dr. Alters testified that the "opt out" feature adds "novelty," thereby enhancing the importance of the disclaimer in the students' eyes.[8] (14:123-25 (Alters)). Moreover, the stark choice that exists between submitting to state-sponsored religious instruction and leaving the public school classroom presents a clear message to students "who are nonadherents that they are outsiders, not full members of the political community." Sante Fe, 530 U.S. at 309-10 (quotation marks omitted).

Accordingly, we find that the classroom presentation of the disclaimer, including school administrators making a special appearance in the science classrooms to deliver the statement, the complete prohibition on discussion or questioning ID, and the "opt out" feature all convey a strong message of religious endorsement.

An objective student is also presumed to know that the Dover School Board

---

[8] In fact, the "opt out" procedure, as will be detailed herein, is itself clumsy and thus noteworthy to students and their parents, as it involves the necessity for students to have a form signed by parents and returned to the classroom before the disclaimer is read. Despite the fact that if properly executed the "opt out" form would excuse a student from hearing the disclaimer, the need to review the form and have some minimal discussion at least between parent and child hardly obviates the impact of the disclaimer, whether heard or not in the classroom.

advocated for the curriculum change and disclaimer in expressly religious terms, that the proposed curriculum change prompted massive community debate over the Board's attempts to inject religious concepts into the science curriculum, and that the Board adopted the ID Policy in furtherance of an expressly religious agenda, as will be elaborated upon below. Additionally, the objective student is presumed to have information concerning the history of religious opposition to evolution and would recognize that the Board's ID Policy is in keeping with that tradition. Consider, for example, that the Supreme Court in Santa Fe stated it presumed that "every Santa Fe High School student understands clearly" that the school district's policy "is about prayer," and not student free speech rights as the school board had alleged, and the Supreme Court premised that presumption on the principle that "the history and ubiquity" of the graduation prayer practice "provides part of the context in which a reasonable observer evaluates whether a challenged governmental practice conveys a message of endorsement of religion." Santa Fe, 530 U.S. at 315; Allegheny, 492 U.S. at 630; see also Black Horse Pike, 84 F.3d at 1486.

Importantly, the historical context that the objective student is presumed to know consists of a factor that weighed heavily in the Supreme Court's decision to strike down the balanced-treatment law in Edwards, specifically that "[o]ut of many possible science subjects taught in the public schools, the legislature chose to affect the teaching of the one scientific theory that historically has been opposed by certain religious sects." 482 U.S. at 593. Moreover, the objective student is presumed to know that encouraging the teaching of evolution as a theory rather than as a fact is one of the latest strategies to dilute evolution instruction employed by anti-evolutionists with religious motivations. Selman, 390 F. Supp. 2d at 1308.

In summary, the disclaimer singles out the theory of evolution for special treatment, misrepresents its status in the scientific community, causes students to doubt its validity without scientific justification, presents students with a religious alternative masquerading as a scientific theory, directs them to consult a creationist text as though it were a science resource, and instructs students to forego scientific inquiry in the public school classroom and instead to seek out religious instruction elsewhere. Furthermore, as Drs. Alters and Miller testified, introducing ID necessarily invites religion into the science classroom as it sets up what will be perceived by students as a "God-friendly" science, the one that explicitly mentions an intelligent designer, and that the "other science," evolution, takes no position on religion. (14:144-45 (Alters)). Dr. Miller testified that a false duality is produced: It "tells students . . . quite explicitly, choose God on the side of intelligent design or choose atheism on the side of science." (2:54-55 (Miller)). Introducing such a religious conflict into the classroom is "very dangerous" because it forces students to "choose between God and science," not a choice that schools should be forcing on them. Id. at 55.

Our detailed chronology of what a reasonable, objective student is presumed to know has made abundantly clear to the Court that an objective student would view the disclaimer as a strong official endorsement of religion or a religious viewpoint. We now turn to whether an objective adult observer in the Dover community would perceive Defendants' conduct similarly.

3. Whether an Objective Dover Citizen Would Perceive Defendants' Conduct to be an Endorsement of Religion

The Court must consider whether an objective adult observer in the Dover community would perceive the challenged ID Policy as an endorsement of religion because the unrefuted evidence offered at trial establishes that although the disclaimer is read to students in their ninth grade biology classes, the Board made and subsequently defended its decision to implement the curriculum change publicly, thus casting the entire community as the "listening audience" for its religious message. Santa Fe, 530 U.S. at 308. We are in agreement with Plaintiffs that when a governmental practice bearing on religion occurs within view of the entire community, the reasonable observer is an objective, informed adult within the community at large, even if the specific practice is directed at only a subset of

The New York Times
nytimes.com

October 31, 1992

After 350 Years, Vatican Says Galileo Was Right: It Moves

By ALAN COWELL,

More than 350 years after the Roman Catholic Church condemned Galileo, Pope John Paul II is poised to rectify one of the Church's most infamous wrongs -- the persecution of the Italian astronomer and physicist for proving the Earth moves around the Sun.

With a formal statement at the Pontifical Academy of Sciences on Saturday, Vatican officials said the Pope will formally close a 13-year investigation into the Church's condemnation of Galileo in 1633. The condemnation, which forced the astronomer and physicist to recant his discoveries, led to Galileo's house arrest for eight years before his death in 1642 at the age of 77.

The dispute between the Church and Galileo has long stood as one of history's great emblems of conflict between reason and dogma, science and faith. The Vatican's formal acknowledgement of an error, moreover, is a rarity in an institution built over centuries on the belief that the Church is the final arbiter in matters of faith.

At the time of his condemnation, Galileo had won fame and the patronage of leading Italian powers like the Medicis and Barberinis for discoveries he had made with the astronomical telescope he had built. But when his observations led him to proof of the Copernican theory of the solar system, in which the sun and not the earth is the center, and which the Church regarded as heresy, Galileo was summoned to Rome by the Inquisition. Forced to Recant

By the end of his trial, Galileo was forced to recant his own scientific findings as "abjured, cursed and detested," a renunciation that caused him great personal anguish but which saved him from being burned at the stake.

Since then, the Church has taken various steps to reverse its opposition to Galileo's conclusions. In 1757, Galileo's "Dialogue Concerning the Two Chief World Systems" was removed from the Index, a former list of publications banned by the Church. When the latest investigation, conducted by a panel of scientists, theologians and historians, made a preliminary report in 1984, it said that Galileo had been wrongfully condemned. More recently, Pope John Paul II himself has said that the scientist was "imprudently opposed."

"We today know that Galileo was right in adopting the Copernican astronomical theory," Paul Cardinal Poupard, the head of the current investigation, said in an interview published this week.

This theory had been presented in a book published in 1543 by the Polish scientist Nicolaus Copernicus in opposition to the prevailing theory, advanced by the second-century astronomer Ptolemy, that the Sun and the rest of the cosmos orbited the Earth. But the contest between the two models was purely on theoretic and theological grounds until Galileo made the first observations of the four largest moons of Jupiter, exploding the Ptolemaic notion that all heavenly bodies must orbit the Earth.

In 1616, the Copernican view was declared heretical because it refuted a strict biblical interpreation of the Creation that "God fixed the Earth upon its foundation, not to be moved forever." But Galileo obtained the permission of Pope Urban VIII, a Barberini and a friend, to continue research into both the Ptolemaic and the Copernican views of the world, provided that his findings drew no definitive conclusions and acknowledged divine omnipotence.

But when, in 1632, Galileo published his findings in "Dialogue Concerning the Two Chief World Systems," the work was a compelling endorsement of the Copernican system.

Summoned to Rome for trial by the Inquisition one year later, Galileo defended himself by saying that scientific research and the Christian faith were not mutually exclusive and that study of the natural world would promote understanding and interpretation of the scriptures. But his views were judged "false and erroneous." Aging, ailing and threatened with torture by the Inquisition, Galileo recanted on April 30, 1633.

Because of his advanced years, he was permitted house arrest in Siena. Legend has it that as Galileo rose from kneeling before his inquisitors, he murmured, "e pur, si muove" -- "even so, it does move."

Empowering people

Empowering people is a vision, a context and a process that supports Master students. The following two notions are central to this vision:

• The Extreme User
• Transformation

Each Master's project has to be very explicit about the extreme user for which a design is being made. In the Empowering People vision, it should always be possible to see things as an opportunity, an invitation to begin to understand a very special lifestyle, a special power, or as a situation in which the extreme user is ahead of the crowd or is in a phase in which others are just not there yet.

Transformation refers to the change in powers of the selected user that becomes possible with the help of the designed product, system, or service. Transformation means “personal growth”. We hope to begin to provide an alternative for the idea of experience (design).The Empowering People approach should trigger a quest for effective technological solutions and for real added-value design.

Intelligent Spaces

Through evolution, people’s way of life changes with the world around them. They respond and interact with curiosity, spontaneity, desire and expression, according to their situation and context. More expressive products stimulate our perception and movement, transforming the space and everything around us. It is an interactive system that couples people and objects together, within a special place for a more holistic experience.

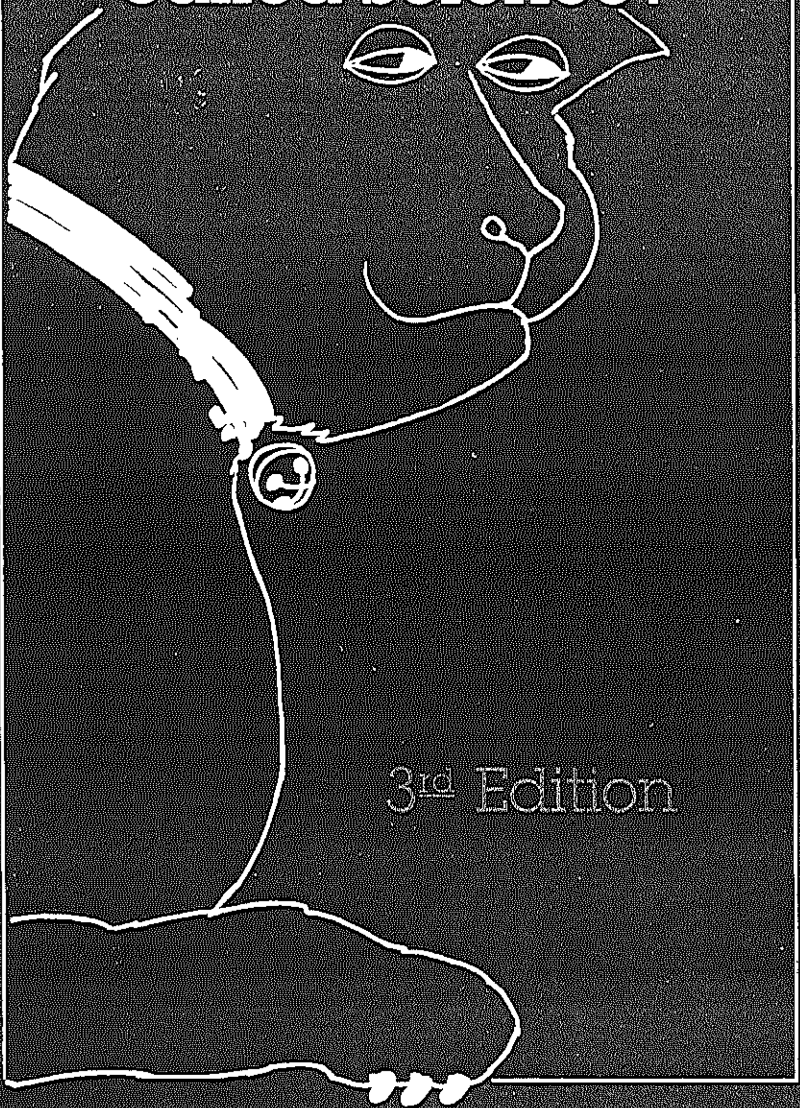
Intelligent Spaces develop new concepts, for a certain ‘event and context’. It includes the design of intelligent products, systems and services that can adapt to the user environment. The form and interaction determines the ‘activity space’ where products and experiences have a complete and connected relationship with the user.

"People change, so spaces need to adapt – new design forms for future experiences”

It is a new modernism for enabling technologies, observing everyday functional products and behaviors, then replacing them with a more ‘human-sense’ of movement and user interaction. Products have surface contours that react to proximity and touch, they have a sensorial feel and a modernist intelligence. People have a more stimulating experience within an intelligent space.

A. F Chalmers

What is this thing called Science?



3rd Edition

Introduction

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. But what, if anything, is so special about science? What is this "scientific method" that allegedly leads to especially meritorious or reliable results? This book is an attempt to elucidate and answer questions of that kind.

There is an abundance of evidence from everyday life that science is held in high regard, in spite of some disenchantment with science because of consequences for which some hold it responsible, such as hydrogen bombs and pollution. Advertisements frequently assert that a particular product has been scientifically shown to be whiter, more potent, more sexually appealing or in some way superior to rival products. This is intended to imply that the claims are particularly well-founded and perhaps beyond dispute. A recent newspaper advertisement advocating Christian Science was headed "Science speaks and says the Christian Bible is provedly true" and went on to tell us that "even the scientists themselves believe it these days". Here we have a direct appeal to the authority of science and scientists. We might well ask what the basis for such authority is. The high regard for science is not restricted to everyday life and the popular media. It is evident in the scholarly and academic world too. Many areas of study are now described as sciences by their supporters, presumably in an effort to imply that the methods used are as firmly based and as potentially fruitful as in a traditional science such as physics or biology. Political science and social science are by now commonplace. Many Marxists are keen to insist that historical materialism is a science. In addition, Library Science, Administrative Science, Speech

Science, Forest Science, Dairy Science, Meat and Animal Science and Mortuary Science have all made their appearance on university syllabuses.[1] The debate about the status of "creation science" is still active. It is noteworthy in this context that participants on both sides of the debate assume that there is some special category "science" . What they disagree about is whether creation science qualifies as a science or not.

Many in the so-called social or human sciences subscribe to a line of argument that runs roughly as follows. "The undoubted success of physics over the last three hundred years, it is assumed, is to be attributed to the application of a special method, 'the scientific method'. Therefore, if the social and human sciences are to emulate the success of physics then that is to be achieved by first understanding and formulating this method and then applying it to the social and human sciences." Two fundamental questions are raised by this line of argument, namely, "what is this scientific method that is alleged to be the key to the success of physics?" and "is it legitimate to transfer that method from physics and apply it elsewhere?".

All this highlights the fact that questions concerning the distinctiveness of scientific knowledge, as opposed to other kinds of knowledge, and the exact identification of the scientific method are seen as fundamentally important and consequential. As we shall see, however, answering these questions is by no means straightforward. A fair attempt to capture widespread intuitions about the answers to them is encapsulated, perhaps, in the idea that what is so special about science is that it is derived from the facts, rather than being based on personal opinion. This maybe captures the idea that, whereas personal opinions may differ over the relative merits of the novels of Charles Dickens and D. H. Lawrence, there is no room for such variation of opinions on the relative merits of Galileo's and Einstein's theories of relativity. It is the facts that are presumed to determine the superiority of Einstein's

innovations over previous views on relativity, and anyone who fails to appreciate this is simply wrong.

As we shall see, the idea that the distinctive feature of scientific knowledge is that it is derived from the facts of experience can only be sanctioned in a carefully and highly qualified form, if it is to be sanctioned at all. We will encounter reasons for doubting that facts acquired by observation and experiment are as straightforward and secure as has traditionally been assumed. We will also find that a strong case can be made for the claim that scientific knowledge can neither be conclusively proved nor conclusively disproved by reference to the facts, even if the availability of those facts is assumed. Some of the arguments to support this skepticism are based on an analysis of the nature of observation and on the nature of logical reasoning and its capabilities. Others stem from a close look at the history of science and contemporary scientific practice. It has been a feature of modern developments in theories of science and scientific method that increasing attention has been paid to the history of science. One of the embarrassing results of this for many philosophers of science is that those episodes in the history of science that are commonly regarded as most characteristic of major advances, whether they be the innovations of Galileo, Newton, Darwin or Einstein, do not match what standard philosophical accounts of science say they should be like.

One reaction to the realisation that scientific theories cannot be conclusively proved or disproved and that the reconstructions of philosophers bear little resemblance to what actually goes on in science is to give up altogether the idea that science is a rational activity operating according to some special method. It is a reaction somewhat like this that led the philosopher Paul Feyerabend (1975) to write a book with the title *Against Method: Outline of an Anarchistic Theory of Knowledge*. According to the most extreme view that has been read into Feyerabend's later writings, science has no special features that render it intrinsically superior to other kinds of knowledge such as ancient myths or voodoo. A high regard for science is seen as a modern religion, playing a similar role to that played by Christianity in Europe in earlier eras. It is suggested that the choices between scientific theories boils down to choices determined by the subjective values and wishes of individuals.

Feyerabend's skepticism about attempts to rationalise science are shared by more recent authors writing from a sociological or so-called "postmodernist" perspective.

This kind of response to the difficulties with traditional accounts of science and scientific method is resisted in this book. An attempt is made to accept what is valid in the challenges by Feyerabend and many others, but yet to give an account of science that captures its distinctive and special features in a way that can answer those challenges.

CHAPTER 1

Science as knowledge derived from the facts of experience

A widely held commonsense view of science

In the Introduction I ventured the suggestion that a popular conception of the distinctive feature of scientific knowledge is captured by the slogan "science is derived from the facts". In the first four chapters of this book this view is subjected to a critical scrutiny. We will find that much of what is typically taken to be implied by the slogan cannot be defended. Nevertheless, we will find that the slogan is not entirely misguided and I will attempt to formulate a defensible version of it.

When it is claimed that science is special because it is based on the facts, the facts are presumed to be claims about the world that can be directly established by a careful, unprejudiced use of the senses. Science is to be based on what we can see, hear and touch rather than on personal opinions or speculative imaginings. If observation of the world is carried out in a careful, unprejudiced way then the facts established in this way will constitute a secure, objective basis for science. If, further, the reasoning that takes us from this factual basis to the laws and theories that constitute scientific knowledge is sound, then the resulting knowledge can itself be taken to be securely established and objective.

The above remarks are the bare bones of a familiar story that is reflected in a wide range of literature about science. "Science is a structure built upon facts" writes J. J. Davies (1968, p. 8) in his book on the scientific method, a theme elaborated on by H. D. Anthony (1948, p. 145):

> It was not so much the observations and experiments which Galileo made that caused the break with tradition as his *attitude* to them. For him, the facts based on them were taken as facts, and not related to some preconceived idea ... The facts of

observation might, or might not, fit into an acknowledged scheme of the universe, but the important thing, in Galileo's opinion, was to accept the facts and build the theory to fit them.

Anthony here not only gives clear expression to the view that scientific knowledge is based on the facts established by observation and experiment, but also gives a historical twist to the idea, and he is by no means alone in this. An influential claim is that, as a matter of historical fact, modern science was born in the early seventeenth century when the strategy of taking the facts of observation seriously as the basis for science was first seriously adopted. It is held by those who embrace and exploit this story about the birth of science that prior to the seventeenth century the observable facts were not taken seriously as the foundation for knowledge. Rather, so the familiar story goes, knowledge was based largely on authority, especially the authority of the philosopher Aristotle and the authority of the Bible. It was only when this authority was challenged by an appeal to experience, by pioneers of the new science such as Galileo, that modern science became possible. The following account of the oft-told story of Galileo and the Leaning Tower of Pisa, taken from Rowbotham (1918, pp. 27–9), nicely captures the idea.

> Galileo's first trial of strength with the university professors was connected with his researches into the laws of motion as illustrated by falling bodies. It was an accepted axiom of Aristotle that the speed of falling bodies was regulated by their respective weights: thus, a stone weighing two pounds would fall twice as quick as one weighing only a single pound and so on. No one seems to have questioned the correctness of this rule, until Galileo gave it his denial. He declared that weight had nothing to do with the matter, and that ... two bodies of unequal weight ... would reach the ground at the same moment. As Galileo's statement was flouted by the body of professors, he determined to put it to a public test. So he invited the whole University to witness the experiment which he was about to perform from the leaning tower. On the morning of the day fixed, Galileo, in the presence of the assembled University and townsfolk, mounted to the top of the tower, carrying with him two balls, one weighing one hundred pounds and the other weighing one pound. Balancing the balls carefully on the edge of the parapet, he rolled them over together; they were seen to fall evenly, and the next instant, with a load clang, they struck the ground together. The old tradition was false, and modern science, in the person of the young discoverer, had vindicated her position.

Two schools of thought that involve attempts to formalise what I have called a common view of science, that scientific knowledge is derived from the fact, are the empiricists and the positivists. The British empiricists of the seventeenth and eighteenth centuries, notably John Locke, George Berkeley and David Hume, held that all knowledge should be derived from ideas implanted in the mind by way of sense perception. The positivists had a somewhat broader and less psychologically orientated view of what facts amount to, but shared the view of the empiricists that knowledge should be derived from the facts of experience. The logical positivists, a school of philosophy that originated in Vienna in the 1920s, took up the positivism that had been introduced by Auguste Comte in the nineteenth century and attempted to formalise it, paying close attention to the logical form of the relationship between scientific knowledge and the facts. Empiricism and positivism share the common view that scientific knowledge should in some way be derived from the facts arrived at by observation.

There are two rather distinct issues involved in the claim that science is derived from the facts. One concerns the nature of these "facts" and how scientists are meant to have access to them. The second concerns how the laws and theories that constitute our knowledge are derived from the facts once they have been obtained. We will investigate these two issues in turn, devoting this and the next two chapters to a discussion of the nature of the facts on which science is alleged to be based and chapter 4 to the question of how scientific knowledge might be thought to be derived from them.

Three components of the stand on the facts assumed to be the basis of science in the common view can be distinguished. They are:

(a) Facts are directly given to careful, unprejudiced observers via the senses.
(b) Facts are prior to and independent of theory.
(c) Facts constitute a firm and reliable foundation for scientific knowledge.

As we shall see, each of these claims is faced with difficulties and, at best, can only be accepted in a highly qualified form.

Seeing is believing

Partly because the sense of sight is the sense most extensively used to observe the world, and partly for convenience, I will restrict my discussion of observation to the realm of seeing. In most cases, it will not be difficult to see how the argument presented could be re-cast so as to be applicable to the other senses. A simple account of seeing might run as follows. Humans see using their eyes. The most important components of the human eye are a lens and a retina, the latter acting as a screen on which images of objects external to the eye are formed by the lens. Rays of light from a viewed object pass from the object to the lens via the intervening medium. These rays are refracted by the material of the lens in such a way that they are brought to a focus on the retina, so forming an image of the object. Thus far, the functioning of the eye is analogous to that of a camera. A big difference is in the way the final image is recorded. Optic nerves pass from the retina to the central cortex of the brain. These carry information concerning the light striking the various regions of the retina. It is the recording of this information by the brain that constitutes the seeing of the object by the human observer. Of course, many details could be added to this simplified description, but the account offered captures the general idea.

Two points are strongly suggested by the forgoing account of observation through the sense of sight that are incorporated into the common or empiricist view of science. The first is that a human observer has more or less direct access to

knowledge of some facts about the world insofar as they are recorded by the brain in the act of seeing. The second is that two normal observers viewing the same object or scene from the same place will "see" the same thing. An identical combination of light rays will strike the eyes of each observer, will be focused on their normal retinas by their normal eye lenses and give rise to similar images. Similar information will then travel to the brain of each observer via their normal optic nerves, resulting in the two observers seeing the same thing. In subsequent sections we will see why this kind of picture is seriously misleading.

Visual experiences not determined solely by the object viewed

In its starkest form, the common view has it that facts about the external world are directly given to us through the sense of sight. All we need to do is confront the world before us and record what is there to be seen. I can establish that there is a lamp on my desk or that my pencil is yellow simply by noting what is before my eyes. Such a view can be backed up by a story about how the eye works , as we have seen. If this was all there was to it, then what is seen would be determined by the nature of what is looked at, and observers would always have the same visual experiences when confronting the same scene. However, there is plenty of evidence to indicate that this is simply not the case. Two normal observers viewing the same object from the same place under the same physical circumstances do not necessarily have identical visual experiences, even though the images on their respective retinas may be virtually identical. There is an important sense in which two observers need not "see" the same thing. As N. R. Hanson (1958) has put it, "there is more to seeing than meets the eyeball". Some simple examples will illustrate the point.

Most of us, when first looking at Figure 1, see the drawing of a staircase with the upper surface of the stairs visible. But this is not the only way in which it can be seen. It can without

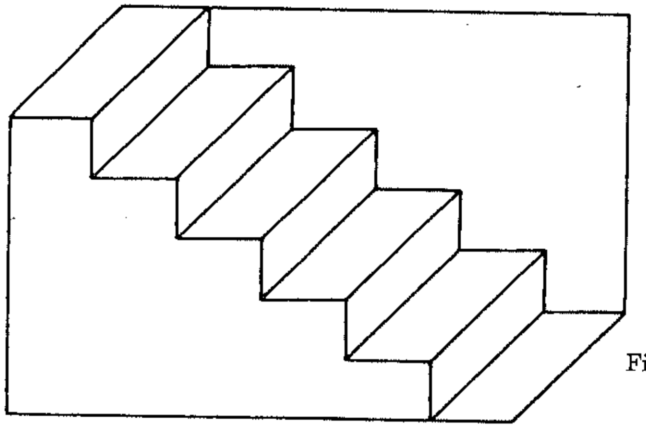


Figure 1

difficulty be seen as a staircase with the under surface of the stairs visible. Further, if one looks at the picture for some time, one generally finds that what one sees changes frequently, and involuntarily, from a staircase viewed from above to one viewed from below and back again. And yet it seems reasonable to suppose that, since it remains the same object viewed by the observer, the retinal images do not change. Whether the picture is seen as a staircase viewed from above or one viewed from below seems to depend on something other than the image on the retina of the viewer. I suspect that no reader of this book has questioned my claim that Figure 1 depicts a staircase. However, the results of experiments on members of African tribes whose culture does not include the custom of depicting three-dimensional objects by two-dimensional perspective drawings, nor staircases for that matter, indicate that members of those tribes would not see Figure 1 as a staircase at all. Again, it seems to follow that the perceptual experiences that individuals have in the act of seeing are not uniquely determined by the images on their retinas. Hanson (1958, chapter 1) contains some more captivating examples that illustrate this point.

Another instance is provided by a children's picture puzzle that involves finding the drawing of a human face among the

foliage in the drawing of a tree. Here, what is seen, that is, the subjective impressions experienced by a person viewing the drawing, at first corresponds to a tree, with trunk, branches and leaves. But this changes once the human face has been detected. What was once seen as branches and leaves is now seen as a human face. Again, the same physical object is viewed before and after the solution of the puzzle, and presumably the image on the observer's retina does not change at the moment the puzzle is solved and the face found. If the picture is viewed at some later time, the face is readily and quickly seen by an observer who has already solved the puzzle once. It would seem that there is a sense in which what an observer sees is affected by his or her past experience.

"What", it might well be suggested, "have these contrived examples got to do with science?" In response, it is not difficult to produce examples from the practice of science that illustrate the same point, namely, that what observers see, the subjective experiences that they undergo, when viewing an object or scene is not determined solely by the images on their retinas but depends also on the experience, knowledge and expectations of the observer. The point is implicit in the uncontroversial realisation that one has to learn to be a competent observer in science. Anyone who has been through the experience of having to learn to see through a microscope will need no convincing of this. When the beginner looks at a slide prepared by an instructor through a microscope it is rare that the appropriate cell structures can be discerned, even though the instructor has no difficulty discerning them when looking at the same slide through the same microscope. It is significant to note, in this context, that microscopists found no great difficulty observing cells divide in suitably prepared circumstances once they were alert for what to look for, whereas prior to this discovery these cell divisions went unobserved, although we now know they must have been there to be observed in many of the samples examined through a microscope. Michael Polanyi (1973, p. 101) describes the changes in a medical student's perceptual experience when he is taught to make a diagnosis by inspecting an X-ray picture.

> Think of a medical student attending a course in the X-ray diagnosis of pulmonary diseases. He watches, in a darkened room, shadowy traces on a fluorescent screen placed against a patient's chest, and hears the radiologist commenting to his assistants, in technical language, on the significant features of these shadows. At first, the student is completely puzzled. For he can see in the X-ray picture of a chest only the shadows of the heart and ribs, with a few spidery blotches between them, The experts seem to be romancing about figments of their imagination; he can see nothing that they are talking about. Then, as he goes on listening for a few weeks, looking carefully at ever-new pictures of different cases, a tentative understanding will dawn on him; he will gradually forget about the ribs and begin to see the lungs. And eventually, if he perseveres intelligently, a rich panorama of significant details will be revealed to him; of physiological variations and pathological changes, of scars, of chronic infections and signs of acute disease. He has entered a new world. He still sees only a fraction of what the experts can see, but the pictures are definitely making sense now and so do most of the comments made on them.

The experienced and skilled observer does not have perceptual experiences identical to those of the untrained novice when the two confront the same situation. This clashes with a literal understanding of the claim that perceptions are given in a straightforward way via the senses.

A common response to the claim that I am making about observation, supported by the kinds of examples I have utilised, is that observers viewing the same scene from the same place see the same thing but interpret what they see differently. I wish to dispute this. As far as perception is concerned, the only things with which an observer has direct and immediate contact are his or her experiences. These experiences are not uniquely given and unchanging but vary with the knowledge and expectations possessed by the observer. What is uniquely given by the physical situation, I am prepared to admit, is the image on the retina of an observer, but an

observer does not have direct perceptual contact with that image. When defenders of the common view assume that there is something unique given to us in perception that can be interpreted in various ways, they are assuming without argument, and in spite of much evidence to the contrary, that the images on our retinas uniquely determine our perceptual experiences. They are taking the camera analogy too far.

Having said all this, let me try to make clear what I do *not* mean to be claiming in this section, lest I be taken to be arguing for more than I intend to be. First, I am certainly not claiming that the physical causes of the images on our retinas have nothing to do with what we see. We cannot see just what we like. However, although the images on our retinas form part of the cause of what we see, another very important part of the cause is the inner state of our minds or brains, which will itself depend on our cultural upbringing, our knowledge and our expectations, and will not be determined solely by the physical properties of our eyes and the scene observed. Second, under a wide variety of circumstances, what we see in various situations remains fairly stable. The dependence of what we see on the state of our minds or brains is not so sensitive as to make communication, and science, impossible. Third, in all the examples quoted here, there is a sense in which all observers see the same thing. I accept and presuppose throughout this book that a single, unique, physical world exists independently of observers. Hence, when a number of observers look at a picture, a piece of apparatus, a microscope slide or whatever, there is a sense in which they are confronted by, look at, and hence see, the same thing. But it does not follow from this that they have identical perceptual experiences. There is a very important sense in which they do not see the same thing, and it is that latter sense on which I base some of my queries concerning the view that facts are unproblematically and directly given to observers through the senses. To what extent this undermines the view that facts adequate for science can be established by the senses remains to be seen.

Observable facts expressed as statements

In normal linguistic usage, the meaning of "fact" is ambiguous. It can refer to a statement that expresses the fact and it can also refer to the state of affairs referred to by such a statement. For example, it is a fact that there are mountains and craters on the moon. Here the fact can be taken as referring to the mountains or craters themselves. Alternatively, the statement "there are mountains and craters on the moon" can be taken as constituting the fact. When it is claimed that science is based on and derived from the facts, it is clearly the latter interpretation that is appropriate. Knowledge about the moon's surface is not based on and derived from mountains and craters but from factual statements about mountains and craters.

As well as distinguishing facts, understood as statements, from the states of affairs described by those statements, it is also clearly necessary to distinguish statements of facts from the perceptions that might occasion the acceptance of those statements as facts. For example, it is undoubtedly the case that when Darwin underwent his famous voyage on the *Beagle* he encountered many novel species of plant and animal, and so was subject to a range of novel perceptual experiences. However, he would have made no significant contribution to science had he left it at that. It was only when he had formulated statements describing the novelties and made them available to other scientists that he made a significant contribution to biology. To the extent that the voyage on the *Beagle* yielded novel facts from which an evolutionary theory could be derived, or to which an evolutionary theory could be related, it was statements that constituted those facts. For those who wish to claim that knowledge is derived from facts, they must have statements in mind, and neither perceptions nor objects like mountains and craters.

With this clarification behind us, let us return to the claims (a) to (c) about the nature of facts which concluded the first section of this chapter. Once we do so they immediately

become highly problematic as they stand. Given that the facts that might constitute a suitable basis for science must be in the form of statements, the claim that facts are given in a straightforward way via the senses begins to look quite misconceived. For even if we set aside the difficulties highlighted in the previous section, and assume that perceptions are straightforwardly given in the act of seeing, it is clearly not the case that statements describing observable states of affairs (I will call them observation statements) are given to observers via the senses. It is absurd to think that *statements* of fact enter the brain by way of the senses.

Before an observer can formulate and assent to an observation statement, he or she must be in possession of the appropriate conceptual framework and a knowledge of how to appropriately apply it. That this is so becomes clear when we contemplate the way in which a child learns to describe (that is, make factual statements about) the world. Think of a parent teaching a child to recognise and describe apples. The parent shows the child an apple, points to it, and utters the word "apple". The child soon learns to repeat the word "apple" in imitation. Having mastered this particular accomplishment, perhaps on a later day the child encounters its sibling's tennis ball, points and says "apple". At this point the parent intervenes to explain that the ball is not an apple, demonstrating, for example, that one cannot bite it like an apple. Further mistakes by the child, such as the identification of a choko as an apple, will require somewhat more elaborate explanations from the parent. By the time the child can successfully say there is an apple present when there is one, it has learnt quite a lot about apples. So it would seem that it is a mistake to presume that we must first observe the facts about apples before deriving knowledge about them from those facts, because the appropriate facts, formulated as statements, presuppose quite a lot of knowledge about apples.

Let us move from talk of children to some examples that are more relevant to our task of understanding science. Imagine a skilled botanist accompanied by someone like myself who is largely ignorant of botany taking part in a field trip into the Australian bush, with the objective of collecting observable facts about the native flora. It is undoubtedly the case that the botanist will be capable of collecting facts that are far more numerous and discerning than those I am able to observe and formulate, and the reason is clear. The botanist has a more elaborate conceptual scheme to exploit than myself, and that is because he or she knows more botany than I do. A knowledge of botany is a prerequisite for the formulation of the observation statements that might constitute its factual basis.

Thus, the recording of observable facts requires more than the reception of the stimuli, in the form of light rays, that impinge on the eye. It requires the knowledge of the appropriate conceptual scheme and how to apply it. In this sense, assumptions (a) and (b) cannot be accepted as they stand. Statements of fact are not determined in a straightforward way by sensual stimuli, and observation statements presuppose knowledge, so it cannot be the case that we first establish the facts and then derive our knowledge from them.

Why should facts precede theory?

I have taken as my starting point a rather extreme interpretation of the claim that science is derived from the facts. I have taken it to imply that the facts must be established prior to the derivation of scientific knowledge from them. First establish the facts and then build your theory to fit them. Both the fact that our perceptions depend to some extent on our prior knowledge and hence on our state of preparedness and our expectations (discussed earlier in the chapter) and the fact that observation statements presuppose the appropriate conceptual framework (discussed in the previous section) indicate that it is a demand that is impossible to live up to. Indeed, once it is subject to a close inspection it is a rather silly idea, so silly that I doubt if any serious philosopher of science would wish to defend it. How can we establish

significant facts about the world through observation if we do not have some guidance as to what kind of knowledge we are seeking or what problems we are trying to solve? In order to make observations that might make a significant contribution to botany, I need to know much botany to start with. What is more, the very idea that the adequacy of our scientific knowledge should be tested against the observable facts would make no sense if, in proper science, the relevant facts must always precede the knowledge that might be supported by them. Our search for relevant facts needs to be guided by our current state of knowledge, which tells us, for example, that measuring the ozone concentration at various locations in the atmosphere yields relevant facts, whereas measuring the average hair length of the youths in Sydney does not. So let us drop the demand that the acquisition of facts should come before the formulation of the laws and theories that constitute scientific knowledge, and see what we can salvage of the idea that science is based on the facts once we have done so.

According to our modified stand, we freely acknowledge that the formulation of observation statements presupposes significant knowledge, and that the search for relevant observable facts in science is guided by that knowledge. Neither acknowledgment necessarily undermines the claim that knowledge has a factual basis established by observation. Let us first take the point that the formulation of significant observation statements presupposes knowledge of the appropriate conceptual framework. Here we note that the availability of the conceptual resources for formulating observation statements is one thing. The truth or falsity of those statements is another. Looking at my solid state physics textbook, I can extract two observation statements, "the crystal structure of diamond has inversion symmetry" and "in a crystal of zinc sulphide there are four molecules per unit cell". A degree of knowledge about crystal structures and how they are characterised is necessary for the formulation and understanding of these statements. But even if you do not have that knowledge, you will be able to recognise that there are other, similar, statements that can be formulated using the same terms, statements such as "the crystal structure of diamond does not have inversion symmetry" and "the crystal of diamond has four molecules per unit cell". All of these statements are observation statements in the sense that once one has mastered the appropriate observational techniques their truth or falsity can be established by observation. When this is done, only the statements I extracted from my textbook are confirmed by observation, while the alternatives constructed from them are refuted. This illustrates the point that the fact that knowledge is necessary for the formulation of significant observation statements still leaves open the question of which of the statements so formulated are borne out by observation and which are not. Consequently, the idea that knowledge should be based on facts that are confirmed by observation is not undermined by the recognition that the formulation of the statements describing those facts are knowledge-dependent. There is only a problem if one sticks to the silly demand that the confirmation of facts relevant to some body of knowledge should precede the acquisition of any knowledge.

The idea that scientific knowledge should be based on facts established by observation need not be undermined, then, by the acknowledgment that the search for and formulation of those facts are knowledge-dependent. If the truth or falsity of observation statements can be established in a direct way by observation, then, irrespective of the way in which those statements came to be formulated, it would seem that the observation statements confirmed in this way provide us with a significant factual basis for scientific knowledge.

The fallibility of observation statements

We have made some headway in our search for a characterisation of the observational base of science, but we are not out of trouble yet. In the previous section our analysis presupposed that the truth or otherwise of observation statements

can be securely established by observation in an unproblematic way. But is such a presupposition legitimate? We have already seen ways in which problems can arise from the fact that different observers do not necessarily have the same perceptions when viewing the same scene, and this can lead to disagreements about what the observable states of affairs are. The significance of this point for science is borne out by well-documented cases in the history of science, such as the dispute about whether or not the effects of so-called N-rays are observable, described by Nye (1980), and the disagreement between Sydney and Cambridge astronomers over what the observable facts were in the early years of radio astronomy, as described by Edge and Mulkay (1976). We have as yet said little to show how a secure observational basis for science can be established in the face of such difficulties. Further difficulties concerning the reliability of the observational basis of science arise from some of the ways in which judgments about the adequacy of observation statements draw on presupposed knowledge in a way that renders those judgments fallible. I will illustrate this with examples.

Aristotle included fire among the four elements of which all terrestrial objects are made. The assumption that fire is a distinctive substance, albeit a very light one, persisted for hundreds of years, and it took modern chemistry to thoroughly undermine it. Those who worked with this presupposition considered themselves to be observing fire directly when watching flames rise into the air, so that for them "the fire ascended" is an observation statement that was frequently borne out by direct observation. We now reject such observation statements. The point is that if the knowledge that provides the categories we use to describe our observations is defective, the observation statements that presuppose those categories are similarly defective.

My second example concerns the realisation, established in the sixteenth and seventeenth centuries, that the earth moves, spinning on its axis and orbiting the sun. Prior to the circumstances that made this realisation possible, it can be said that the statement "the earth is stationary" was a fact confirmed by observation. After all, one cannot see or feel it move, and if we jump in the air, the earth does not spin away beneath us. We, from a modern perspective, know that the observation statement in question is false in spite of these appearances. We understand inertia, and know that if we are moving in a horizontal direction at over one hundred metres per second because the earth is spinning, there is no reason why that should change when we jump in the air. It takes a force to change speed, and, in our example, there are no horizontal forces acting. So we retain the horizontal speed we share with the earth's surface and land where we took off. "The earth is stationary" is not established by the observable evidence in the way it was once thought to be. But to fully appreciate why this is so, we need to understand inertia. That understanding was a seventeenth-century innovation. We have an example that illustrates a way in which the judgment of the truth or otherwise of an observation statement depends on the knowledge that forms the background against which the judgment is made. It would seem that the scientific revolution involved not just a progressive transformation of scientific theory, but also a transformation in what were considered to be the observable facts!

This last point is further illustrated by my third example. It concerns the sizes of the planets Venus and Mars as viewed from earth during the course of the year. It is a consequence of Copernicus's suggestion that the earth circulates the sun, in an orbit outside that of Venus and inside that of Mars, that the apparent size of both Venus and Mars should change appreciably during the course of the year. This is because when the earth is around the same side of the sun as one of those planets it is relatively close to it, whereas when it is on the opposite side of the sun to one of them it is relatively distant from it. When the matter is considered quantitatively, as it can be within Copernicus's own version of his theory, the effect is a sizeable one, with a predicted change in apparent diameter by a factor of about eight in the case of Mars and

about six in the case of Venus. However, when the planets are observed carefully with the naked eye, no change in size can be detected for Venus, and Mars changes in size by no more than a factor of two. So the observation statement "the apparent size of Venus does not change size during the course of the year" was straightforwardly confirmed, and was referred to in the Preface to Copernicus's *On the Revolutions of the Heavenly Spheres* as a fact confirmed "by all the experience of the ages" (Duncan, 1976, p. 22). Osiander, who was the author of the Preface in question, was so impressed by the clash between the consequences of the Copernican theory and our "observable fact" that he used it to argue that the Copernican theory should not be taken literally. We now know that the naked-eye observations of planetary sizes are deceptive, and that the eye is a very unreliable device for gauging the size of small light sources against a dark background. But it took Galileo to point this out and to show how the predicted change in size can be clearly discerned if Venus and Mars are viewed through a telescope. Here we have a clear example of the correction of a mistake about the observable facts made possible by improved knowledge and technology. In itself the example is unremarkable and non-mysterious. But it does show that any view to the effect that scientific knowledge is based on the facts acquired by observation must allow that the facts as well as the knowledge are fallible and subject to correction and that scientific knowledge and the facts on which it might be said to be based are interdependent.

The intuition that I intended to capture with my slogan "science is derived from the facts" was that scientific knowledge has a special status in part because it is founded on a secure basis, solid facts firmly established by observation. Some of the considerations of this chapter pose a threat to this comfortable view. One difficulty concerns the extent to which perceptions are influenced by the background and expectations of the observer, so that what appears to be an observable fact for one need not be for another. The second source of difficulty stems from the extent to which judgments about the truth of observation statements depend on what is already known or assumed, thus rendering the observable facts as fallible as the presuppositions underlying them. Both kinds of difficulty suggest that maybe the observable basis for science is not as straightforward and secure as is widely and traditionally supposed. In the next chapter I try to mitigate these fears to some extent by considering the nature of observation, especially as it is employed in science, in a more discerning way than has been involved in our discussion up until now.

Further reading

For a classic discussion of how knowledge is seen by an empiricist as derived from what is delivered to the mind via the senses, see Locke (1967), and by a logical positivist, see Ayer (1940). Hanfling (1981) is an introduction to logical positivism generally, including its account of the observational basis of science. A challenge to these views at the level of perception is Hanson (1958, chapter 1). Useful discussions of the whole issue are to be found in Brown (1977) and Barnes, Bloor and Henry (1996, chapters 1–3).

CHAPTER 2

Observation as practical intervention

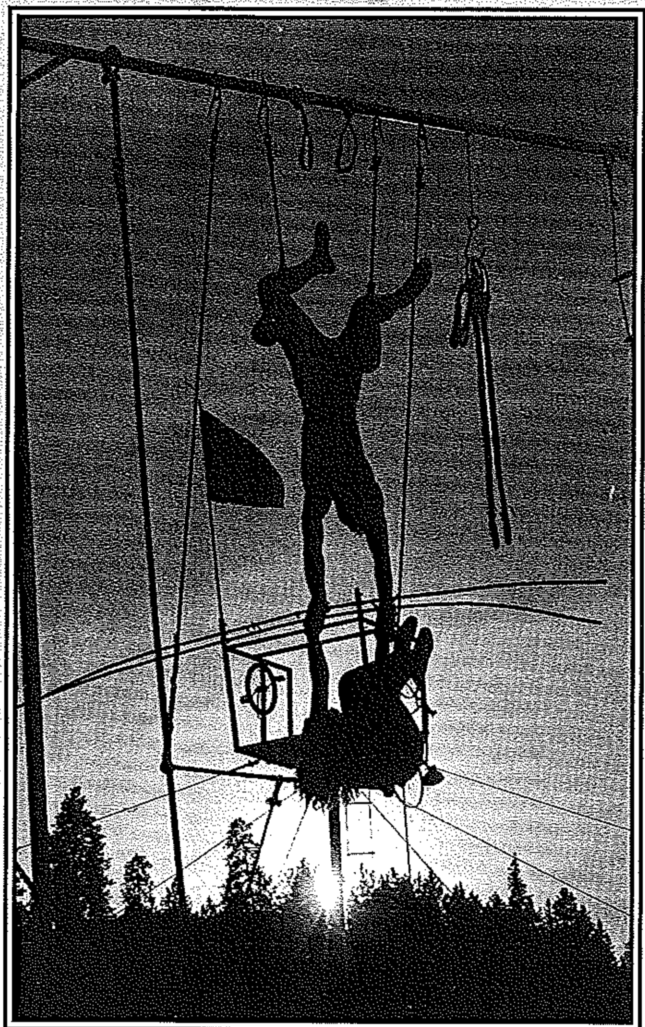
Observation: passive and private or active and public?

A common way in which observation is understood by a range of philosophers is to see it as a passive, private affair. It is passive insofar as it is presumed that when seeing, for example, we simply open and direct our eyes, let the information flow in, and record what is there to be seen. It is the perception itself in the mind or brain of the observer that is taken to directly validate the fact, which may be "there is a red tomato in front of me" for example. If it is understood in this way, then the establishment of observable facts is a very private affair. It is accomplished by the individual closely attending to what is presented to him or her in the act of perception. Since two observers do not have access to each other's perceptions, there is no way they can enter into a dialogue about the validity of the facts they are presumed to establish.

This view of perception or observation, as passive and private, is totally inadequate, and does not give an accurate account of perception in everyday life, let alone science. Everyday observation is far from passive. There are a range of things that are *done*, many of them automatically and perhaps unconsciously, to establish the validity of a perception. In the act of seeing we scan objects, move our heads to test for expected changes in the observed scene and so on. If we are not sure whether a scene viewed through a window is something out of the window or a reflection in the window, we can move our heads to check for the effect this has on the direction in which the scene is visible. It is a general point that if for any reason we doubt the validity of what seems to be the case on the basis of our perceptions, there are various actions we can take to remove the problem. If, in the example

SAGE *foundations of psychology*

Statistics *with* Confidence



Michael Smithson

assignment is due. If you find yourself getting lost, make a note of where you got lost and why before going to someone for help.

Something else that you can do to aid your learning is to make use of available support and resources. The electronic media accompanying this book provide helpful materials in addition to the questions and problems at the end of major sections and chapters. First, there are data-files corresponding to the appropriate problems and most of the main examples. These are in SPSS, Excel, and ASCII formats, so they should be readable by most statistics packages.

Second, there is a suite of tutorial modules collectively called *StatPatch* and *Demos*, and you will see references to them throughout this book. StatPatch is a mix of exploratory and problem-generating modules designed to build understanding and intuition in statistics. Demos is a collection of Excel workbooks, serving much the same purposes. One advantage that they have over textbook problems is that these modules generate infinitely many problems as well as providing immediate feedback, so you can learn at your own pace and have as much practice or exploration as you wish.

Finally, I maintain a website associated with this book. It may be found at http://psy.anu.edu.au/staff/mike/Statbook/TOC.html. It has links to other sites and helpful resources for psychology students studying statistics and research methods. While I would have liked to include web addresses in the book, I decided not to, mainly because many addresses change fairly often and new resources appear all the time. You are more likely to obtain the latest sites and correct addresses if I maintain them on the web.

Paths to knowledge or belief

Researchers are sometimes called 'knowledge workers.' This phrase suggests that their main contributive goals are adding to what we know or believe about the world, and correcting erroneous beliefs. Since many psychologists and other kinds of researchers claim to be using 'scientific methods' in their pursuit of knowledge, it is worthwhile briefly considering what they mean by this claim and how scientific methods are related to and distinguished from other ways of acquiring knowledge.

The following list is not exhaustive, nor is it the only possible list of different methods of knowledge acquisition. It is partly based on debates in the philosophy and sociology of science over whether science is an institution that is truly distinguishable from other institutions, and whether it has a defensible claim to superiority over those other institutions in getting us closer to truth or even reducing untruth. Fascinating as those debates are, we will not go into them here. Instead, the intention is to provide a rough guide to various paths to knowledge and belief, including scientific research:

- Personal (first-hand) experience
- Authority and/or consensus
- Intuition
- Common sense and tradition
- Rationalism and reasoning
- Scientific methods

DEFINITION

Personal experience encompasses events that we describe with phrases such as 'I saw it with my own eyes' or 'hands-on.' For many people, first-hand personal experience is synonymous with reality-testing. It is virtually impossible for us to see ourselves holding false beliefs here and now; the best we can do is to realize retrospectively that we once held a belief that we now consider false. This 'blind-spot' points towards one of the main drawbacks to reliance on personal experience, namely that without adequate precautions and comparisons with others' experiences, we may easily be led astray.

For one thing, personal experiences are necessarily very circumscribed and may not even comprise a representative sampling of the totality of experiences. Our experiences of blind people, avalanches, dugongs, and snowflakes encompass only a tiny and unrepresentative fraction of the blind people, avalanches, dugongs, and snowflakes to be found anywhere and for all time. Nevertheless, we sometimes overgeneralize on the basis of our experiences, as in making inferences to the entire population of dugongs from the only one we ever saw.

Worse still, we may be deluded or fall prey to illusions in our own experiences. You undoubtedly already know that your senses (sight, for instance) can be fooled by a magician or an optical illusion. We suffer from cognitive illusions as well, some of which we will become acquainted with in this book.

DEFINITION

Personal experience is nevertheless a crucial component of any scientific method, because scientific methods are grounded in empiricism. **Empirical methods** are those based on first-hand experiences of the world, so personal experience is a necessary component of those methods. **Empiricism** is a doctrine that ascribes superior truth-status to things that have been directly observed or manipulated over things that cannot be observed or manipulated. Most scientists are empiricists of one kind or another.

There are prescriptions in the scientific versions of personal experience that distinguish them from the usual versions. Most importantly, a scientist is supposed to adopt a stance of **impartiality** (or **disinterestedness**) towards all competing opinions or theories, including their own. That does not mean they cannot have values or pet ideas, although many writers confuse impartiality with the notion of being value-free (whereupon they rightly contend that no one is value-free and then wrongly conclude that scientists cannot adopt an impartial stance). It does mean that a scientist should take precautions in their research so that someone with different values and opinions could repeat their investigations and arrive at the same conclusions. An experiment set up so that the experimenter is 'blinded' with regard to which subjects have been assigned to which treatment condition is an example of such a precaution. Another example is designing a study expressly for investigating conditions under which the scientist's theory should fail if it is incorrect.

DEFINITION

In direct contrast with personal experience, using **authority or consensus** as a path to knowledge entails relying on second- or third-hand accounts of others' experiences. Authorities are sources with high status in our eyes. Parents, teachers, scientific experts, and religious leaders are examples of authorities. So are encyclopedias, scientific journals, television news programs, and computer programs. When every relevant authority agrees on a proposition (e.g., 'the world is round, not flat'), we have a *consensus* that makes that proposition appear indubitable. It is not difficult to see that the vast majority of what we think we know is based on appeals to authority and/or consensus.

Authorities can, of course, be wrong. A Dean of the Harvard Medical School was renowned for declaring to incoming first-year students that before they graduated they would have to commit some 40,000 'facts' to memory. Within 10 years of their graduation, about half of those 'facts' would be shown to be wrong. Unfortunately, he was fond of concluding, we never know which half. We have no way of knowing which of today's authoritatively established truths will become tomorrow's laughingstock. Moreover, the greater the authoritative consensus behind a belief, the less likely anyone will buck the tide to find out whether it is wrong after all.

DEFINITION

Scientific methods rely on appeals to authority, and agreement among relevant authorities is a legitimate goal in scientific work. All scientists are members of one or more scientific *communities* and none of them remain uninfluenced by those communities. Scientific communities have norms and institutions that many have argued make them less likely than other communities to fall prey to a misleading authoritative consensus. A **norm** means a usual or expected practice, rather like a custom. One of the most popular and also widely criticized lists of scientific norms is Robert K. Merton's (1973). I have added one more to his original four (Honesty).

- **Universalism**: Research and theory are to be judged on their own merits, regardless of the scientist's gender, ethnicity, creed, political affiliation, or any other characteristic. Blind peer review of research papers is an example of this norm in action, since the authors of the paper and the reviewers are unknown to each other.
- **Organized skepticism**: All ideas and evidence should be carefully scrutinized and subjected to skeptical inquiry. No results or conclusions should

be accepted other than provisionally, and even then subject to replication by other independent researchers.

- **Communalism**: Scientific knowledge should be shared freely with everyone. Proprietary secrecy is contrary to this norm. Where ethically possible, research practices, processes, data, and other 'raw' instruments or products should be publicly available for scrutiny.
- **Disinterestedness**: Alternative ideas are to be considered and tested on an equal footing with one's own, in such a way that someone with other views could repeat the tests or investigations and arrive at the same conclusion.
- **Honesty**: Cheating or dissembling is an especially strong taboo in scientific communities, so much so that an instance of it may result in banishment or ostracism.

This list of norms has provoked heated debate, both about whether scientists really adhere to them and whether they should. While there are plenty of counterexamples against each of these norms (e.g., instances of prejudice, discrimination, credulity, secrecy, or fraud among scientists), defenders of the scientific community point to the institutional practices that embody them and observe that to the extent that anyone adheres to those norms they are adopting a scientific outlook and attitude (cf. Grinnell, 1987).

Now let us turn to **intuition**. In one sense, having an intuitive understanding of or belief about something entails not being able to ascribe that understanding or belief to a legitimate basis. Another sense of this term refers to the sudden, blinding insight that seems to arrive from nowhere. Both of these meanings amount to tacit knowledge, knowing something without knowing how we know it. It is here, perhaps, that scientists most closely resemble everyone else. While some famous scientists have written popular accounts of having flashes of intuition and while many scientists prize good intuition as highly as the rest of us, they also happily confess that they don't know how it happens either! There is a widely held view among scientists that intuition alone is not sufficient to justify an idea, but that is not news to most people.

There is one respect in which scientists may diverge somewhat from popular views about intuition and common sense. They tend to be fascinated with research outcomes that fly in the face of intuition or common sense. It is possible that the fascination with counter-intuitive findings simply reflects a shrewd judgment that such findings are unlikely to have been discovered before and quite likely to advance one's scientific career, but there seems to be more to it than that.

Common sense and traditional truths are repositories of second- and third-hand knowledge loosely organized into theories and explanations of how the world works. Common sense is certainly a good place to begin but may be a bad place to end for scientific research. Moreover, psychology probably has one of the most difficult relationships of any discipline with common sense. The main reason for this is that most of us are pretty good common-sense psychologists, at least within our own cultures. Otherwise, we could not make our way through everyday life. In contrast, most of us are rather poor common-sense chemists and very poor common-sense subatomic particle physicists. Fortunately, we have little need to depend on our common sense in those areas. We can leave them to experts.

Psychological research often is accused of not going any farther than common sense while taking much longer to get there. There are two lines of defense against such accusations. One is that common sense contains mutually contradictory propositions that are not recognized as contradictory because people use them at different times. It is not difficult to think of opposing proverbs that demonstrate this, for instance:

- Look before you leap, vs. He who hesitates is lost.
- Opposites attract, vs. Birds of a feather flock together.
- Absence makes the heart grow fonder, vs. Out of sight, out of mind.
- Many hands make light the work, vs. Too many cooks spoil the broth.
- It's never too late to learn, vs. You can't teach an old dog new tricks.
- No one is an island, vs. We die alone.

Haslam & McGarty (1998) make amusing and instructive use of the third pair of proverbs to demonstrate how one might build up a research program to investigate which one is correct under various conditions. The other line of defense refers back to scientific norms of organized skepticism and disinterestedness. No matter how many people have endorsed a common-sensical assertion and no matter how long it has been believed, if it has not been properly tested then it is not scientific knowledge.

Finally, we turn to rationality and reasoning. **Rationality** involves adherence to a system of reasoning (usually standard logic). A popular view of science and, to a greater extent, mathematics, is that it relies heavily on logical reasoning and thereby rationality. While scientific research does make use of logic, logic is by no means sufficient on its own. Traditionally, rationality (along with rationalism) has been linked with knowledge and certainty. While ancient canons of rationality comprised substantive contents and told people what to believe, those versions were gradually supplanted by procedural and algorithmic prescriptions. Instead of directing people to specific conclusions, modern versions of rationality tell them how to reach conclusions. That is why most widely accepted versions of rationality boil down to some kind of logical consistency and coherency. DEFINITION

What is **rationalism**? It amounts to faith that rationality is the 'best' guide to decision making. Anything else (i.e., the nonrational, irrational, or anti- DEFINITION rational) is considered to be worse. Rationalists are anti-Heraclitans, which means they think there is sufficient regularity and stability in the universe for us to learn generalizable lawlike properties of it. They share this view with many empirical scientists (and much common-sense reasoning as well!). If we do not have a learnable world in some minimal sense, then rationality has no use. The usefulness of logical consistency assumes predictable, stationary relations among things in the real world. So does much scientific research. Some of the debates about whether psychology can or should be a science hinge on just this issue.

Where does statistics fit into all of this? Statistics is the offspring of a liaison between empiricism and rationalism. Statistical techniques are derived from general frameworks for understanding empirical data, so statistics and empirical research go hand-in-hand. Statistical models are based on theories of probability that, in turn, have some rationalistic and mathematical foundations. The marriage of empiricism and rationalism has not always been a peaceful one, and there are competing theories of statistics and probability. The versions we will use in this book are the most popular in psychology and work quite well under a wide range of conditions, but it is always wise to bear in mind that they are not the only approaches that could be used.

SUMMARY

The alternative **paths to knowledge** and belief reviewed in this section include:

- Personal (first-hand) experience
- Authority and/or consensus
- Intuition
- Common sense and tradition
- Rationalism and reasoning

Science makes use of all of these, albeit in ways that differ from their uses in everyday life.

Scientific communities have **norms** that many have argued make them less likely than other communities to fall prey to a misleading authoritative consensus:

- Universalism
- Organized skepticism
- Communalism
- Disinterestedness (or impartiality)
- Honesty

SUMMARY

Scientific methods are grounded in **empirical methods,** based on first-hand experiences of the world.

Empiricism *is a doctrine that ascribes superior truth-status to things that have been directly observed or manipulated over things that cannot be observed or manipulated.*

Rationality involves adherence to a system of reasoning (usually standard logic).

Rationalism is a faith that rationality is the best guide to decision making.

Statistics and probability are a combination of empiricist and rationalist ideas.

Uncertainties in research

The phrase 'psychological research' claims a large and diverse terrain, perhaps larger and more diverse than at any time in the history of the discipline. Pick 500 psychologists at random, ask them to describe how they do their research, and the answers will probably include experiments, surveys, case studies, in-depth interviews, ethnographies, test construction, discourse analysis, and computer simulations. You might encounter some who are doing historical or archival investigations, or even archeological studies.

It may seem as if there are no concepts or methods shared by all of these approaches to studying human beings. There are certainly practitioners of specific approaches who say that their approach has absolutely nothing in common with others. Nevertheless, there are good reasons to be suspicious of this sort of territorial statement and to think that researchers may share a few common bonds after all.

First, all researchers engage with the unknown in one sense or another. They begin by claiming that there really is something new under the sun and they are going to return from their voyaging to tell us something about it. Accordingly, they grapple with uncertainties, trade in novelties, map uncharted seas, and make discoveries. For all researchers, ignorance and uncertainty are both friend and foe, sometimes simultaneously. Without ignorance or uncertainty, there is nothing new to discover and the research game is over. In the grip of ignorance or uncertainty, however, the researcher is seldom in a position to demonstrate or prove anything conclusively. The physicist Richard Feynman captured this essential characteristic of scientific work in his 1955 address to the American National Academy of Science:

> The scientist has a lot of experience with ignorance and doubt and uncertainty, and this experience is of very great importance, I think. When a scientist doesn't know the answer to a problem, he is ignorant. When he has a hunch as to what the result is, he is uncertain. And when he is pretty

> darned sure of what the result is going to be, he is in some doubt. We have found it of paramount importance that in order to progress we must recognize the ignorance and leave room for doubt. Scientific knowledge is a body of statements of varying degrees of certainty – some most unsure, some nearly sure, none absolutely certain. (Feynman, 1988: 245)

Second, all researchers are members of one or more research communities. These are collections of people who agree sufficiently with one another to be able to share a conceptual framework, but whose discourse within that framework is characterized by vigorous argument, disputation, and conflict. Like ignorance, disagreement is both friend and foe to the researcher. Researchers crave consensus, but only up to a point. Complete agreement is a disaster for research, because no one is able to move outside the accepted way of thinking and there is nothing genuinely creative going on. Disagreement, while essential for motivating research, is often also agonizing for the researcher. Given the stylistic conventions of the time, a medical researcher turned philosopher of science, Ludwik Fleck, expressed this very well in 1935: 'At the moment of scientific genesis [discovery], the research worker personifies the totality of his physical and intellectual ancestors and of all his friends and enemies. They both promote and inhibit his search.' (Fleck, 1935/1979: 95.)

Third, all researchers make mistakes, both in their own eyes and the eyes of others. Here again, error is friend and foe. Anyone who gets it right the very first time really has learned nothing new. Making an error and realizing that it is an error are necessary components of any learning process, and therefore any process of discovery or creation as well. Again, Fleck is right on the mark: 'Discovery is thus inextricably interwoven with what is known as error. To recognize a certain relation, many another relation must be misunderstood, denied, or overlooked.' (Fleck, 1935/1979: 30.)

This does not mean, of course, that making any kind of mistake leads to discovery or learning. It *does* mean that reluctance to take a step for fear of making a misstep will surely impede discovery and learning. All researchers strive against indoctrinated fears of failure, error, and ridicule, much of it traceable to years of formal education that has rewarded them only for correct answers to problems whose solutions already are known. We can always dream of a system of education that does not penalize students for making mistakes! On a slightly more sober note, we can reward ourselves and others for risking productive and interesting mistakes, along with careful descriptions of them and our current states of ignorance. The University of Arizona's Medical Curriculum on Ignorance (Kerwin, 1993) is a salutary (and, alas, nearly solitary) example of a curriculum that invites students to describe and study not only knowledge but also what they don't know, and sustains ignorance as an object of study throughout their entire degree program.

This book is about uncertainty in psychological research and some widely shared methods for understanding and coping with it. It is also about how to make productive mistakes by taking strategic risks in designing and conducting research. That said, this book does not cover anything like the full gamut of research styles and techniques, the varieties of uncertainty, or their sources. That would require many books. None the less, there are some reasons behind the choices of research styles and uncertainties that inform this book's core.

To start with, we can place this book's focus in the context of various kinds of ignorance and uncertainty. The following list is adapted from Smithson (1989) and divides ignorance into two major chunks:

Types of ignorance and uncertainty

1. Distortion
   - Qualitative: Confusing one thing for another.
   - Quantitative: Systematic inaccuracy.
2. Incompleteness
   - Absence: Missing information.
   - Uncertainty: Indeterminate information.
     - Probability and statistical uncertainty: Likelihood of an event.
     - Ambiguity or vagueness: Multiple possible meanings or a range of values.

The first chunk, **distortion**, is usually taken to be some kind of systematic descriptive error. Its qualitative version consists of **confusion**, mistaking one thing for something else (as in a misdiagnosis), and its quantitative version consists of **inaccuracy** (as in a miscalibrated weighiug scale). DEFINITIONS

The second chunk, **incompleteness**, refers to information that is **absent** (missing) or **uncertain** (indeterminate). Indeterminacy of information is then subdivided into two categories: probabilistic and ambiguous or vague. **Ambiguity** and **vagueness** refer to ways in which information can be blurry, have multiple interpretations, or have shades and degrees of meaning. **Probability**, on the other hand, refers to the likelihood that something will happen.

Psychological research (indeed, perhaps all research) necessarily traffics in all of these kinds of ignorance and uncertainty. In psychology, problems of distortion are usually the province of measurement and ascribing meaning to our observations. We will introduce some of the basic concepts of measurement in Chapter 2, and discuss some issues concerning distortion in measurement there. Entire textbooks are devoted to measurement, however, and that is not the primary focus of this book (see Kaplan & Saccuzzo, 1989, on psychological testing and measurement, for example, or Foddy, 1993, on designing questions for surveys and experiments).

Incompleteness, on the other hand, is mainly the province of data-analytic techniques, particularly statistical techniques. Most of this book focuses on incompleteness, especially probabilistic or statistical uncertainty. Uncertainty is generally held to be more difficult to deal with than distortion, and less likely to be eliminated even from the best research. However, uncertainty can be described, sometimes quantified and estimated, and even manipulated in the service of the researcher. In this book we will encounter these three strategies for managing uncertainty many times.

Another viewpoint on uncertainty in research emerges once we distinguish among the sources of uncertainty that become salient during the research process. One way to understand this is to begin with a schematic guide to that process. Like almost any schematic, the one in Figure 1.1 is oversimplified. It begins with the researcher defining a topic and research goals, developing questions and/or hypotheses, and then going on to design the study, collect and analyze data, and finally interpreting the findings and revising what is known about the topic on the basis of those findings.

The feedback loops indicated in this figure are not the only possible feedback effects, and to some extent the revisability of the earlier stages depends on the kind of research being conducted and the norms of the research community involved. A rigorous experiment, for example, designed to test very specific hypotheses, leaves little room for the researcher to revamp those hypotheses in midstream. None the less, the feedback loops represent the fact that the research process may not be a one-way sequence but can involve successive iterations, whereby the researcher oscillates between stages until satisfied enough to move on.

DEFINITION

The six stages in this schematic also provide convenient labels for sources of uncertainty. **Topical uncertainty**, to begin with, concerns how the researcher is to describe the object of their investigations. Consider psychological research on affect or emotions. Are we studying emotional traits such as temperament, or more temporary states like moods, or even briefer episodes?

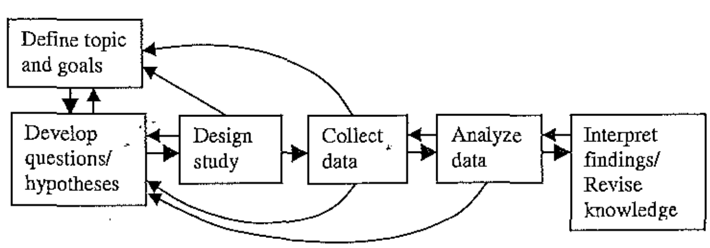


FIGURE 1.1 Research process

Are emotions best thought of as physiological, somatic, or socially based? What does the topic of emotions include and what does it exclude? Until we have at least tentative answers to questions such as these we cannot be sure about our topic.

DEFINITION

Investigative uncertainty concerns research questions and hypotheses. The kinds of questions that are sensible to ask about a topic are influenced to some extent by prior views and assumptions about the topic. For instance, if we assume that emotions are primarily products of physiological processes, then the most obvious starting-point would include aiming at an accurate description of the physiological concomitants of each distinct emotion. If, on the other hand, we assume that emotions are mainly interpretive constructs that arise from social interactions, then inquiring about physiological states and arousal would seem less relevant.

DEFINITION

Methodological uncertainty refers to the design of the study and whether it will suit our purposes, answer our questions, or test our hypotheses. This term is taken from Haslam & McGarty (1998, Ch. 11) but used somewhat more broadly here. They distinguish between two kinds of methodological uncertainty, whereas I will use three:

- Design uncertainty
- Internal uncertainty
- External uncertainty

DEFINITIONS

Design uncertainty concerns the overall method to be used in the study. Should we set up an experiment? A survey? What about a qualitative field study? **Internal uncertainty** refers to whether the researcher can interpret the outcomes of the study correctly, and **external uncertainty** refers to whether the study's results can be generalized to other populations or settings. These concepts may seem quite abstract now, but they will become clearer in the chapters to follow. For the present time, let's consider a few examples of research that has involved methodological uncertainty.

One of the chief concerns for anyone using more than one method to study something is that the methods might produce conflicting findings. Researchers must then sift through the studies and findings for possible explanations of the differences among them. For example, self-report or 'subjective' measures may differ from 'objective' indicators, as is often the case on topics such as risk or quality of life. A researcher faced with this situation will attempt to weave an explanation for it into a general account of risk perception, or health consequences of perceived versus objective quality of life. This explanation might refer to how people's perceptions and attitudes differ from so-called 'objective' measures, or to differences between the phenomena being studied, or even effects of studying the same phenomenon in different ways.

As an example of the latter case, in a study in which I was involved (Smithson *et al.*, 1991) we found that survivors of suicide attempts were much more likely to admit to consuming alcohol at the time of their attempt if we asked them directly about it than if we left them to tell about the event in their own words. Any reasonable attempt to explain this contrast would require a theory of how people reconstruct such memories and then edit them into narratives about themselves.

A much earlier example is Hovland's (1959) review of the experimental versus survey research evidence on attitude change subsequent to exposure to a change-inducing message. He found that surveys were less likely to find evidence of attitude change than experiments. He ascribed this divergence to a tendency of the surveys to attract people who already favored the view advocated by the message, the shorter time-intervals used in experimental studies, and differences in the kinds of attitudinal issues used.

Sometimes the issues involved in comparing different methods may simply be too complex or researchers may be too biased in their views for a resolution to be achieved. Consider, for instance, the contrast between one group of primarily quantitative studies of mental patients and another group of mainly qualitative studies reviewed by Weinstein (1979) whose review is discussed at length in Bryman's (1988) textbook. The first group used structured survey questionnaires with rating scales, multiple-choice questions, sentence-completion tests, and the like. For the most part, these studies reported that mental patients had favorable attitudes towards their institutions, benefited from and even enjoyed hospitalization. The qualitative studies, on the other hand, used unstructured interviews with patients who in turn answered questions in their own words, observation by researchers masquerading as patients, and observations on hospital wards. These studies consistently found evidence of debasement and oppression by hospital authorities; and patients' feelings of anxiety, boredom, powerlessness, and betrayal.

Weinstein's review provoked controversy over how best to account for these different findings. He found fault with the methodologies of the qualitative studies. Critics of his review responded that Weinstein had tried to force a comparison between two relatively noncomparable sets of studies. They pointed to differences between the two groups on the admission status of the patients, and whether patient experiences or patient attitudes and outcomes were the object of investigation.

As Cook & Campbell (1979: 66) observe, relying on just one method to stndy a phenomenon when we know little abont it lays the researcher open to accusations of 'mono-method bias.' The challenge in using more than one method is to do so in a strategic and even-handed fashion so that contrasting findings suggest new avenues for research and ways of integrating the findings. That way lies progress in any field.

DEFINITIONS

The major portion of this book is devoted to **statistical uncertainty**, which is uncertainty related to the analysis and, to some extent, interpretation of data. **Descriptive statistics** characterize the data themselves, and so **descriptive uncertainty** concerns those characterizations. Suppose you have taken an exam in a cognitive psychology class, and the instructor is about to distribute the exam results. Before doing so, she mentions that the class average score was 64%. How well would this describe each student's score, including your own? The less variability in the scores, the closer the average would be to describing individual students' scores.

DEFINITION

Inferential statistics, on the other hand, are used for drawing conclusions about populations or underlying processes from the sample of data at hand. **Inferential uncertainty**, therefore, arises when we are not sure what kinds of statistical inferences we can make from our data. For example, the instructor might wonder whether your class has scored higher than last year's class, whose average was only 59%. She would realize that the difference between class averages of 59% and 64% might occur simply by chance, and she would use inferential statistics to address that possibility.

DEFINITION

Finally, **interpretive uncertainty** arises when, despite having good data and sound statistics, we are still unable to decide between competing interpretations for what we have found. Suppose the instructor finds out that your class probably did perform better on the exam than last year's class. Is your class more intelligent? Did they work harder? Did she do a better job of teaching? Were the exams equivalent, or was this year's exam easier? These are plausible alternative explanations for her findings, and she would want to eliminate all but one of them if possible.

Given all of these different sources of uncertainty, research might appear to be a very daunting enterprise. Most research, however, does not deal with all of them simultaneously. In fact, researchers routinely distinguish one kind of research undertaking from another in terms of which of these uncertainties are being dealt with. The labels we will work with here are 'exploratory,' 'descriptive,' and 'explanatory' research (Neuman, 1997, also uses these).

DEFINITION

Exploratory studies deal primarily with topical and investigative uncertainties. If we do not know anything about a topic, if little or nothing has been written about it, then we need to refine our comprehension of it and develop questions that may be used to guide future research. Until topical uncertainty has been reduced to some extent, little progress on the other sources of uncertainty can be made. This is not to say that topical uncertainty must or can be completely eliminated. Some of the most interesting topics in psychology, creativity and consciousness being two examples, still are quite vague and sharply disputed even though they are the objects of long-running mature research programs.

DEFINITION

Descriptive studies focus mainly on investigative uncertainty, although the research may also end up dealing with methodological and descriptive statistical uncertainties. The goal of description is to provide an accurate portrayal of the phenomena that leads to further questions, hypotheses, and eventually explanations and theories. Developing a better way of measuring anxiety would be a good example of a descriptive study. Descriptions may be in either qualitative or quantitative form, and quite often the researcher will use this research to organize understanding of the phenomena.

As you might already have imagined, some studies can be both exploratory and descriptive. In the mid-1980s I worked with a former stomatherapist nurse, Therese Turner, who wanted to do her Honours research project on how colostomy patients managed their stigmatized condition in everyday life after their operations. Since a colostomy entails rerouting the colon so that it empties involuntarily into an external plastic bag instead of via the rectum, people who have had a colostomy are often at risk of public embarrassment. A search of the literature at the time revealed almost no relevant studies, so she elected to conduct a descriptive study based on in-depth interviews of former colostomy patients. She began by asking them what they thought were the major problems they faced and how they dealt with them. The matters raised by these people in the interviews generated further questions, and she returned to her informants for additional information. In the end, their accounts provided many fruitful suggestions for future research as well as advice that could be provided to such patients before and after the operation.

DEFINITION

Given a topic about which something is known and some descriptions of it, we tend to wonder why it is so. In **explanatory** research, the principal objects are reasons, causes, and interpretations. Explanatory studies therefore concentrate on statistical and interpretive uncertainties. They often test hypotheses or theories. We conduct such studies when we already have a good idea of the nature of our topic and what methods to use in studying it. Experiments are perhaps the best examples of explanatory research, because they require enough prior knowledge about a phenomenon to be able to manipulate some aspects of it in order to observe the effects that follow.

A number of concepts and terms have been introduced in this section, some of which may seem abstract and unfamiliar. If you can bear with it, these ideas will become clearer and form the basis for a genuine overview of psychological research that will stand you in good stead, not just for learning the material in this book but for understanding the diverse kinds of research throughout psychology.

SUMMARY

Research of any kind has three things in common that are both friend and foe:

1. Dealing with ignorance and uncertainty,
2. Disputation and conflict within a framework shared by other researchers, and
3. Learning and discovery through errors.

The kinds of uncertainty dealt with in research include the following:

1. **Distortion:** systematic error.
   - **Confusion**: Mistaking one thing for another.
   - **Inaccuracy**: Systematic miscalibration.
2. **Incompleteness**: Missing or indeterminate information.
   - **Absence**: Missing information.
   - **Uncertainty**: Indeterminate information.
   - **Probability** and statistical uncertainty: Likelihood of an event.
   - **Ambiguity or vagueness:** Multiple possible meanings or a range of values.

Sources of uncertainty in research arise at each of its six stages:

Topical uncertainty concerns how the researcher is to describe the object under investigation.

Investigative uncertainty concerns the nature of research questions and hypotheses.

Methodological uncertainty refers to the design of the study and whether it will answer our questions or test our hypotheses.

- **Design uncertainty** concerns the overall method to be used in the study.
- **Internal uncertainty** refers to whether the researcher can interpret the outcomes of the study correctly.
- **External uncertainty** refers to whether the study's results can be generalized to other populations or settings.

Statistical uncertainty concerns the analysis and, to some extent, interpretation of data.

- **Descriptive statistics** characterize the data themselves, and so **descriptive uncertainty** concerns those characterizations.
- **Inferential statistics** are used for drawing conclusions about populations or underlying processes from the sample of data. **Inferential uncertainty** concerns what kinds of statistical inferences we can make.

Interpretive uncertainty arises when the researcher is unable to decide between competing interpretations of the research findings.

Quantifying and counting

Since statistics are closely allied to quantification and counting, we should examine both of those practices before sailing off into areas where we take them for granted. We will start with counting, since that is the more venerable of the two and easier to conceptualize.

Counting assumes that the things being counted all belong to the same category. Its main advantage over using words is obvious once we grant that assumption. Saying that 'many' people in the class are right-handed is a nearly useless description compared to saying that 72 out of 83 are. Moreover, we can perform arithmetic operations with counts that are impossible with linguistic terms. Numbers and mathematics are not arbitrary social conventions. They have been successful because they help us think more clearly about certain things. To get a quick appreciation of this assertion, try multiplying 'three hundred and twenty-five by one hundred and twelve' versus 325 by 112, or try doing division with Roman numerals (if you are curious about other counting systems, take a look at Barrow, 1992).

Before counting behaviors, manifestations of cognitive processes, or the like, we need to be sure that they really do belong to the same category. For instance, consider the act of choosing the correct alternative on a true–false question in an exam. If we count the number of people who chose that alternative, we are lumping together those who knew the answer and those who happened to guess it. There may be no harm in counting how many got the question right, but we would be mistaken if we went on to say that was the number of people who knew the answer.

Quantification involves assigning numbers to distinguishable observations. While some concepts such as speed, duration, or length seem 'naturally' quantifiable, many psychological concepts provoke debates over whether they are quantifiable and if so, how best to quantify them. We will explore concepts in this book that inform those debates. When it is successful, quantification has much the same advantages as counting. It enables us to say not just that change or differentiation has taken place, but *how much* of it has occurred.

There are some popular arguments against quantification and counting, and we should examine them before moving on. One of the most pervasive is that 'reducing people to numbers' is anti-humanistic. It degrades people by ignoring their uniqueness as individuals. It is true that quantifying and counting require that we lump people together in some respects, thereby ignoring unique features. All general descriptions and theories do this. However, careful description and measurement never degrades anyone or anything. Also, words can just as easily and far more tellingly debase people by distorting or glossing over important characteristics.

Another related argument is that many important, observable things cannot be quantified. Characteristics that are not quantifiable tend to be ignored or discounted in favor of those that are quantified. There is some truth to this argument too, but the fault does not lie with quantification itself. After all, even qualitative characteristics ultimately must be codified, summarized, and counted once sufficiently many instances of them have been collected. Problems about quantification arise mainly when numbers become separated from their contexts. Good researchers know that every numerical datum has a context that needs to be considered before combining it with other numerical data, such as realizing when an EEG is showing 'artifact', that a rat went to sleep in the middle of a maze-running trial, or when a child is not paying attention in a reaction-time task.

A more general overview of the tradeoffs involved here might refer to 'data reduction' versus 'data enhancement.' Ragin (1994: 92) provides slightly different terms, and aptly observes that data reduction enables a researcher to see the big picture at the expense of attending to details, while data enhancement provides surrounding contextual information about the data that enables the researcher to better understand a particular case. **Data reduction** entails *combining* or *truncating* data. In order to perform either operation, we must treat the data as if they are combinable or comparable, or as if their qualitative differences are irrelevant. **Data enhancement** entails elaborating a set of data by *dissecting* it into components, or *supplementing* it with related data. Data enhancement involves an assumption that each datum is unique or distinctive in some relevant way that needs further explication.

DEFINITIONS

For instance, suppose we take weekly measurements of 100 people's levels of self-esteem using a well-established self-esteem index that consists of 15 questions, for a period of 11 weeks. An example of data reduction via combination would be averaging each person's self-esteem scores over the 11-week period. An example of reduction by truncation would be to rank their scores from lowest to highest, and use the middle (6th-ranked) score as their 'typical' score, ignoring all the other scores above and below it.

Likewise, an example of data enhancement via dissection would be breaking each score into the responses people gave on every one of the 15 questions in the self-esteem index. Data enhancement via supplementation, on the other hand, might consist of having people keep a diary of self-esteem-influencing events that would then be listed along with their score for the week (e.g., being reprimanded at work, or winning a ribbon in a local fun-run). Another kind of supplementation would be asking people to record their thoughts, feelings, and reasons for responding to the self-esteem questions, so that we have an elaboration of their accounts of the meanings behind their responses.

Many of the statistical techniques covered in this book have been designed to effectively condense or reduce data in various ways. In Chapter 3, for example, we will explore various kinds of summary statistics (such as the average, or arithmetic mean) that reduce a collection of scores to a few pieces of information about the properties of those scores. Other techniques, mainly those concerned with various ways of displaying data in graphs or tables, involve data enhancement as well as reduction. Although Ragin (1994) is oversimplifying somewhat, there is some truth in his claim that quantitative research techniques are mostly data condensers.

The tradeoff between these two ways of treating data is fairly obvious. Reducing large volumes of data to a few pieces of information is grist for any scientist's mill because it is compatible with the scientific goal of generalizable explanations and theories. A general theory effectively tells us that we may treat numerous specific cases as if they are essentially identical. When appropriately and intelligently applied, data reduction can reveal hidden order, regularity, or relationships among data in a powerful and even elegant fashion.

On the other hand, reducing means combining or truncating information and therefore ignoring it. The researcher who condenses data thereby risks ignoring important details or distinctions among particular cases. If taken to extremes, data reduction techniques can 'reduce people to numbers' by omitting crucial information about where the numbers come from or what they mean. In a somewhat different sense of the word, Dennett (1994) coined the term 'greedy reductionism' to refer to excessive reductionistic ambitions. Researchers can sometimes get carried away by the power of their data-reducing techniques, especially since the advent of computers. So can consumers of research. A friend who is an analyst in a large government department concerned with health and safety has repeatedly told me that she is always under pressure from her superiors to 'boil it down to *one* number.'

Data enhancement provides ways of grounding information in context. Even quantitative data may require data enhancement in order to be properly understood. For example, consider the effect of an income increase of $100 per week on someone with a $150 per week income versus someone whose income is $10,000 per week. Or compare the student who has scored 70 on an exam with 70 on both the 'technical' and 'conceptual' components, with another student whose score of 70 is the average of 95 on the technical and 55 on the conceptual components.

The disadvantages and pitfalls of data enhancement are twofold. First, the researcher may become overwhelmed by elaboration and thereby unable to see the forest for the trees. This is simply the opposite side of the reductionism coin as outlined earlier.

Second, inappropriate or irrelevant contextual distinctions can mislead us into separating data that should be combined. This is a somewhat subtler point that the first one, but a simple example can illustrate it. Shoe-sizes are numbered according to different conventions in the U.S. than they are in Australia. I wear a size 11 shoe if I purchase it in the U.S. but only a size 9 if I buy it in Australia. A survey of shoe-sizes with samples from both Australia and the U.S. would therefore require that we record where the respondent's shoes were purchased. However, if the survey were restricted to Australia, then separating shoe-sizes by the state in which they were purchased would be irrelevant.

When should we choose data reduction or data enhancement? There is no simple answer. It depends on the researcher's goals and what is already known about the area. One strategy that is frequently used is to begin at one extreme (either reduction or enhancement) and then work back towards the other as far as is needed. To conclude this section, here is an example of a debate that is frequently found in psychology, namely whether a psychological concept should have more than one dimension or not. The concept in question is risk.

The editorial in a recent issue of the *Royal Statistical Society News* (October 1998) bemoaned the fact that people choose their risks in an 'irrational' way. According to the editor, people 'refuse to engage in activities which have known, but quite negligible, risks yet fearlessly participate in those whose dangers are orders of magnitude greater' (p. 1). He recounted the solution put forward by the past president of the RSS, which was that a 'Richter-type' scale of risk be constructed whereby people could compare known risks in a systematic way. A very similar proposal was made by the mathematician John Allen Paulos in his book, *Innumeracy* (1988). Such a scale would be an example of data reduction, siuce it would collapse all risk evaluation down to one dimension.

Eating a peanut-butter sandwich every day for one month, working in a coalmine for a few hours, and living next to a nuclear power plant for five years all involve an increase in risk of death of about one in a million, so if we were weighing risks on just that basis, we should equally value these three. But most of us do not. A large empirical research literature demonstrates that people perceive and evaluate risk along several dimensions. The list below contains the influences on risk prefereuce identified in studies of risk perception (Otway & von Winterfeldt, 1982):

1. Involuntariness of exposure to the risk.
2. Lack of personal control over outcomes.
3. Uncertainty about probabilities or consequences of exposure.
4. Lack of experience or familiarity with the risk.

5. Difficulty in imagining consequences.
6. Delayed effects.
7. Genetic effects.
8. Catastrophic size of consequences (either geographically or numbers of people affected).
9. Benefits are not visible.
10. Benefits go to others but not oneself.
11. Human-caused rather than naturally caused.

The Richter-type risk scale does not provide a valid way of characterizing people's valuations of risk. It does, however, provide a worthwhile benchmark against which to compare how people do evaluate risks, because it orders risks along the continuum that we would use if probability of injury or death were our only concern in risk assessment. Given a person who has accurate information about such probabilities, we may use the Richter-type scale as a way of determining whether they are evaluating risks solely on the basis of probabilities. If we find that their preferences for risks disagree with the rank-ordering of those risks on the scale, then we know that we need to take more than just probabilities into account when attempting to describe people's risk preferences.

The moral to this example is that data reduction and enhancement can work together in getting a start in an unknown area. The field of risk assessment began with attempts to 'reduce' risk perception to a one-dimensional scale and the failure of that simple model stimulated the search for additional dimensions. This is an example of a pattern commonly encountered in scientific research, namely beginning with a simple, reduced model and then complicating it as necessary to fit the phenomena. The reverse process also can be found in some areas, whereby researchers begin with elaborate data enhancement and then systematically eliminate unnecessary features until they arrive at a more parsimonious model.

SUMMARY

Quantifying or even counting should not be undertaken without first considering whether these are sensible given possible arguments to the contrary.

Counting assumes that the things being counted all belong to the same category. *Quantification* involves assigning numbers to distinguishable states.

- **Data reduction** entails *combining* or *truncating* data.
- **Data enhancement** entails elaborating a set of data by *dissecting* it, or *supplementing* it with related data.

Two common strategies in research are to begin with data reduction and then enhance as much as necessary, or to start with data enhancement and then reduce as much as possible.

Questions and exercises

Q.1.1. Which are examples of scientific, rationalistic, intuitive, and authoritative methods of gaining or creating knowledge?

(a) Using a voltmeter to see whether your torch battery is flat.
(b) Your doctor says you have dermatitis, and you decide that you have dermatitis.
(c) Figuring that because only dogs make barking sounds in your neighborhood, the source of the barking sound outside your front door is a dog.
(d) Even though you haven't got a formal definition of creativity, you know it when you see it.
(e) Looking up the meaning of an English word in the *Oxford English Dictionary*.
(f) The curried chicken was too hot last week, so you try using one half-teaspoon less Madras powder this time.

Q.1.2. Give two examples of statistical descriptions.

Q.1.3. Give two examples of statistical inferences.

Q.1.4. Suppose a psychological researcher points out that everyone sees the same distinct bands in the rainbow regardless of how they classify or name colors, and uses this observation as an example of categorization that is independent of culture. Another psychologist argues that this phenomenon is not categorization at all. What kind of uncertainty is involved here?

Q.1.5. Give an argument for why happiness should be measured on a single scale, and an argument for why it should be measured on two separate scales (one 'positive' and the other 'negative').

Variables and Measurement

2

CONTENTS

Observational and measurement strategies

Why should we concoct systematic strategies for observing or measuring anything? There are at least four compelling reasons. First, all of us have only very limited first-hand knowledge of anything about the world, including human existence. The vast majority of what we think we know or believe is based solely on second- and third-hand accounts by authorities such as parents, teachers, and the media. Often, that is the best we can do. Nevertheless, without first-hand experience, second- and third-hand accounts require us to make assnmptions about their truth-status. Even our own experiences are sometimes of doubtful pedigree – all of us are potentially fallible observers and recorders, to say nothing of memorizers. Moreover, onr first-hand experiences are not just haphazardly constrained, but systematically truncated by social conventions, matters of interpersonal attraction, and political and other instrumental agendas. To gain a wider experience of what a representative sample of people thinks, feels, or does is no small undertaking and requires far more time, strategic work, and resources than most of us can bring to bear.

A second reason is that systematic measurement strategies and instruments are ways of extending our senses. We cannot directly see electrical activity in the brain when someone is listening to music, but we can measure that activity at least indirectly by using sensing and imaging techniques beyond our own capacities. Likewise, we cannot hang around a temple for 500 years and count the number of worshippers, but we can arrive at the scene 500 years after the temple was built and measure the wear on the steps at its entrance. Finally, while we do not have immediate access to any animal's preferences for one food over another, we may infer those preferences by watching which foods the animal selects when it is given a choice.

A third reason for strategic measurement refers to the demands on our time and attention. We cannot pay attention to everything, and we do not have much time. Nevertheless, we may make some careful choices about what to pay attention to, how to set about it, and how much time to devote to it. Depending on the researcher's goals and theoretical orientations, some kinds of observation and measurement are more relevant, important, or probative than others.

Suppose you are investigating the mental processes involved in reading, and you are debating with your colleagues about whether the capacity to recognize the meaning of a word operates independently of the capacity to recognize how it should sound. Then you should be very interested in finding people who have suffered head traumas that have left one capability intact but not the other, since that would unambiguously support the separate capacity theory. Unfortunately for opponents of that theory, head-trauma victims who have lost both capabilities do not provide unarguable support for their position. Likewise, you should be more interested in native readers of languages such as Chinese or Japanese, where the symbols representing meaning are sometimes separate from those representing sound, than readers of languages such as English, where the same symbols do both kinds of work.

A fourth reason for strategic observation and measurement is overcoming or guarding against biases and hidden assumptions. At one time some people claimed that scientific measurement could be freed of bias and uninfluenced by values. These days, some people have gyrated to the opposite extreme of claiming that all measurement is inherently biased and driven by researchers' values, ideological orientations, and preconceptions. Neither of these positions is true, and both invite intellectual laziness born of complacency in the first case and nihilistic relativism in the second. A more viable position that also gives us something to work with is that at least some biases can be identified and many of those can be overcome, even though we must bear in mind that there probably is no infallible method for doing so. Two kinds of perennial biases are those that direct our attention toward certain phenomena and away from others, and those that compel us to explain certain events but not others. One thing that makes these biases important is that they can entrap nearly everyone regardless of their ideological orientation or even their motives for doing their research.

We shall turn first to attentional biases. In 1986, the American space-shuttle 'Challenger' exploded shortly after take-off, killing all on board. A key cause of this tragedy was the failure of 'O-rings' that held the booster's fuel tank to the rest of the rocket. A *post-hoc* investigation revealed that these O-rings tended to fail when the temperature outside the rocket fell below a certain level. Why didn't the highly trained engineers who designed and tested the rocket figure this out beforehand? It turned out that they had considered the possibility that the O-rings might be sensitive to temperature. They had even examined the relationship between the number of O-ring failures and temperature for all previous space-shuttle flights when one or more O-rings had failed. What they had omitted to do was check the temperatures involved with flights where *no* O-rings failed. Had they done so, they would have seen that all those flights had temperatures above a critical level.

This is an example of our predilection for being much better at detecting the presence of something than its absence. We are greatly inclined to see objects and not the spaces around them – indeed, a famous basic exercise for beginning artists is to learn to see and draw those spaces – and we, along with other animals, learn much better from a cue linked to the presence of something than a cue linked to an absence. Thus, a common mistake in early medical and clinical psychological research was to focus exclusively on the clinical cases instead of also studying people who did not have the clinical condition.

Let us also consider a related bias that befalls even trained researchers. Suppose each of the cards in Figure 2.1 has a number on one side and a letter on the other, and someone tells you 'If a card has a vowel on one side then it has an even number on the other side.' Which of the cards *must* you turn over in order to decide whether the person is right? Try deciding which cards these would be before reading on.

Now imagine that you are a forensic criminal psychologist specializing in serial murderers, and your experiences in the field have inspired a hypothesis: 'All serial murderers kill domestic pets before turning to killing people.' If you were going to investigate whether this hypothesis is true, which of the following kinds of people would be most important to incorporate in your study? Try rank-ordering them from first to fourth most important, before reading on.

E	**K**	**4**	**7**

FIGURE 2.1 Card Task

- Serial murderers
- People who are not serial murderers
- People who have killed pets
- People who have not killed pets

If you are like most people, your response in the card task was to choose to turn over E and 4. Relatively few people give the logically correct answer, which is E and 7. The reason this answer is correct is that turning over the E card will disconfirm the rule if you find an odd number on its other side, and turning over the 7 will disconfirm the rule if you find a vowel on its other side. The other two cards are not relevant, since the rule states 'if vowel then even number' (and not, for instance, 'if even number then vowel'). The card task was first studied by Wason & Johnson-Laird in 1972 and is a classic example of the **confirmation bias**, namely the tendency for people to attend to instances that confirm their hypotheses or beliefs and ignore instances that might disconfirm them. Likewise, to rank pet-killers as more important to study than people who have not killed pets would be to fall prey to this bias. DEFINITION

Now let's briefly consider explanatory biases. Perhaps the most widespread is what might be called the **expectancy bias**, whereby we feel compelled to explain the unexpected but not the expected. Events that follow what we believe is the 'natural' or 'typical' order do not seem to require explanation, but 'unnatural' or 'atypical' ones do. Thus, we have seen many more studies of the causes of homosexuality than of heterosexuality, and one of the reasons is that a large and influential sector of Western societies perceives homosexuality to be 'unnatural'. Furthermore, we tend to think that atypical phenomena require special explanations over and above those for typical phenomena. Thus, from time to time people have proposed special theories of criminal behavior or of psychological disorders that are largely separate from theories of ordinary behavior or nonpathological psychology. The seductive nature of explanatory biases owes a great deal to the fact that if any phenomenon is merely what we expected it to be or what seems natural to us, then that is because we already have a theory about that phenomenon. However comfortable with our theory we might be, we should be prepared to subject it to as much rigorous testing as any other, and so we should be willing to investigate the ordinary, mundane, or expected along with their extraordinary counterparts. DEFINITION

Basic concepts of measurement

Whenever we measure anything, it is reasonable to distinguish between the record we make of this activity and the thing we are measuring. The record incorporates our measurements and any other relevant information; the thing

DEFINITION

we are measuring often is called a 'construct.' A **construct** may be an abstract concept such as verbal latency, information search strategy, anxiety, or appetite; but it may also be something that is quite tangible, such as reaction time, the trajectory of eye movement, galvanic skin response, or the amount of food consumed in a day. Some constructs, moreover, may serve as **indicators** of other constructs so that by knowing about one construct we obtain information about another. Thus, we might use reaction time as an indicator of latency, eye movement to indicate something about information search strategy, galvanic skin response to indicate nervousness, or daily food consumption to indicate appetite. Indicators that involve recording something about the real world are often called **operationalizations** of their respective constructs, and if we are convinced that one or more indicators truly represent everything

DEFINITION

about a construct, then we claim that they **measure** that construct. These two terms are important because they involve strong claims on the part of a researcher. Anyone who says they have *operationalized* a construct is claiming that they have a way of observing that construct's real-world manifestations, and if they further say they can *measure* the construct then they are claiming access to the totality of its real-world manifestations.

DEFINITION

A **variable** is an operationalization of a construct that can take on different values or states for different people (or even for the same person on different occasions). What is or is not a variable depends on the population being studied. If we are studying a group of 10-year-olds' reading ability, age is not a variable in that study. But if we are studying reading ability in children from an entire primary school, then age is a variable. Moreover, we must specify the conditions under which the construct can vary. Some constructs may change over time, or differ across people, or across situations for the same person.

As with most other aspects of research, deciding what will be treated as a variable is a matter of judgment and may be controversial. Gender, for example, is not a variable for most people during their lives, but it can be (transsexuals are those whose gender has varied at least once). In most, but not all, research on people it is sensible to treat gender as varying across people but not over time for the same person. However, a current issue in the psychology of the self is the extent to which an individual's personality or self-concept can vary throughout one's lifetime or even fleetingly from one situation to another.

All measurement strategies require the researcher to take at least some theoretical stances and risks, because in the absence of any theory the researcher has no idea what can or should be measured, let alone how to set about it. If we think that personality is fully formed and constant throughout adulthood, then the idea of measuring personality traits makes good sense. On the other hand, if we think that personality changes from one situation to another, the most we can hope for is to measure personality states at various points in time.

Even constructs themselves may be altered or redefined as theoretical frameworks develop. For instance, definitions of 'mental disorder' changed sufficiently throughout the 1970s that homosexuality was eliminated from the list of disorders in the more recent editions of the *Diagnostic and Statistical Manual* of the American Psychiatric Association.

It should not be surprising, then, that psychologists can differ dramatically on what is possible or appropriate to measure, and what the same measurements signify. Radical behaviorists, for instance, claim that we have no access to anyone's internal mental states or processes, so it is impossible to measure any such thing as a belief, attitude, or emotion. Their focus is exclusively on behavior and its causes, and unlike cognitive psychologists they do not infer cognitive states from behaviors. Likewise, where cognitively oriented researchers argue over whether various paper-and-pencil tests reveal people's underlying beliefs or attitudes.

The main point here is not simply that there is no such thing as theory-free measurement, although that is a popular view among psychologists and social scientists. Almost any measurement can be interpreted via many theories regardless of its origins, so that the same measurement may be given different meanings or evidential status by researchers using different theoretical perspectives. In fact, many controversies in psychology really are propelled by disputes over what certain data or measurements mean and imply. When conducting research or reading others', *it is important not only to understand* the researcher's theoretical standpoint on what is being measured, how, and why; but also how other theoretical perspectives would answer those questions and what uses they would make of the same measurements.

SUMMARY

The **confirmation bias** is a tendency to attend to instances that confirm prior hypotheses or beliefs and ignore instances that might disconfirm them.
The **expectancy bias** *is the inclination to* explain the unexpected but not the expected, or the unnatural but not the natural.

A **construct** is a concept, usually a characteristic or property, that underlies measurement.
A construct is an **indicator** of another construct if knowing about one provides information about the other.
Indicators are **operationalizations** of their respective constructs if they give us access to some aspect of their real-world manifestation.
If an operationalization captures the totality of a construct's real-world manifestation, then we claim that it **measures** that construct.
A **variable** is an operationalization of a construct that can change.

Measurement validity and error

Nearly every psychological perspective or theory incorporates stipulations of what are valid ways of measuring psychological phenomena and what constitutes measurement error. **Measurement error** occurs whenever measurements are influenced by something other than what the researcher intends to measure. **Validity**, on the other hand, is a general term denoting the extent to which measurement is not contaminated by error. Researchers who use randomized assignment in experimental studies or random sampling from populations distinguish between two kinds of measurement error. The first kind, **systematic error**, refers to influences on measurement that contain regularities and therefore bias the measurement outcomes. The second kind, **random error**, is not regular and therefore does not bias measurement outcomes, but nevertheless renders them less precise.

DEFINITIONS

DEFINITIONS

As an example of systematic error, suppose we have a test of verbal intelligence that is written in Russian. For those of us who do not read and write Russian fluently, this test will systematically underestimate our verbal intelligence. An example of random error would be a test of verbal intelligence in the test-taker's native language, comprising 25 questions randomly chosen from a large bank of such questions. By luck of the draw, some of the 25 questions will be easier than average and some will be more difficult, but the error in this case is distributed randomly throughout the test.

DEFINITION

A related term that is often used in psychological research is **reliability**, the extent to which measurement is free of random error. That is, a measure is reliable if it produces the same result every time under identical conditions. Although reliability and validity might seem similar at first, they are not synonymous at all. Our verbal intelligence test consisting of randomly chosen questions may be valid but it will not be perfectly reliable because of random error. On the other hand, we can have perfectly reliable measures that are nevertheless invalid. A verbal intelligence test in Russian administered to people who know no Russian at all will be very reliable, since each person will get a low score no matter how many times they take the test without learning more Russian in the meantime. No reasonable assessor would claim such a test is a valid measure of verbal intelligence for those people.

There is a large variety of systematic errors that psychologists have to contend with, and because validity is such a huge topic, they often subdivide it into several specific kinds of validity to be dealt with more or less separately. We will discuss errors first, and then return to the different kinds of validity. In their influential textbook, Rosenthal & Rosnow (1991) divide systematic errors into those arising in such a way as to not affect the respondent's responses (which they call 'noninteractional') and those that do ('interactional').

DEFINITIONS

Noninteractional errors usually are caused by the researcher or measuring instrument, whereas **interactional errors** may be caused either by the researcher, measuring instrument, or the respondent. The distinction is important because for noninteractional errors, the responses of the people being studied may be valid and so only the researcher needs correcting; whereas interactional errors contaminate the responses themselves and are usually more difficult to disentangle or correct.

Perhaps the most common kind of error arises from the researcher's expectations about how the findings ought to turn out. These **expectancy effects**, although pointed out as early as 1933 (by Rosenzweig), were not seriously examined until Rosenthal and his students took up the topic in the 1960s. An example of a noninteractional expectancy effect is researchers who, when hand-timing rats' progress through a maze, consistently click their stopwatches just before rats in one experimental condition reach the maze exit, but click their watches just after rats in another condition reach the exit. The result will be erroneously fast times for the former rats and slow times for the latter. An example of an interactional expectancy effect is a medical researcher who is friendlier, more optimistic, and more positive when administering a new experimental drug to patients than when administering a placebo to the control group patients. The patients in each condition might respond accordingly, and the resulting better outcomes for the experimental group will then reflect the researcher's behavior toward them rather than just the effects of the drug or placebo.

Another widespread kind of effect stems mainly from the respondent, and is sometimes called the **good subject effect** (Orne, 1962). People who participate in experiments or surveys usually are aware of the fact that they are being studied, and often they try to anticipate what the researcher is trying to study, or what the hypotheses are. Many (but not all) respondents are obedient or altruistic, and may attempt to provide researchers what they are looking for. These are adopting what Orne referred to as the 'good subject' role. A more general notion along these lines is the concept of **demand characteristics**, which refer to aspects of the tasks in a study that motivate respondents toward certain behaviors and away from others in ways that are extraneous to what is being measured. One important example of a demand characteristic is social desirability, which usually arises when people are asked to make choices in a setting where some choices clearly will make them seem nicer or better than others. Another demand characteristic that frequently occurs in survey research is a bias toward positive (or negative) responses induced by presenting the respondent with a list of questions whose response formats all point in the same direction.

DEFINITIONS

A number of strategies have been proposed for eliminating systematic errors that arise from researchers and/or respondents, and the interested reader will want to consult Rosenthal & Rosnow's (1991) text and the works cited therein. Several of the most effective strategies boil down to unobtrusiveness and blindness. **Unobtrusive measures** are those that do not alert people to the fact that they are being studied or measured, thereby minimizing the likelihood that any respondent-driven errors will arise. **Blindness** means keeping the people who run the experiment and/or the respondents participating in it ignorant of the object of inquiry, of what experimental conditions respondents are assigned to (e.g., which one is the drug and which the placebo), and of the hypotheses – thereby minimizing expectancy effects.

DEFINITIONS

Let us return to validity. There are two key motives for identifying different kinds of validity. First, measurements may be valid in some respects but not others, so it is incumbent on researchers to specify the way(s) in which their measures are valid. Second, many important psychological constructs are so indirectly observable or measurable that researchers can claim only quasi-validity or something approximating validity. Much of the literature on measurement validity distinguishes among three types: content validity, criterion validity, and construct validity.

DEFINITION

Content validity (sometimes also called 'face validity') is attained when measurements refer exclusively and exhaustively to the things that the researcher intends to measure. Clearly this is a subjective claim, requiring judgments by appropriate experts or informants. Researchers could claim to have evidence for content validity in their battery of tests of artistic creativity, for instance, if a panel of artists agreed that its contents referred to artistic creativity as they understood it. The researchers' claim would be stronger yet if the panel of experts concurred that *all* aspects of artistic creativity were captured by the test battery.

DEFINITION

Criterion validity is the degree to which a measure is corroborated by a criterion that is known to be valid already. At first glance, this might seem silly – why bother to invent a new measure if we already have a valid one? The most common reason is that an established valid measure may not be widely usable for practical or ethical reasons. Nevertheless, it may serve as a criterion against which to validate another measure that can be used widely.

In clinical psychology, diagnoses of psychological disorders by qualified clinicians are often time-consuming, expensive, and even intrusive. A paper-and-pencil test, on the other hand, is quick, inexpensive, and more private. A game that incorporates the test tasks and elements may not only be quicker and more pleasant, but even therapeutic! If the test yields the same diagnoses for clients whose disorders have already been ascertained by clinicians, then the test has shown criterion validity. Likewise, any aptitude or ability test may be said to have criterion validity if it retrospectively distinguishes between people who are known to perform well on tasks requiring the ability and those who perform poorly.

DEFINITIONS

Construct validity refers to whether one measurement relates to other measurements in the ways that mirror the relationships among their respective underlying constructs. If two measures are supposedly alternative ways of measuring the same construct, then they should correspond to one another (this is called **convergent validity**). Suppose, for example, that we have two tests of eye–hand coordination: throwing a ball at a target, and tracing a figure on paper while looking at it in a mirror. If both of these tasks are indicators of the same construct, then someone who gets close to the target in the ball-throwing task should take only a short time to complete the mirror-tracing task; and likewise someone who throws wide of the target should also take a long time to trace the figure in the mirror. On the other hand, if two measures pertain to distinct constructs, then they should not covary (this is called **discriminant validity**). If musical ability is distinct from mathematical ability, for instance, then not everyone who scores high on a music test of one should score high on a mathematical test. So a test of musical ability would be said to demonstrate discriminant validity if people's scores on it did not mimic their scores on a well-established test of mathematical ability.

Now, let us return at last to random error in measurement. This is the kind of error that statistics handles, and we will deal with it at some length in this book. Because random error is *not* systematic, we may still get reasonable estimates of 'true' measurements in the presence of random errors because they tend to cancel each other out in the long run. One way to put random error in perspective with everything else we have discussed so far is to imagine taking an examination on your knowledge of quantitative methods in psychology after reading this book. No test is perfect, so this hypothetical exam might very well suffer from systematic error (such as being too difficult or too easy). However, even if there were no systematic errors, your score on the test could be influenced by a host of randomly occurring factors – be they illness, lapses in concentration, or lucky guessing on some questions. So, if we were to construct a miniature theory or model to account for your score on the exam, we would have to hypothesize that there is the score you should have obtained given your knowledge and preparatory efforts, plus the influence of random error. That is,

$$\text{Your Score} = \text{True Score} + \text{Random Error}.$$

DEFINITION

If this model were accurate, then in a science-fiction universe where we could erase your memory of having read this book and get you to start again, if we did that over and over, your scores on the same exam should average out to your True Score. We would be assessing what is called **test–retest reliability**, which is the extent to which a measurement gives the same result when repeated under identical conditions. As you may have surmised, test–retest reliability may be ascertained only when tests or tasks are repeatable under identical conditions. This requires that the respondent be able to forget the previous occasion sufficiently that memory does not affect their current performance, and that there be no effect from practice.

Of course, no one can erase a person's memory that selectively, but psychological testers can present people with a batch of approximately equivalent items in a test, and obtain their average score on all of them. In a test of ability, for instance, classical test theory assumes this model of a person's score on each item:

$$\text{Item Score} = \text{True Score} + \text{Random Error}.$$

By collecting a person's scores on a large number of approximately equivalent items, the tester obtains estimates of two things: the person's true score, and the items' average departure from the estimated true score. The latter indicates the test's reliability, insofar as the smaller the average deviation away from the true score the more reliable the test is as a whole. Methods for assessing the reliability of an entire test-bank of items by analyzing the extent

DEFINITION

to which people score similarly on those items are said to be assessing **internal consistency reliability**.

SUMMARY

Measurement error occurs whenever measurements are influenced by something other than what the researcher intends to measure, in other words extraneous influences.
Validity denotes the extent to which measurement is not contaminated by error.
Systematic error refers to extraneous influences on measurement that contain regularities and therefore bias the measurement outcomes.
Systematic errors that do not affect the respondent's responses are **noninteractional**, whereas those that do are **interactional**.
Demand characteristics are aspects of the tasks in a study that motivate respondents toward certain behaviors and away from others in ways that are extraneous to what is being measured.

Unobtrusive measures are those that do not alert people to the fact that they are being studied or measured.
Blindness means keeping people involved in a study ignorant of certain characteristics of it, such as the object of inquiry, the experimental conditions respondents are assigned to (e.g., which one is the drug and which the placebo), and the hypotheses.

SUMMARY

Random error is not regular and therefore does not bias measurement outcomes, but nevertheless renders them less precise.
Reliability is the extent to which measurement is free of random error.
Test-retest reliability is the extent to which a measurement gives the same result when repeated under identical conditions.
Internal consistency reliability is the extent to which items that are supposed to measure the same construct yield similar results under identical conditions (see also construct validity).

Content validity (sometimes also called 'face validity') is attained when measurements refer exclusively and exhaustively to the things that the researcher intends to measure.
Criterion validity is the degree to which a measure is corroborated by a criterion that is known to be valid already.
Construct validity refers to whether one measurement relates to other measurements in the ways that mirror the relationships among their respective underlying constructs.

Types of measurement and varieties of data

We have already seen that in most psychological perspectives, measurements and observations are not the same as the construct being measured or observed. Theories of measurement, therefore, are accounts of the imperfect correspondence between measurements or observations and the construct itself. So far, we have concentrated mainly on the psychological aspects of measurement. Now we will borrow some useful concepts from the mathematical side.

Most of the mathematical aspects of measurement boil down to what Funtowicz & Ravetz (1990) call 'craft skills with numbers.' Measurement amounts to using symbols such as numbers to represent states or values of the construct being measured, so these skills include knowing how to use those symbols and numbers appropriately and meaningfully. For instance, those people wearing digital watches who read out the time right down to the second are displaying craft ignorance about numbers if their watches are not accurate to within one second (and they usually are nowhere nearly that accurate).

One crucial, though risky, judgment that each of us as researchers must make is what we may presume about the relationships between our measurements and the constructs to which they pertain. Suppose we try to measure people's knowledge about HIV by setting them a test with 100 multiple-choice questions designed by appropriately qualified experts on HIV. The test scores have a range from 0 to 100 correct. At first glance, it might seem that we could assume a perfect correspondence between their score and the amount of knowledge they have about HIV. Were that the case, a statement such as 'a person who scores 80 is twice as knowledgeable as someone who scores 40' would be true. After all, 80 is twice 40.

A bit of thought, however, should convince us otherwise. Even a 100-question test would be unlikely to include all relevant questions about HIV. So someone who got a score of 0 might not be totally ignorant about this topic – they just were not asked the 'right questions.' If 0 does not correspond to total ignorance, then 80 does not necessarily indicate twice as much knowledge as 40. Likewise, someone who scores 100 does not necessarily know everything about HIV.

What about a statement such as 'an increase in score from 80 to 100 indicates the same increase in knowledge as an increase from 0 to 20?' Both are 20-point gains in test score. Nevertheless, it might be more difficult to attain the knowledge required to move from a score of 80 to 100 than that required to move from 0 to 20. The statement would be plausible only if we were convinced that the 100 questions were equally difficult. This example demonstrates that it is quite possible to reduce an apparently fully quantified scale, in the harsh light of honest examination, to a scale that merely indicates whether someone knows more than someone else. It also underscores the necessity of carefully investigating the measurement properties of any scale before coming to conclusions about how it may be used.

Discrete and continuous variables

One of the most fundamental properties of any variable is whether its domain consists of distinct possible states or values, or a range of possible values.

DEFINITION **Discrete** variables can only take particular states or values and no others. Gender is usually treated as a discrete variable with two states. An employee's rank in her or his organization's hierarchy is a discrete variable with ordered states, each corresponding to a particular rank. Many numerical variables are discrete as well. Any counted variable is a discrete variable because it can only have integer values such as 1, 5, or 23 (e.g., a family cannot have 2.3 children). The same is true of rank-order, as in the finishing order in a race – even taking ties into account.

DEFINITION **Continuous** variables are those which (theoretically) could take any value within their range. Weight, strength, agreement with an attitude, happiness, and task motivation all are examples of continuous variables, because they are not only matters of degree but in principle infinitely fine-grained. Some confusion can arise from the fact that many variables are *made* discrete even though their underlying constructs are continuous. A widespread example is the Likert scale of agreement or disagreement, as shown in Figure 2.2. The underlying construct is degree of agreement or disagreement, even though there are only five possible values associated with the variable itself. Whereas discrete variables are accurately measurable in principle, continuous variables can only ever be measured approximately (i.e., with limited precision). In practice, therefore, *all* variables are rendered discrete because of limits to precision. So why is this distinction important?

First, how we think of a construct in terms of variables influences how we treat it statistically and what can be meaningfully said about it (hence the gibe about 2.3 children). A continuous variable can take any value in its range, so reporting fractional values (such as saying someone is 42.5 years old) may be meaningful. Second, in some areas of psychology the issue of whether a construct is inherently discrete or continuous constitutes a worthwhile topic in itself. Early work in psychophysics, for example, addressed the issue of whether the sensation of intensity (as in loudness, brightness, or heaviness) was best regarded as continuous or having discrete thresholds (see Luce, 1997, for a thoughtful reminiscence concerning that research). Recent investigations of how people categorize things or think about categories have drawn psychologists to realize that people do so using both discrete and continuous (or 'fuzzy') heuristics (for an entertaining, brief, but informative exegesis on human category heuristics, see Pinker, 1997: 308–311). Discrete categories are characterized by hard-and-fast rules about what is or is not a member of that category, whereas fuzzy categories have blurry boundaries and are described in terms of prototypical members, key features, and family-like resemblances.

Likewise, some variables that once were assumed to be discrete have been found to be continuous. Handedness is a good example. I write, draw, throw, thread a needle, and use a hammer with my left hand, but I wield a racquet, bat, computer mouse, and use chopsticks with my right hand. Am I left-handed or right-handed? I am not simply ambidextrous since in each of those activities I have a definite preference for one hand over the other. These days most researchers use measures of handedness that permit degrees of preference for the left or right.

Finally, some of our most important social issues hinge on decisions about whether a variable should be thought of in discrete or continuous terms. Racial or ethnic identity, employment status, gender, and sexual preference

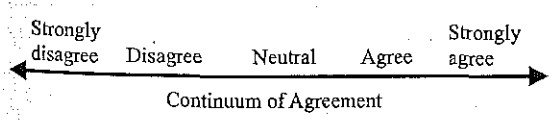


FIGURE 2.2 Likert scale example

are good examples of constructs that often are treated as if they are discrete for certain political purposes and continuous for others. Political agendas may be ubiquitous, but researchers must take responsibility, especially where politicians and bureaucrats will not, for critically assessing the status of the variables they use in their investigations. In clinical psychology, many 'pathologies' are treated as if they are discrete for diagnostic purposes (e.g., either you are clinically depressed or you are not), even when the tests or scales used to measure them use continuous variables. Educational institutions often do the same – students either fail or they do not, regardless of whether they fail by four points or 40.

In summary, the issue of whether a construct is best thought of as discrete or continuous is fundamental and often not easily settled. Luce's (1997) retrospection on unresolved matters in mathematical psychology includes the discreteness–continuity question, along with the issue of how to handle uncertainty and vagueness, in his list of thorny problems. In the next section, on levels of measurement, we will revisit this issue in the form of constructs that seem to be both categories and continuous scales at the same time.

Levels and types of measurement

Not only do we need to decide whether a variable is best thought of as discrete or continuous, but also what **level of measurement** is involved. The main reason for this is so that we may ascertain what constitutes meaningful or appropriate uses of our data, and what kinds of questions may be answered with those data.

DEFINITIONS

The conventional wisdom in psychological research is that there are four levels of measurement:

- **Nominal**, in which a variable has categories or states that differ in kind but not in order or degree;
- **Ordinal**, in which a variable has states or ranks that are ordered but whose differences are not quantifiable;
- **Interval**, in which a variable has values on a scale whose differences are quantifiable but which lacks a true zero-point;
- **Ratio**, in which the differences are quantifiable and the scale has a true zero-point.

Nominal variables are by definition discrete. The simplest kind is a binary variable that has two categories; presence and absence. More complicated nominal variables have more categories, and they may also have nested categories, i.e., where one category contains others. We could classify the teachers at a university according to which department or division they are in, which would give us a nominal variable with departments as the categories. Suppose that every department also is assigned to one faculty. Then departments are nested inside faculties, so the entire framework has two levels. We now have two nominal variables, faculty and department, with the former containing the latter.

Constructing nominal variables is often more difficult than it seems. The most important criteria to bear in mind are these:

- *Exhaustiveness*: Do the categories cover all conceivable cases?
- *Exclusiveness*: Are the categories unambiguously distinct, or do they overlap?
- *Appropriate specificity*: Are the categories as fine-grained as they need to be, without being hair-splitting?

In many research situations, it seems tempting or even practical to sacrifice exhaustiveness, especially when we do not know all of the possible categories in advance or when there are a great many of them. The most common method of truncating an undetermined or lengthy list of categories is to stick a **catch-all** category at the end, labeled 'Other' or 'Miscellaneous.' However, we take two risks in doing this. One is that the catch-all category will obscure important differences among cases that it includes. The second risk arises when the catch-all category is used as one of the alternatives that may be chosen by the respondent in answering a question, and stems from what is called the 'catch-all underestimation bias,' or CAUB (Fischhoff, Slovic & Lichtenstein, 1978). This is a tendency for people to recall fewer instances or to underestimate the likelihood of the catch-all option in comparison with explicitly listed options.

The issue of appropriate specificity is related to the CAUB problem. People are more likely to recall events that are explicitly represented in a list of alternatives than those that are not (Tversky & Koehler, 1994). Likewise, unpacking a category into specific subcategories may remind people of possibilities they have overlooked. Consequently, our ability to recall instances of some category may be boosted by unpacking that category into subcategories. For example, if people are asked how many times they have been dishonest in the past month, they will generally remember fewer instances than if you ask them how many times in the past month they have lied, stolen anything, concealed anything from someone, led someone to believe something that was not true, and so forth. However, the solution is not always to provide exhaustive lists of specific alternatives. Instead, we need to take specificity into account when drawing inferences about people's recollections of events, especially if we are comparing events or different studies.

Since many of the categories that humans use in natural language are 'fuzzy' in the sense that they have blurry boundaries and overlap with one another, constructing exclusive categories is not always easy and may seem contrived. Consider the difficulties involved in deciding whether a person's death from a self-inflicted gunshot wound while under the influence of alcohol is a suicide, a self-injury without an intention to die, or an accident. Even if we had a great deal of information about the person's emotional state, thoughts, and actions at the time, it is quite possible that we would still be undecided about how to classify this event.

The problems of category exclusiveness and ensuring that there is one category per case are related, but not identical. Exclusive categories are necessary but not sufficient to ensure that there is one category per case. For instance, in Huston *et al.*'s (1981) study of who was likely to intervene in a violent crime incident, the investigators classified respondents according to whether they had received emergency training or not. The classification scheme they used included first aid, life-saving, self-defense, police, and medical. The potential problem here is not that these categories overlap but that one person may have received more than one of these kinds of training.

Now let us turn to **ordinal** variables. Although it makes almost no difference in how we analyze ordinal data, there are two rather distinct ways of thinking about these variables. One is in terms of ordinal categories, which are ordered from lowest to highest. For example, the widely-used Likert-style format indicating degree of agreement or disagreement with a statement is a categorical ordinal scale because it places respondents in one of five ordered categories: strongly agree, agree, undecided, disagree, or strongly disagree. The other, more general way is in terms of rank-ordering cases from lowest to highest, as in the finishing order of a race. The main reason for mentioning this distinction now is to avoid confusion later on, since some statistical techniques refer to values on an ordinal scale and others to 'ranks.'

Ordinal variables (and their more quantifiable cousins) may be further divided into **bipolar** and **monopolar** variables. The agree–disagree format is an example of a bipolar variable, because it has two opposite ends with a 'neutral' midpoint. The so-called 'bipolar personality disorder' of manic depression is an example of a bipolar construct, ranging from extreme mania to extreme depression. Monopolar variables, on the other hand, range from low to high on attributes that are unidirectional. A variable measuring degree of anxiety that ranges from 'low' to 'very high' is an example of a monopolar variable.

Bipolar variables are very popular, but they are also more easily misused than monopolar ones. First, the researcher must be sure that the opposite poles really should be located on the one dimension and not separate dimensions. There have been extensive debates in psychology over such issues as whether happiness and unhappiness require separate monopolar rating scales, or whether masculine–feminine is a bipolar or two-dimensional construct. Second, the researcher should ensure that the midpoint on these scales has only one meaning for respondents. It is possible for someone to place themselves at the midpoint of an agree–disagree scale because they have no opinion on the issue concerned, or because they are in great conflict and therefore ambivalent about it, or because they are genuinely balanced between agreement and disagreement. If we anticipate getting two or more of these responses from people, then we should provide separate response categories for them or filter those responses out with a prior question.

All ordinal variables, whether monopolar or bipolar, must have an unambiguous and agreed-upon ordering of their response categories. For instance, in a variable designed to measure level of satisfaction with a service, respondents might not agree on the order among options that are too similar, such as 'so-so' versus 'adequate' versus 'satisfactory.' Options with similar meanings should therefore be avoided when constructing these scales.

Likewise, specificity and vagueness may need to be taken into account. Experimental studies of the meanings people attach to probability phrases such as 'very likely' or 'low probability' (e.g., Wallsten *et al.* 1986) have found that these phrases usually refer to a range of probabilities, and the ranges may overlap and differ in width. A term such as 'possible' is so vague for many people that it encompasses probabilities from nearly 0 to nearly 1, so it is useless as a response option. The lesson to be learned here is to try to ensure that response categories not only have distinct meanings but are similar in their specificity.

Finally, the meanings of response categories may vary considerably with context. 'Very often' could mean more than twice in one year when referring to earthquakes on the Australian East Coast but not in New Zealand, and certainly not when referring to how frequently citizens of either country brush their teeth. One way around this problem is to standardize and specify the context by providing a **reference frame** and **anchors**.

Suppose we want to evaluate the performance of public transport bus drivers, and we are designing a question to elicit customers' perceptions of their driving performance. If we simply ask them 'How skillful is X as a driver on this route?' and request them to choose a response on an ordinal scale ranging from 'very unskillful' to 'very skillful,' we will be unable to tell what *reference frame* the passengers have used in answering this question. Skill, after all, may be judged with reference to:

- what they expect,
- the average bus driver,
- the average automobile driver,
- their ideal driver,
- their beliefs about what is adequate driving skill,
- other bus drivers they have known, etc.

A better version of this question would include the references of our choice in the question and appropriate *anchors* within the scale. For instance, we could ask ‘Compared to other bus drivers you have known, how skillful is X as a driver on this route?’, and then incorporate anchors in the scale, such as ‘About average in skill’ or ‘Less skillful than any of them.’

The ordered categories in an ordinal scale usually denote differences in magnitude or intensity, as in the illustrations thus far. Another kind of ordering is *cumulative* or inclusive, so that someone who is described by one category is also described by the categories that lie below it. For instance, someone who has completed secondary schooling also has completed primary schooling, but not vice versa. This kind of scale is called a **Guttman scale**, although examples of them had been used before Guttman described their properties in the 1940s. An example that commonly occurs in educational psychology and testing generally is a scale composed of several questions that differ in difficulty. People who answer a particular question correctly should answer all the less difficult questions correctly too. This kind of scaling in educational testing was further elaborated by Rasch in the 1960s and others since then.

DEFINITION

By analyzing the response patterns to a set of items, a research may determine whether they form a cumulative scale. We won’t plumb the details of that analysis here, but instead will settle for an appreciation of what is involved through a simple example. Imagine investigating people’s attitudes about capital punishment, and asking them whether they would prefer that the death penalty be legally available for various criminal offenses. If all of our respondents believe that the death penalty should be available only for the most serious crimes and if they also agree on which crimes in our list are more serious than others, then we should be able to find a cumulative scaling pattern in their responses. For illustration, consider this very short list: multiple homicides, one homicide, multiple rapes. If our respondents agree that this list is ordered from the most grave offense downward, then each should provide one of response patterns 1–4 in Table 2.1. All of the other possible patterns would disconfirm the cumulative scaling hypothesis.

TABLE 2.1 Response patterns

Pattern	> 1 homicide	1 homocide	> 1 rape	
1	No	No	No	
2	Yes	No	No	Confirm
3	Yes	Yes	No	scaling
4	Yes	Yes	Yes	
5	No	Yes	Yes	
6	Yes	No	Yes	Disconfirm
7	No	No	Yes	scaling
8	No	Yes	No	

Sometimes we would like to treat a construct as if it was both a scale and a category at the same time. Many clinical diagnostic constructs have this property because they are matters of degree and therefore measurable by using scales. Depression is an example, because even though people may be depressed to a greater or lesser degree, certain cutoff points on indicators of depression must be exceeded before someone is diagnosed as 'clinically' depressed. Under these conditions, it is worthwhile to think of ordinary scales in a cumulative way, by considering the set of people who score at least as high as each of the points on the scale. The further up the scale we go, the smaller (and therefore more exclusive) the set of people we will find. As we move down the scale, we include more and more people. To distinguish this viewpoint from the Guttman scale, we may refer to it as the **cumulative scale** viewpoint. By way of illustration, suppose we ask 150 people to indicate their level of subjective fear of snakes on the five-point scale shown in the next table, and then we ask them to do the same for their fear of dogs. The number of responses in each category is displayed in the leftmost two columns.

DEFINITION

The right-hand columns with the boldfaced numbers redisplay the same information from a cumulative scale point of view. Each level is associated with a set of people who have scored at least as high as that level. For instance, 60 people have indicated that they fear dogs at least at a Moderate level. If a specialist in phobias were to say that anyone who expressed at least a 'severe' fear level was phobic, then we could immediately see that there are three times as many snake phobics in this sample as dog phobics (75 vs. 25). In fact, at any level we care to choose, snake fear is a *more inclusive phenomenon* than dog fear, since the number of people whose snake fear is a certain level or higher is greater than the number of people whose dog fear is the same level or higher. So an advantage of the cumulative scale view is that we are able to treat fear as if it is both a category and a scale, thereby enabling us to address questions such as whether more people fear snakes than dogs.

Finally, let us consider **interval** and **ratio** levels of measurement. We have already seen that an interval level of measurement requires that we have

TABLE 2.2 Snake and dog fears

	Number at each level		**Cumulative viewpoint**		
Fear level	snakes	dogs	snakes	dogs	
Profound	25	5	**25**	**5**	**at least Profound**
Severe	50	20	**75**	**25**	**at least Severe**
Moderate	45	35	**120**	**60**	**at least Moderate**
Slight	20	40	**140**	**100**	**at least Slight**
Not at all	10	50	**150**	**150**	

intervals between scale values that are quantifiable and therefore comparable. Conventionally, the intervals are equal, which entails that they have a *linear relationship* with their underlying construct. The percentage of questions a student answers correctly on an exam, for instance, has this property if we agree that the underlying construct is quantity of knowledge about the subject matter of the exam and all of the questions are equally weighted.

The issue of whether we can ascertain when we have interval-level measurement has provoked debates in many areas in psychology. A brief review of the more common pitfalls in these debates may help you in your own research and in critically assessing others' efforts. First, it is not always clear that we can establish whether there is a simple linear relationship between our constructs and our measures. Revisiting our example of the percentage of correct exam answers, someone could take a cumulative scale viewpoint and point out that whereas 30 members of the class got at least 60% of the questions correct, 20 got 70% correct and only seven got 80% correct. If instead of mere quantity of correct answers, we consider difficulty as well, we might conclude that percentage of correct questions is not linearly related to knowledge of (more difficult) subject material after all. Likewise, as in the earlier HIV exam example, we might want to investigate whether some questions are more difficult than others.

It is also easy to confuse the properties of the numbers on a scale with the scale's measurement properties. A famous suggestion that dates from 1738 illustrates this point quite well. The suggestion by Daniel Bernoulli hinges on the observation that increasing a poverty-stricken person's wealth by \$100 is much more highly valued by that person than increasing a billionaire's wealth by \$100. People's subjective valuation of money declines with increasing wealth. Even though the monetary scale has equal intervals, it is not linearly related to people's subjective valuation and therefore *as a measure of subjective utility*, money does not provide an equal-interval measure.

The HIV examination example discussed earlier also highlighted a frequently misunderstood distinction between having a zero on a scale and having that zero correspond to a true zero in the construct. While there is nothing wrong with saying that someone got no questions correct on an exam, we usually cannot infer from this that they know nothing about the subject. The same is true of many psychological scales, such as self-esteem or depression. Zero scores on such scales do not entitle us to conclude that someone has no self-esteem or zero depression.

That said, it is also worth bearing in mind that some scales may have an absolute zero and yet be ordinal. Psychological state and trait questionnaires have legions of items whose response formats are something like 'Never,' 'Rarely,' 'Sometimes,' 'Frequently.' Moreover, some counted data must be coded in unequal intervals for practical or ethical reasons. A typical case is a question about drug usage, in which a respondent is asked to indicate how many times they have used a particular drug: 'Never,' 'Once,' '2–4 times,' and 'more than 4 times.' Neither of these scales has equal intervals, but both have true zeroes.

How do researchers decide what level of measurement they can or should use? To some extent they tailor their variables to the hypotheses they want to test or the questions they want to answer. Many hypotheses and research questions contain implicit assumptions or requirements about measurement. A good researcher will figure out what these are before constructing a measurement instrument.

Some examples may help illustrate the issues involved. Suppose we are interested in whether women and men are differentially treated in the engineering profession. Here are some of the ways in which we might formulate hypotheses about this discrimination:

1. Women are underrepresented in engineering.
2. Female engineers tend to have lower status than male engineers.
3. Female engineers tend to earn less than male engineers.

What kind of variable do we need for the first hypothesis? One thing we could do would be to compare the percentage of women in the labor force who are engineers and compare that with the percentage of men in the labor force who are engineers. We could contrast that with similar comparisons in other professions. In either case, we just need a variable that measures whether an employed person is an engineer or not. That is a nominal variable, since it has two categories (Engineer, Not engineer) which are not ordered.

The second hypothesis has a key concept, 'status.' Since there are higher and lower-status positions, we know that we must measure status at least at the ordinal level. We would find that there are standard designations for engineers' occupational positions that distinguish between the lower and higher levels in the profession. Or we might want to construct a measure based on public perceptions of the status of engineers. That is all we need if we want only to determine whether females occupy lower-status positions than males.

If we wanted to go further and ask 'how much lower?' then we would need an interval- or ratio-level measure of status. Unfortunately, we would find that engineer grades form only an ordinal-level scale. We would have to conclude that 'how much lower?' cannot be meaningfully addressed with the measure of status available. The third hypothesis also requires only an ordinal scale for earnings. But in this case if we want to ask 'how much less?' perhaps we can get an answer if we are willing to use salary as a measure of status.

SUMMARY

Discrete variables can only take particular states or values and no others. **Continuous** variables are those which (theoretically, at least) could take any value over their range.

The conventional four **levels of measurement** are:

- **Nominal**, in which a variable has categories or states that differ in kind but not in order or degree;
- **Ordinal**, in which a variable has states or ranks that are ordered but whose differences are not quantifiable;
- **Interval**, in which a variable has values on a scale whose differences are quantifiable but which lacks a true zero-point;
- **Ratio**, in which the differences are quantifiable and the scale has a true zero point.

In a **Guttman scale**, anyone who is described by one category also is described by the categories that lie below it.
In the **cumulative scale** viewpoint, each value on a scale is associated with the number of people who have scored as high as that level or higher.

Missing data, absent people, and the fallacy of false precision

When data are only partly known or imprecise, there is an understandable temptation to represent them as if they are nevertheless fully known and precise. In the psychology of judgment under uncertainty, several researchers have studied what they call 'ambiguity aversion' (Ellsberg, 1961; Einhorn & Hogarth, 1985), which is a tendency for people to prefer precise information over imprecise information, even when imprecision is more realistic. Ambiguity aversion has been found not only in hypothetical situations but also in gambles with real payoffs, and even after people are exposed to written arguments persuading them not to indulge in it.

Aversion to imprecision of any kind manifests itself at the institutional and organizational levels as well. Some organizational agendas also promote illusory precision because it makes the organization appear to be in control and more authoritative (Downs, 1966; Linnerooth, 1984). As the scientific policy analysts Funtowicz & Ravetz (1990: 11) point out, the 'simplest, and still most common response of both the decision-makers and the public is to demand at least the appearance of certainty. The scientific advisors are under severe pressure to provide a single number, regardless of their private reservations.'

Unfortunately, once data are recorded as if they were precise, any further analysis of them will treat them as if they actually are precise. Sloppy estimates will appear as if they are error-free. Forecasts or decisions based on them will be made with greater confidence than is warranted, and so on. I refer to this trap as the 'fallacy of false precision' (Smithson, 1990). Avoiding false precision entails *recording* both qualitative and quantitative information about uncertainty, including the likely causes of it, on a case-by-case basis.

Missing data may arise from a variety of causes, especially in the early stages of a research project when measures are still being refined and data collection methods are most error-prone. Many popular statistical packages designate missing data with only one symbol, but this is often inadequate. The following list contains the most common reasons why data end up missing. Recording the type and cause of missing data enables a researcher to ascertain whether there is any pattern or regularity in missing data that might invalidate their measurements or conclusions.

- Respondent refused to answer.
- Question is not applicable to respondent.
- Respondent did not understand the question or task.
- Respondent's answer does not fit any available categories.
- Respondent's answer varies among categories depending on circumstances.
- Respondent's answer is incomprehensible.
- The data were lost or misrepresented.

Not only can data go missing, so can people. Participants in experiments may drop out before the conclusion of the study, potential respondents may refuse to be interviewed for a survey, and former patients may be impossible to track down for follow-up studies. Rather than ignoring these issues, patterns in missing data and/or missing people should be studied and understood. There are at least two reasons for doing so. The most common reason is that missing data may bias the findings of a study, and any conscientious researcher will want to ascertain whether that is a possibility. A typical example is a longitudinal study of cognitive impairment in elderly Japanese (Liang *et al.*, 1996), in which the researchers found that although urban residents were more likely than rural residents to drop out of the study, the dropouts did not differ from the stayers in mental or physical health.

The second reason for analyzing missing data patterns is that sometimes they may constitute an important research finding in themselves. For instance, in a social psychological study of gender solidarity in organizations, Fajak & Haslam (1998) presented members of an organization with hypothetical organizational charts in which people were identified by gender and position, and asked who should be promoted if a position at a certain level became vacant. They analyzed the patterns of *refusals to nominate a candidate*, as part of the evidence of when a participant identified with or declined to identify with their own gender.

Questions and exercises

Q.2.1. In an educational psychological study, the amount of time each participant spent reading the material in preparation for a test was recorded to measure motivation. What is the *construct* in this study?

Q.2.2. Give an example of systematic error in measurement. Give an example of random error in measurement.

Q.2.3. What kind of effect is likely to contaminate people's responses to the following question:
'How much money did you donate to charities during that last financial year?'

Q.2.4. Why is having a rat run the same maze several times *not* a good assessment of test–retest reliability?

Q.2.5. A well-known statistics textbook author was criticized for using the 'number of hairs on a goat' as an example of a continuous variable. He corrected this mistake in a later edition. Why was the criticism correct?

Q.2.6. Which is true of the variable reaction-time, measured in milliseconds?

(a) It is continuous.
(b) It is discrete.
(c) It has ordinal-level measurement.
(d) It has nominal-level measurement.
(e) It has interval-level measurement.

Q.2.7. For what purpose(s) would your score on a statistics exam have interval-level measurement? For what purpose(s) would it have ratio-level measurement?

Q.2.8. Which is true about 'self-esteem as measured by how many out of 10 positively worded statements about the self a person agrees with?'

(a) It is continuous.
(b) It can never equal 0.
(c) It has interval-level measurement.
(d) It has ordinal-level measurement.
(e) It has nominal-level measurement.

Q.2.9. Which of the following variables are discrete and which continuous?

(a) Religious denomination.
(b) The temperature of a person's palm.
(c) Response time in a test of reflexes.
(d) Number of successes in a series of puzzles.
(e) Subjectively judged distance from self to an object.

Q.2.10. From the standpoint of measurement error, what is problematic with this question?
'What was your income during the most recent financial year?'

What Is Good? – A Comparison Between The Quality Criteria Used In Design And Science

ABSTRACT

The human-computer interaction community is an umbrella for many disciplines. Conflicts occur from time to time, in particular between scientists and designers. This article compares the quality criteria used in design with those used in science, in order to gain insight into what design can contribute to the development of science. From the scientific perspective, the weakest point of design knowledge is its limited generalizability.

Author Keywords

Design, science, quality, criteria

ACM Classification Keywords

H5.0. Information interfaces and presentation (e.g., HCI): General.

INTRODUCTION

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. At times, open conflicts between the different groups emerge, in particular between scientists and designers, since they have the least common ground [2]. 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 CHI2006 conference ignited a shouting match between academics and practitioners [1]. At the conference itself the conflict recurred 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 [12]. Both groups felt the need to defend themselves, which indicated that both had the feeling of being under attack to start with. Stuart Feldman, the president 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 and they include a high proportion of scientific studies. Non-scientific studies, such as experience reports and case studies, are more often found in the extended abstracts. Furthermore, the main proceedings use the “archival format” whereas 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 hence published in the archival main proceedings, while non-scientific 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.

Even though science is highly esteemed, Chalmer [3] 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 & Walker [4] even suggested that the confusing epistemology of science may be unable to function as a blueprint for the epistemology of design. Levy [8] 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 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 [9]. 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. In this paper I would therefore 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 may also provide insights into the direction in which design methods have to evolve to become more scientific.

This comparison of quality criteria does not imply that design should use the classical scientific method. Cross provided an excellent historical review of the developments in the various design methodologies [5]. 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 CHI2007 workshops on “Converging on a science of design through the synthesis of design methodologies” and on “Exploring design as a research activity”.

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 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 well 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.

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 is therefore 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" then this may hold true only for people who share my paranoia about small creatures with many legs. Objectivity is therefore a good method for increasing the generalizability of a statement. Generalizability is also related to the repeatability of an experiment. If the results of an experiment are objective, meaning that they are not dependent on the experimenter, then others should be able to repeat the experiment with exactly the same results. Furthermore, time itself should not matter. Repeating the experiment at a later point in time should yield the same results. For the design community time does matter and hence the CHI conference has the section “Contemporary Trends”. Like fashion, the results of design work are expected to change over time, which makes them less generalizable. However, some design classics, such as the Tizio lamp, appear to be timeless.

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 is usually 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 the 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 to a different solution.

Falsifiability is another important criterion that is known to both scientists and designers. Originally proposed by Karl Popper (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 its value.

The use of falsifiability in design is very similar. A solution must admit 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 are usually easy to observe, and this criterion overlaps 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, multiple definitions of truth exist. The Wikipedia lists many theories of truth including correspondence, coherence, constructivist, consensus, pragmatic, performative, semantic, and Kripke's theory. The correspondence and coherence theories are probably the most acknowledged, and hence 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 then they are considered to be true. This direction in the relationship between truth and reality is usually 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 been built. The United States Patent and Trademark Office (USPTO) has made an official policy of refusing 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 have often 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 understood only in 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 laws 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, 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 20 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.

Lastly, 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 left). 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 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 right).

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 and hence it should not be judged by the fuel consumption criteria. The Hummer SUVs are not designed for driving fuel-efficiently from A to B.

CONCLUSION

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. Often they are not even interested in it. Jon Kolko, editor of ACM's <interactions> magazine, rejected this very manuscript based on its academic format:

"However, the submission is in a very academic format, while Interactions Magazine is intended to read in a more approachable and casual manner - specifically, it is intended to be of worth to practitioners, who may not be familiar with or interested in the very specific and grounded citations and discourse you have provided." [7]

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 [11] that such a method would lead to knowledge that is "far more reliable, secure, and trustworthy than scientific knowledge". Currently, designers who want to work as scientists have to become either engineers or psychologists. Since they often lack training in these disciplines they have a natural disadvantage.

Until considerable progress has been made in defining a suitable epistemology for design, 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 hence of lesser value. Scientists, on the other hand, need to acknowledge that the highly general knowledge they produce is often too abstract to improve society. It requires a skilled designer to translate this knowledge to a specific context of use.

As for the CHI conference, 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 static quality patterns [10] appear suitable for defining the sections, but the terms "intellectual" and "social" carry different meanings in the various sub-communities, and hence may cause misunderstandings. Maybe the sections could be called "Discovery" and "Invention". The latter would collect contributions that are aimed at improving society. The discovery section would gather contributions that present scientific insights. Whatever principle is used to divide the proceedings, it should be made explicit.

The use of 'best paper' awards is another ranking method. 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 multi–disciplinary community. The sub-communities of design, education, engineering, management, research, and usability will co-exist, but future shouting matches cannot be excluded.

REFERENCES

[1] Arnowitz, J. and E. Dykstra-Erickson, *CHI and the Practitioner Dilemma*. Interactions, 2005. **12**(4): p. 5-9.

[2] Bartneck, C. and M. Rauterberg, *HCI Reality - An Unreal Tournament*. International Journal of Human Computer Studies, 2007. **65**(8): p. 737-743.

[3] Chalmers, A.F., *What is this thing called science?* 3rd ed. 1999, Indianapolis: Hackett. xxii, 266.

[4] Cross, N., J. Naughton, and D. Walker, *Design method and scientific method*. Design Studies, 1981. **2**(4): p. 195-201.

[5] Cross, N., *Science and design methodology: A review*. Research in Engineering Design, 1993. **5**(2): p. 63-69.

[6] Dourish, P. *Implications for design*. in *SIGCHI conference on Human Factors in computing systems*. 2006. Montreal, Quebec, Canada: ACM Press.

[7] Kolko, J., *Email from Jon Kolko to Christoph Bartneck on August 21st, 2007*. 2007.

[8] Levy, R., *Science, technology and design*. Design Studies, 1985. **6**(2): p. 66-72.

[9] Pirsig, R.M., *Zen and the art of motorcycle maintenance: an inquiry into values*. 1974, New York: Morrow. 412.

[10] Pirsig, R.M., *Lila : an inquiry into morals*. 1991, New York: Bantam Books. 409.

[11] Pitt, J., C., *What Engineers Know*. Techne, 2001. **5**(3): p. 17-30.

[12] Wolf , T.V., et al. *Dispelling "design" as the black art of CHI*. in *SIGCHI conference on Human Factors in computing systems (CHI)*. 2006. Montreal, Quebec, Canada: ACM Press.

Research in Engineering Design (1993) 5: 63–69

Research in
Engineering
Design

Science and Design Methodology: A Review

Nigel Cross

Design Discipline, Faculty of Technology, The Open University, Milton Keynes MK7 6AA, UK

Abstract. *Design methodology has always seemed to have a problematic relationship with science. The "design methods movement" started out with intentions of making design more "scientific", but the more mature field of design methodology has resulted in clarifying the differences between design and science. This paper reviews the relatively short history of design methodology and its relationship with science, maps out some of the major themes that have sustained it, and tries to establish some agreed understanding for the concepts of scientific design, design science and the science of design.*

Keywords. Design methodology; Design science; Science of design

1. Introduction

It is now more than thirty years since the first conference on design methods was held in London in 1962 (Jones and Thornley 1963). This conference is generally regarded as the event which marked the launch of the "design methods movement", which in turn led to the emergence of design methodology as a subject or field of enquiry. Of course, the field was based on some earlier work (the earliest reference in design methodology literature is probably Zwicky's work on morphological method published in 1948 (Zwicky 1948)), but the 1962 conference was the first time that "design methods" received substantial academic recognition.

So the history of design methodology is still rather a brief one. Some previous "history" reviews have been by Broadbent (1979) and Cross (1980 1984). In 1986 the Design Methods Group celebrated its twentieth anniversary with some special reviews in its journal *Design Methods and Theories*. Finger and Dixon (1989) opened the first issue of *Research in Engineering Design* with a review of previous research.

Correspondence and offprint requests to: Nigel Cross, The Open University, Milton Keynes MK7 6AA, UK.

2. A Brief Overview

The origins of the emergence of new design methods in the 1950s and 1960s lay in the application of novel, "scientific" methods to the novel and pressing problems of the Second World War—from which came OR and management decision-making techniques—and in the development of creativity techniques in the 1950s. (The latter was partly, in the USA, in response to the launch of the first sputnik, which seemed to convince American scientists and engineers that they lacked creativity.)

The new "design methods movement" developed through a series of conferences in the 1960s and 1970s – London, 1962 (Jones and Thornley 1963); Birmingham, 1965 (Gregory 1966); Portsmouth, 1967 (Broadbent and Ward 1969); Cambridge, Mass., 1969 (Moore 1970); London 1973; New York 1974 (Spillers 1974); Berkeley, Calif., 1975, Portsmouth again in 1976 (Evans *et al.* 1982) and again in 1980 (Jacques and Powell 1981).

The first design methods or methodology books also appeared in this period – Hall (1962), Asimow (1962), Alexander (1964), Archer (1965), Jones (1970) and Broadbent (1973), together with the first creativity books – Gordon (1961) and Osborn (1963).

However, the 1970s also became notable for the rejection of design methodology by the early pioneers. Christopher Alexander said: "I've disassociated myself from the field.... There is so little in what is called 'design methods' that has anything useful to say about how to design buildings that I never even read the literature any more.... I would say forget it, forget the whole thing.... If you call it 'It's A Good Idea To Do', I like it very much; if you call it 'A Method', I like it but I'm beginning to get turned off; if you call it 'A Methodology', I just do not want to talk about it." (Alexander 1971), and J. Christopher Jones said: "In the 1970s I reacted against design methods. I disliked the machine language, the behaviourism, the continual attempt to fix the whole of life into a logical framework." (Jones 1977)

These were pretty harsh things for these two founding fathers to say about their offspring, and were potentially devastating to those who were still nurturing the infant. To put the quotations of Alexander and Jones into context, it may be necessary to recall the social–cultural climate of the late 1960s – the campus revolutions, the new liberal humanism and rejection of previous values. But also it had to be acknowledged (and it was) that there had been a lack of success in the application of "scientific" methods to design. Fundamental issues were also raised by Rittel and Webber (1973), who characterised design and planning problems as "wicked" problems, fundamentally unamenable to the techniques of science (and engineering), which dealt with "tame" problems.

Design methodology was temporarily saved, however, by Rittel's (1973) brilliant proposal of "generations" of methods. He suggested that the developments of the 1960s had been only "first-generation" methods (which naturally, with hindsight, seemed a bit simplistic, but none the less had been a necessary beginning) and that a new, second generation was beginning to emerge. This suggestion was brilliant because it let the new methodologists escape from their commitment to inadequate "first-generation" methods, and it opened a vista of an endless future of generation upon generation of new methods.

We might wonder what has happened to Rittel's theory of "generations". The first generation (of the 1960s) was based on the application of systematic, rational, "scientific" methods. The second generation (of the early 1970s) moved away from attempts to optimise and from the omnipotence of the designer (especially for "wicked problems"), towards recognition of satisfactory or appropriate solution-types (Simon (1969) had introduced the notion of "satisficing") and an "argumentative", participatory process in which designers are partners with the problem "owners" (clients, customers, users, the community). However, this approach tends to be more relevant to architecture and planning than engineering and industrial design, and meanwhile these latter fields were still developing their methodologies but in a different direction.

Engineering design methodology developed strongly in the 1980s; for example, in Europe, through ICED – the series of International Conferences on Engineering Design – and the work of the VDI (Verein Deutscher Ingenieure) in Germany. These developments were especially strong in England, Germany and Japan (Hongo and Nakajima, 1991), if not in the USA. (Although there may still have been only limited evidence of practical applications and results.) A series of books on engineering design methods and methodology began to appear. Just to mention some English-language ones, these included Hubka (1982), Pahl and Beitz (1984), French (1985), Cross (1989) and Pugh (1991).

It should also be acknowledged that in the USA there were some important conferences on design theory, and the National Science Foundation initiative on design theory and methods (perhaps in response to German and Japanese progress – like the earlier response to the first sputnik?) led to substantial growth in engineering design methodology in the late 1980s. ASME, the American Society of Mechanical Engineers, launched a series of conferences on Design Theory and Methodology (Stauffer, 1991).

So the development of "second generations" of design methodology in architecture and engineering appeared to diverge from each other in the 1970s and 1980s. Roozenburg and Cross (1991) have pointed out that these two fields have tended to diverge especially in their models of the design process, to the detriment of both. Perhaps a third generation of the 1990s might be based on a combination of the previous two; or, as in the model proposed by Cross (1989), on understanding the "commutative" (Archer 1979) nature of problem and solution in design. There was also a broader renewal of interest in design methodology in the late 1980s – especially in AI developments, where hope springs again for design automation and/or intelligent electronic design assistants.

A particularly significant development has been the emergence of new journals of design research, theory and methodology. Just to refer, again, to publications in English, we have had *Design Studies* (UK) since 1979, *Design Issues* (USA) since 1984, *Research in Engineering Design* (USA) since 1989, the *Journal of Engineering Design* (UK) since 1990 and the *Journal of Design Management* (USA) since 1990.

3. Relationships Between Design Methodology and Science

From the earliest days, design methodologists have sought to make distinctions between design and science, as reflected in the following quotations.

> Scientists try to identify the components of existing structures, designers try to shape the components of new structures. (Alexander 1964)

> The scientific method is a pattern of problem solving behaviour employed in finding out the nature of what exists, whereas the design method is a pattern of behaviour employed in inventing things ... which do not yet exist. Science is analytic; design is constructive. (Gregory 1966)

> The natural sciences are concerned with how things are ... design on the other hand is concerned with how things ought to be. (Simon 1969)

Glynn (1985) has pointed out that the above distinctions tend to be based on a positivistic (and possibly simplistic) view of the nature of science, and that scientists too, like designers, create their hypotheses and theories, and use these theories to guide their search for facts. Hillier *et al.* (1972) also criticized design methodologists for basing their ideas on outmoded concepts of scientific method and epistemology.

Cross *et al.* (1981) went so far as to suggest that the current epistemology of science is in some confusion and therefore is a most unreliable guide for an epistemology of design. This conclusion was challenged by Levy (1985), who suggested that transformations within the epistemology of science should be seen as active growth and development rather than simply chaos, and that it would be naïve to try to isolate design and technology from science and society.

However, there may still be a critical distinction to be made: method may be vital to science (where it validates the results) but not to design (where results do not have to be repeatable).

It is also clear that practitioners, whether in science or design, do not have to be methodologists. As Sir Frederick Bartlett pointed out, "The experimenter must be able to use specific methods rigorously, but he need not be in the least concerned with methodology as a body of general principles. Outstanding 'methodologists' have not themselves usually been successful experimenters." (Bartlett 1958) If "designer" is substituted for "experimenter", this observation also holds true in the context of design.

The Design Research Society's "Design: Science: Method" conference of 1980 (Jacques and Powell 1981) gave an opportunity to air many of these considerations. The general feeling from that conference was that it was time to move on from making simplistic comparisons and distinctions between science and design; that perhaps there was not so much for design to learn from science after all; and that perhaps science rather had something to learn from design. As Archer (1981) wrote in his paper for that conference, "Design, like science, is a way of looking at the world and imposing structure upon it." Both science and design, as Glynn (1985) pointed out, are essentially based on acts of perception, and "it is the epistemology of design that has inherited the task of developing the logic of creativity, hypothesis innovation or invention that has proved so elusive to the philosophers of science."

More informed views of both science and design now exist than they did in the 1960s. As Levy (1985) wrote, "Science is no longer perceived in terms of a single fixed methodology focused on a specific view of the world. It is more an expanded rationality for problem-identifying, -structuring and -solving activities." This makes scientific methodology sound indistinguishable from design methodology. Thus the simple dichotomies expressed in the 1960s are being replaced by a more complex recognition of the web of interdependencies between knowledge, action and reflection.

But in some places, old attitudes die hard! The editorial in the very first issue of this journal was clear about the journal's aim to change design from an art to a science: "For the field of design *to advance from art to science* (emphasis added) requires research. ..." (Dixon and Finger 1989)

Let us at least try to clarify some of the terminology that is used in discussing concepts such as "scientific design", "design science" and "the science of design".

3.1. Scientific Design

As already noted above, the origins of design methods lay in "scientific" methods, similar to decision theory and the methods of operational research. The originators of the "design methods movement" also realised that there had been a change from pre-industrial design to industrial design – and perhaps even to post-industrial design? The reasons advanced for developing new methods were often based on this assumption; modern, industrial design is too complex for intuitive methods.

The first half of this century had also seen the rapid growth of scientific underpinnings in many types of design – e.g. materials science, engineering science, building science, behavioural science. A relatively simple view of the design – science relationship is that, through this reliance of modern design upon scientific knowledge, through the application of scientific knowledge in practical tasks, design "makes science visible" (Willem 1990).

So we might suggest that "scientific design" refers to modern, industrialised design – as distinct from pre-industrial, craft-oriented design – based on scientific knowledge but utilising a mix of both intuitive and non-intuitive design methods.

3.2. Design Science

"Design science" was a term perhaps first used by Gregory (1966), in the context of the 1965 conference on "The Design Method". Others, too, have the development of a "design science" as their aim; for example, the originators of the ICED conferences, the "Workshop Design Konstruction" (WDK), are also the International Society for Design Science. The concern to develop a design science has led to attempts to formulate *the* design method – a single rationalised method, based on formal languages and theories. We have even had presented the concept of "creativity as an exact science" (Altshuller 1984).

But a desire to "scientise" design can be traced back to ideas in the modern movement of design. The designer Theo van Doesburg wrote in the 1920s: "Our epoch is hostile to every subjective speculation in art, science, technology, etc. The new spirit, which already governs almost all modern life, is opposed to animal spontaneity, to nature's domination, to artistic flummery. In order to construct a new object we need a method, that is to say, an objective system." (van Doesberg 1923) And a little later, the architect Le Corbusier wrote: "The use of the house consists of a regular sequence of definite functions. The regular sequence of these functions is a traffic phenomenon. To render that traffic exact, economical and rapid is the key effort of modern architectural science." (Le Corbusier 1929)

Hansen (1974), quoted by Hubka and Eder (1987), has stated the aim of design science as being to "recognise laws of design and its activities, and develop rules." This would seem to be design science constituted simply as "systematic design" – the procedures of designing organised in a systematic way. Hubka and Eder regard this as a narrower interpretation of design science than their own: "Design science comprises a collection (a system) of logically connected knowledge in the area of design, and contains concepts of technical information and of design methodology.... Design science addresses the problem of determining and categorising all regular phenomena of the systems to be designed, and of the design process. Design science is also concerned with deriving from the applied knowledge of the natural sciences appropriate information in a form suitable for the designer's use."

This definition extends beyond "scientific design", in including systematic knowledge of design process and methodology as well as scientific/technological underpinnings of design of artefacts. For Hubka and Eder the important constituents of design science are: (1) applied knowledge from natural and human sciences; (2) theory of technical systems; (3) theory of design processes; (4) design methodology.

Andreasen (1991) points to two important areas of theory in design science that are delineated by Hubka (for mechanical engineering): theory of the design process (general procedures, methods, tools) and theory of machine systems (classification, modelling, etc. of technical systems). This helps to define design science as including both process and product knowledge and theory.

So we might conclude that "design science" refers to an explicitly organised, rational and wholly systematic approach to design: not just the utilisation of scientific knowledge of artefacts, but design also in some sense as a scientific activity itself.

3.3. Science of Design

There is some confusion between concepts of "design science" and of a "science of design", since "science of design" seems to imply (or for some people has an aim of) the development of a "Design Science". For example, we have praxeology, "the science of effective action", and in *The Sciences of the Artificial* Simon (1969) defined "the science of design" as "a body of intellectually tough, analytic, partly formalisable, partly empirical, teachable doctrine about the design process".

This view is controversial. As Grant (1979), wrote: "Most opinion among design methodologists and among designers holds that the act of designing itself is not and will not ever be a scientific activity; that is, that designing is itself a non-scientific or a-scientific activity." However, Grant also made it clear that "the study of designing may be a scientific activity; that is, design as an activity may be the subject of scientific investigation."

A similar view of "the science of design" has also been clearly stated by Gasparski (1990): "The science of design (should be) understood, just like the science of science, as a federation of subdisciplines having design as the subject of their cognitive interests."

In this latter view, therefore, the science of design is the *study* of design – something similar to what I have elsewhere defined as "design methodology"; the study of the principles, practices and procedures of design. For me, design methodology "includes the study of how designers work and think, the establishment of appropriate structures for the design process, the development and application of new design methods, techniques and procedures, and reflection on the nature and extent of design knowledge and its application to design problems". (N. Cross 1984)

So let us conclude here that the "science of design" refers to that body of work which attempts to improve our understanding of design through "scientific" (i.e. systematic, reliable) methods of investigation.

4. Progress in Design Methodology

I conclude here with a brief review of developments in the science of design (design methodology) over the last decade. I will use categories of work similar to those I used in *Developments in Design Methodology* (N. Cross 1984), which covered the period 1962 – 82.

4.1. The Development of Design Methods: Origination and Application of Systematic Methods

In this category, the last decade has been notable for the development of product quality assurance methods, such as Taguchi methods (Ross 1988) and quality function deployment (Hauser and Clausing 1988).

There has also been significant new work in design automation, using expert systems and other artificial intelligence techniques. A new series of conferences on AI and design has been established, where this work is reported (Gero 1991).

4.2. The Management of Design Process: Models and Strategies for Executing Design Projects

We have had a new generation of systematic models of the design process, particularly in engineering design, and particularly from Germany (Hubka (1982), Pahl and Beitz (1984), Verein Deutscher Ingenieure (VDI) (1987)). We have also seen the emergence of "concurrent" models of product planning and development (Andreasen 1991, Pugh 1991).

In architecture and planning there has been development of the "argumentative" process models (e.g. McCall 1986).

4.3. The Structure of Design Problems: Theoretical Analysis of the Nature of Design Problems

There has been significant new work on problem "types", for example by Schön (1988) and by Oxman (1990). In this category we might also include the new work on formal languages and grammars of design (Stiny 1980, Flemming 1987).

4.4. The Nature of Design Activity: Empirical Observations of Design Practice

There have been many more protocol and case studies made in this period. Examples include Schön (1984), Rowe (1987), Davies and Talbot (1987), Wallace and Hales (1987), Stauffer *et al.* (1987), Eckersley (1988), Waldron and Waldron (1988). A recent conference in the Netherlands on "research in design thinking" brought together several related approaches and recent new work (Cross *et al.* 1992).

4.5. The Philosophy of Design Method: Philosophical Analysis and Reflection on Design Activity

Some of the comparative discussions of design and science have already been referred to earlier in this paper (Levy 1985, Glynn 1985). There have been several new studies in the epistemology of design (Buchanan 1989, Zeng and Cheng 1991, Roozenburg 1992), and we should also include here work in the praxeology of design (Gasparski 1990).

Some of us have also been theory-building around the concept of "designerly" ways of thinking and acting (A. Cross 1984, 1986, Tovey 1986, N. Cross 1990), although some aspects of this work have been challenged by Coyne and Snodgrass (1991).

5. Conclusion

Design methods emerged in the early 1960s as part of a desire (with a history going back to the 1920s) to make design somehow more "scientific". The subject became adopted as an academic area of study known as design methodology.

For some people, design methodology appeared to have died in the 1970s; however, we can now see that it survived, and that there has been some particularly strong and healthy growth in the 1980s, especially in the engineering and product design fields. There is still some confusion and controversy over the use of terms such as design science, but I hope that the discussion here has helped to clarify this.

Design methodology has become a much more mature academic field, but still suffers from a lack of confidence in it by design practitioners, and it has had little (acknowledged) practical application.

Note

This is a revised version of a paper, "A History of Design Methodology", prepared for a conference on "Design Methodology and Relationships with Science" (de Vries *et al.* 1993).

References

Alexander, C. (1964) *Notes on the Synthesis of Form.* Cambridge, Mass., Harvard University Press

Alexander, C. (1971) The State of the Art in Design Methods. *DMG Newsletter* **5**(3): 3 – 7

Altshuller, G.S. (1984) *Creativity as an Exact Science.* London, Gordon & Breach

Andreasen, M.M. (1991) Design Methodology. *Journal of Engineering Design* **2**(4): 321 – 335

Archer, L.B. (1965) *Systematic Method for Designers.* London, Design Council

Archer, L.B. (1979) Whatever Became of Design Methodology? *Design Studies* **1**(1): 17 – 18

Archer, L.B. (1981) A View of the Nature of Design Research. In R. Jacques and J. Powell (ed.) *Design: Science: Method.* Guildford Westbury House

Asimow, M. (1962) *Introduction to Design.* Englewood Cliffs, N.J., Prentice-Hall

Bartlett, F.C. (1958) *Thinking: an Experimental and Social Study.* London, George Allen & Unwin

Broadbent, G. (1973) *Design in Architecture.* Chichester, John Wiley

Broadbent, G. (1979) The Development of Design Methods. *Design Methods and Theories* **13**(1): 41 – 45

Broadbent, G.; A. Ward (ed.) (1969) *Design Methods in Architecture.* London: Lund Humphries

Buchanan, R. (1989) Declaration by Design: rhetoric, argument and demonstration in design practice. In V. Margolin (ed.) *Design Discourse.* Chicago, Ill., University of Chicago Press

Coyne, R.; A. Snodgrass (1991) Is Designing Mysterious? Challenging the Dual Knowledge Thesis. *Design Studies* **12**(3): 124 – 131

Cross, A. (1984) Towards an Understanding of the Intrinsic Values of Design Education, *Design Studies* **5**(1): 31 – 39

Cross, A. (1986) Design Intelligence: the use of codes and language systems in design, *Design Studies* **7**(1): 14 – 19

Cross, N. (1980) The Recent History of Post-industrial Design Methods, in N. Hamilton (ed.) *Design and Industry.* London, Design Council

Cross, N. (ed.) (1984) *Developments in Design Methodology.* Chichester, John Wiley.

Cross, N. (1989) *Engineering Design Methods.* Chichester, John Wiley

Cross, N. (1990) The Nature and Nurture of Design Ability, *Design Studies* **11**(3): 127 – 140

Cross, N.; Dorst, K.; Roozenburg, N. (eds) (1992) *Research in Design Thinking.* Delft, Netherlands, Delft University Press

Cross, N.; Naughton, J. *et al.* (1981) Design Method and Scientific Method. In R. Jacques and J. Powell (ed.) *Design: Science: Method.* Guildford, Westbury House

Davies, R.; Talbot, R.J. (1987) Experiencing Ideas: Identity, Insight and the Imago, *Design Studies* **8**(1): 17 – 25

de Vries, M.J.; Cross, N.; Grant, D.P. (ed.) (1993) *Design Methodology and Relationships with Science.* Dordrecht, Kluwer

Dixon, J.R.; Finger S. (1989) Editorial, *Research in Engineering Design* **1**(1): 1

Eckersley, M. (1988) The Form of Design Processes: a Protocol Analysis Study, *Design Studies* **9**(2): 86 – 94

Evans, B.J.; Powell, J.; Talbot, R. (eds) (1982) *Changing Design.* Chichester, John Wiley

Finger, S.; Dixon, J.R. (1989) A Review of Research in Mechanical Engineering Design. *Research in Engineering Design* **1**(1): 51 – 67

Flemming, U. (1987) The Role of Shape Grammars in the Analysis and Creation of Designs. In Y.E. Kalay (ed.) *Computability of Design.* New York, John Wiley

French, M.J. (1985) *Conceptual Design for Engineers.* London, Design Council

Gasparski, W. (1990) Editorial: Contributions to Design Science. *Design Methods and Theories* **24**(2): 1186 – 1194

Gasparski, W. (1990) On the General Theory (Praxeology) of Design. *Design Methods and Theories* **24**(2): 1195 – 1215

Gero, J. (ed.) (1991) *Artificial Intelligence in Design '91.* Oxford, Butterworth-Heinemann

Glynn, S. (1985) Science and Perception as Design. *Design Studies* **6**(3): 122 – 133

Gordon, W.J.J. (1961) *Synectics.* New York, Harper & Row

Grant, D.P. (1979) Design Methodology and Design Methods. *Design Methods and Theories* **13**(1): 46 – 47

Gregory, S.A. (ed.) (1966) *The Design Method.* London: Butterworth

Gregory, S.A. (1966) Design Science. In S.A. Gregory (ed.). *The Design Method.* London, Butterworth

Hall, A.D. (1962) *A Methodology for Systems Engineering.* Princeton, N.J., Van Nostrand

Hansen, F. (1974) *Konstruktionswissenschaft.* Munich, Carl Hanser

Hauser, J.R.; Clausing, D. (1988) The House of Quality. *Harvard Business Review* (May/June): 63 – 73

Hilier, B.; Musgrove, J. *et al.* (1972) Knowledge and Design. In W.J. Mitchell (ed.) *Environmental Design: Research and Practice.* Los Angeles, Calif., University of California Press

Hongo, K.: and Nakajima, N. (1991) Relevant Features of the Decade 1981 – 1991 for Theories of Design in Japan. *Design Studies* **12**(4): 209 – 214

Hubka, V. (1982) *Principles of Engineering Design.* Guildford, Butterworth

Hubka, V.; Eder W.E. (1987) A Scientific Approach to Engineering Design. *Design Studies* **8**(3): 123 – 137

Jacques, R.; Powell, J. (eds) (1981) *Design: Science: Method.* Guildford, Westbury House

Jones, J.C. (1970) *Design Methods.* Chichester, John Wiley

Jones, J.C. (1977) How My Thoughts About Design Methods Have Changed During the Years. *Design Methods and Theories* **11**(1): 48 – 62

Jones, J.C.; Thornley, D.G. (eds) (1963) *Conference on Design Methods.* Oxford, Pergamon Press

Le Corbusier (1929) CIAM 2nd Congress, Frankfurt

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

McCall, R. (1986) Issue-Serve Systems: a Descriptive Theory of Design. *Design Methods and Theories* **20**(3): 443 – 458

Moore, G.T. (ed.) (1970) *Emerging Methods in Environmental Design and Planning,* Cambridge, Mass.: MIT Press

Osborn, A.F. (1963) *Applied Imagination – Principles and Procedures of Creative Thinking.* New York: Charles Scribner's Sons

Oxman, R. (1990) Prior Knowledge in Design. *Design Studies* **11**(1): 17 – 28

Pahl, G.; Beitz, W. (1984) *Engineering Design.* London, Design Council

Pugh, S. (1991) *Total Design: Integrated Methods for Successful Product Engineering.* Wokingham, Addison-Wesley

Rittel, H. (1973) The State of the Art in Design Methods. *Design Research and Methods (Design Methods and Theories)* **7**(2): 143 – 147

Rittel, H. and M. Webber (1973) Dilemmas in a General Theory of Planning. *Policy Sciences* **4**: 155 – 169

Roozenburg, N. (1992) On the Logic of Innovative Design, in N. Cross, C. Dorst and N. Roozenburg (eds). *Research in Design Thinking.* Delft, Delft University Press

Roozenburg, N.: Cross, N. (1991) Models of the Design Process: Integrating Across the Disciplines. *Design Studies* **12**(4): 215 – 220

Ross, P.J. (1988) *Taguchi Techniques for Quality Engineering.* New York, McGraw-Hill

Rowe, P. (1987) *Design Thinking.* Cambridge, Mass., MIT Press

Schön, D. (1984) Problems, Frames and Perspectives on Designing. *Design Studies* 5(3): 132 – 136

Schön, D. A. (1988) Designing: Rules, Types and Worlds. *Design Studies* 9(3): 181 – 190

Simon, H.A. (1969) *The Sciences of the Artificial.* Cambridge, Mass., MIT Press

Spillers, W.R. (ed.) (1974) *Basic Questions of Design Theory.* Amsterdam – New York, North-Holland – Elsevier

Stauffer, L.; Ullman, D. *et al.* (1987) Protocol Analysis of Mechanical Engineering Design. In W.E. Eder (ed.), Proceedings of International Conference on Engineering Design, Boston, New York, American Society of Mechanical Engineers

Stauffer, L.A. (ed.) (1991) *Design Theory and Methodology – DTM '91*, New York: American Society of Mechanical Engineers

Stiny, G. (1980) Introduction to Shape and Shape Grammars. *Environment and Planning B* 7: 343 – 351

Tovey, M. (1986) Thinking Styles and Modelling Systems. *Design Studies* 7(1): 20 – 30

van Doesberg, T. (1923) Towards a Collective Construction, *De Stijl*: quoted in Naylor, G. (1968) *The Bauhaus.* London, Studio Vista

Verein Deutscher Ingenieure (VDI) (1987) *Systematic Approach to the Design of Technical Systems and Products: Guideline VDI 2221*, Berlin, Beuth Verlag

Waldron, M.B.; Waldron, K.J. (1988) A Time Sequence Study of a Complex Mechanical System Design. *Design Studies* 9(2): 95 – 106

Wallace, K.; Hales, C. (1987) Detailed Analysis of an Engineering Design Project. In W.E. Eder (ed.) *Proceedings of the International Conference on Engineering Design, Boston.* New York, American Society of Mechanical Engineers

Willem, R.A. (1990) Design and Science. *Design Studies* **11**(1): 43 – 47

Zeng, Y.; Cheng, G.D. (1991) On the Logic of Design. *Design Studies* **12**(3): 137 – 141

Zwicky, F. (1948) The Morphological Method of Analysis and Construction. In *Studies and Essays*, New York: Interscience

Science, technology and design

R Levy

Faculté de l'Aménagement, Université de Montréal, Montréal, Québec, Canada

It is argued, through a historical review of human inquiry systems, that contemporary science is not in epistemological chaos as suggested by some authors. It should not, therefore, be overlooked as a source of constitutive knowledge relevant to the activity of design. It is suggested that the present state of scientific epistemology embodies the principles not only of science but also those of technology, industry and society, and that design should be viewed as an activity in conjunction with this paradigmatic organization and not solely as a technological activity.

Keywords: epistemological chaos, mind–body, inquiring systems, techne, epistemological incertitude

Stemming from the knowledge, both implicit and explicit, that there is an autoorganizational interaction between science and technology, technology and industry, industry and society and society and science[1], three interrelated statements are derived which set the tone for this paper and form some of the preconditions for the arguments to follow. Firstly, it is believed unwise, from a methodological point of view, to isolate technological praxis from any or all systems of inquiry. Secondly, it is shortsighted to view *techne* as being, in some way or another, separate from scientific epistemology. And thirdly, one should avoid falling into the reductivistic trap of believing that technology is an independent ontological 'object', and as such an object of specific cultural value, i.e. a cult object.

Cross *et al.*[2], in a paper on design method and scientific method, argued that design should be considered more a technological activity than a scientific activity. Evidence cited by these authors for this conjecture was derived from an historical perspective of the received view of science. They concluded that there is a fundamental difference between the act of designing and the activity of science, between design method and scientific method and between design values and scientific values. In addition they proposed the notion that, owing to the numerous and diverse schools of scientific thought that presently hold centre stage, contemporary science is therefore in epistemological chaos. As a result of this perceived chaos, they argued that any inferences to be drawn between design and science must wait for science to achieve epistemological coherence and historical validity. Consequently design, in the meantime, should be projected through the lens of the more practical and stable technological model of human action.

The objective of the present paper, is to question the assertions made by Cross *et al.*, and to make other kinds of conjectures. In the first part it will be argued, by way of an historical analysis of the major inquiry systems developed this century, that scientific epistemology has never been coherent nor historically valid. It is suggested that rather than waiting (and we may have an infintely long wait on our hands) for science to achieve epistemological coherence and historical validity, it may be more fruitful to participate in its multiple, parallel and continuous structural transformations, and to use the insights gained from each transformation to help sharpen the epistemological coherence and historical validity of design itself.

In the second part, some aspects of the contemporary movements in the philosophy of technology are discussed in which scientific inquiry systems and technological problem-solving systems are welded together in an epistemological unification of science, technology and society. In this context, methodology is reflected in terms

 0142–694X/85/020066–07 $03.00

of the intentionalities of both science and technology and defined more in relation to concepts such as strategy, initiative, invention and art[3]. Through this general perspective of method and interplay of contemporary inquiry systems, the claim that design be based on an implicit model of designing and viewed solely as a technological activity will be questioned. In its place, it is proposed that design be viewed in more expansive epistemological terms, based on an integrated model of human inquiry systems and a unification of science, technology and society.

SYSTEMS OF INQUIRY

Over the past century (and indeed throughout history) epistemology has manifested a continuous process of structural transformation. Varela's[4] definition of ontogeny is used to describe this process, i.e. the 'history of the structural transformation of a unity'. If we look at the history of the structural transformation of scientific epistemology and at its present state of affairs, we will find many instances of change, rupture, divergence and pluralism. To view these structural transformations and the present divergent viewpoints of science as akin to a state of chaos[2], is equivalent to thinking of the concept of change as being pejorative. On the other hand, it is conceded that it is possible to see chaos in contemporary scientific epistemology, if science is viewed only within the framework of logical or empirical positivism. However, it is important to enlarge this perspective of science to include other parallel epistemologies which have had their own ontogenic developments and which have provided much of the material for the widespread criticism of the dominant hypothetico-deductive approach.

One need merely cite Kuhn[5], Lakatos[6], Popper[7], Polanyi[8], Toulmin[9] and Feyerabend[10], to provide ample subjective evidence that the history of science is a continuous and complex process of revolutionary, evolutionary or structural transformation of its epistemological underpinnings. In this context, it is suggested that any analysis of the received view of science must include, at least, the parallel epistomologies of the general theory of systems, structuralism, human action inquiry systems, and the existential-phenomenological and hermeneutic approaches. It is the interaction of these various inquiry systems and their respective epistemological frameworks which helps to construct a vision of science continually undergoing structural and value transformations.

The following historical review takes this kind of approach, based on an excellent analysis of the subject by Polkinghorne[11], and identifies some of the essential issues in the ontogeny of the major inquiry systems over the past 60 years. Neopositivism is used as the benchmark for the discussion, and follows the five phases of its development beginning in the 1920s with the Vienna Circle. This provides a sequence of events which helps to contextualize and situate the other parallel inquiry systems.

The Vienna Circle was essentially a renewal of positivist thought in reaction to the strong antipositivist trends at the end of the nineteenth century. the Vienna Circle, consisting of a group of philosophers, mathematicians and scientists (Schlick, Carnap, Neurath, Feigl and Gödel), met on a regular basis essentially to discuss how to eliminate all metaphysical and speculative errors from knowledge. The outcome of these discussions was the reformulation of positivism in strict accordance with the epistemological chain stemming back through Bacon, Galileo, Newton, Hobbes, Comte, Mill and Mach. In particular, the neo-positivism of the Vienna Circle defined a formalistic empiricism based on Mach's sensationism and Russel and Whitehead's formal system of logic. This neopositivistic scientific epistemology became known as logical positivism or the received view of science. However, because of the restrictive nature of the phenomenalistic language used in this inquiry process and because this process lent itself to excesses in solipsism (private experience of individuals) the movement was forced, by its own admission, to make structural changes. Thus, in 1933, a second stage of the Vienna Circle's vision of scientific inquiry was formulated.

This change was precipitated by a new doctrine, called physicalism, which replaced that of phenomenalism. The adjustment in the epistemology was done in recognition of the criticism that statements referring to scientific inquiry should be related to objects and relationships and not solely to private experience. In distinction to the first stage, where knowledge (*epistémé*) was seen to be in the form of words and statements and their relationship to the world (i.e. a purely axiomatic deductive inquiry system), the second stage opened up scientific inquiry to intersubjective verification and provided the mechanism to use the everyday working language of scientists. This second stage became known as logical empiricism and with it the whole modality of classical empircal thought was changed. It removed the insistence that knowledge (what is known with certainty) can only be in the form of words and statements, and admitted into the terms of reference all statements that were ultimately linked to some observable phenomenon.

This received view of science of the 1930s, which declared that science must limit its statements only to descriptions of regularities that are observable, was expanded, in the early 1940s, to include statements that referred to nonobservable entities as well. Thus the second phase in the development of modern scientific epistemology was born and acceptance of axiom-based networks of universal theoretical statements was sanctioned. This approach held sway into the 1950s where all 'sciences' were limited to statements deductively linked to descriptions of direct observation and a system of axioms. Statements that were inferred by induction were recognized only as approximations which had not yet attained the certainty of deductively valid statements. In essence, the goal of science (logical empiricism) was to develop comprehensive theories which then formed the presuppositions for explaining phenomena as a set of deductively connected empirical generalizations. It was in fact this formulation of scientific inquiry which began the storm of criticism which continues until today.

The criticism of the received view of science in the 1950s was sparked by a number of issues. In the main, the question of induction was not adequately integrated into the narrow modalities of the theoretical network concept of science. As such, logical positivism (and the subsequent logical empiricism), could not provide a basis for the unification of science grounded on theoretical statements of the sense experience. In addition, criticism was levelled at this dominant paradigm from a historical perspective. Philosophers such as Toulmin[9] and Polanyi[8] called for a view of science based on a more pragmatic problem-orientated approach.

The debate in the 1950s was therefore centred on whether or not positivistic science as the dominant inquiry system could continue to uphold the principle of certainty for scientific knowledge. The proposed alternative approach, a more pragmatic science, viewed science as a human activity in which understanding the universe was guided by a process of conjecture and refutation[6,7].

Together with these structural changes that were brought to bear on scientific epistemology, other parallel developments were occurring which had an equally dramatic effect on the philosophy of science. The influence of the systemic system of inquiry, the human action system of inquiry and the existential–phenomenological–hermeneutic inquiry system cannot be underestimated.

The importance of the systemic approach or the movement dedicated to the development of a general theory of systems[12–15], stemming from biology and engineering, relates to the fact that it uses a form of logic that differs from deductive or inductive thinking. Simply stated, it represents a more open form of reflection and calls for an expanded notion of reason. It was, and is, an approach which focuses on the structure, organization, and interaction of a particular phenomenon. The systemic inquiry activity involves modelling phenomena as totalities and in so doing integrating into a single system modes of epistemological abstraction, methodological organization, ontological substantiality and teleological forms of purpose[16]. The interdisciplinary movement which gravitated around the development of a general theory of systems had considerable influence on the ongoing critical reflection of scientific epistemology and on the development of the subsequent third, fourth and fifth phases of scientific philosophy from 1950 to the present time.

In addition, it is important to include in this discussion the ontogenic developments in the field of the humanities and especially the area of human action. Human science methodology has traditionally looked towards both the humanities and the physical sciences for guidance and approaches in the study of social and individual phenomena. This last century has also seen rapid structural transformations in this field. In the 1920s and 30s the format of deductive-nomological explanations[17] was the dominant school of thought where human action was explained by causal laws[18,19]. In the 1950s, criticism of this approach was very pronounced, fed by studies derived from Wittgenstein's later work, in which language was used to describe the rules of social interaction[11]. A later and more recent development has been the refinement of formal logic systems including decision theory, deontic logic and nonformal approaches of practical reasoning used by individuals and groups in making decisions. This latter approach is important since it has influenced philosophers of science who see science as '. . . an endeavour which engages in practical reasoning rather than in a formal system of demonstrative logic'[11].

Another sector which has influenced the structural transformation of science over the past decades is the Continental movement of phenomenology and hermeneutics. The approaches link such people as Husserl, Heidegger, Merleau-Ponty, Ricoeur and Habermas. On the one hand, phenomenology takes a descriptive approach to inquiry as distinct from the formal method of scientific empiricism. On the other hand, hermeneutics takes an interpretive approach. Thus in terms of scientific inquiry, the formal logical empiricism of the physical and social sciences has been in constant juxtaposition with the nonformal descriptive and interpretive approaches since the turn of the century. With its roots in Husserl's phenomenology and descriptions of the structures of consciousness, contemporary existential-phenomenology aims to understand the basic structures of human existence. Hermeneutics and its interpretive approach is essentially a supplement to phenomenology by seeking to understand social and individual actions. Hermeneutics has also undergone considerable structural change during this century and together with phenomenology has developed a system of inquiry which provides refined and discriminating explanations of human organization and expression.

The development of the third, fourth and fifth phases of scientific epistemology should be seen in the context of these various inquiry systems that existed in parallel over the past sixty years. Considering this epistemological 'soup' of the fifties and early sixties, it is no wonder that various trends, or schools of thought, emerged in quick succession. This changing nature of science should not be viewed as a form of epistemological chaos but more as a process of healthy epistemological ontogeny.

The criticism of the received view of the 1950s produced the third phase or the pragmatic movement which essentially addressed the problem of induction. It was argued that scientific observations cannot only be connected by correspondence rules to one theoretical statement, as in deductive reasoning, but rather are connected to a plethora of theories and assumptions[20]. In this way scientific epistemology changed its unique descendent hypothetico-deductive formalization to one that also included ascendent inductive intuition[21]. In other words, pure sense data and formal logic were married to provide a more amenable, malleable pragmatic science.

In a parallel movement, the fourth phase was developed which described science as expressions of various world outlooks derived from the notion of the 'weltanschauungen' (and is strongly identified with the systemic inquiry system). The proponents of the fourth phase charged that neither the pure senses nor formal logic could provide a knowledge absolute. From the perspective of the 'weltanschauungen' the characteristics of an individual's knowledge are biased by their historically given cognitive framework. The outcome of this approach was a perception of scientific inquiry in which all observations, meanings and facts are believed to be ultimately theory-laden[10]. In this way, the fourth phase served as a corrective to the traditional logical positivism in pointing out the sociocultural influences on scientific inquiry. As such, knowledge can be acquired through the deployment of various perspectives, all perspectives being considered equal. It then follows that *epistémé* (knowledge of things as they really are) is ultimately a futile objective, as we live in a universe of multiple and parallel truths.

This brings us to the fifth and present phase, which Suppe[22] calls 'historical realism'. This phase essentially interconnects the third and fourth phases by accepting the utility of a science based on pragmatic reasoning and the importance of augmenting the pragmatism by including the historical conditions of scientific inquiry. Thus the fifth phase is characterized by the belief that deductive/inductive scientific reasoning coupled with worldview concepts can provide reliable information about our universe. It is important to note that this approach does not only seek to explain what things are but also seeks to explain how things are[23–26].

> What is showing through in this new examination is that the central and most characteristic activity of science is the use of various patterns of reason rather than simple logic. It is through the process of reasoning that hypotheses are suggested and developed, and it is through the process of reasoning that knowledge claims are evaluated. Among contemporary philosophers of science it is believed that the core of scientific activity is to be found in these patterns of reasoning, and come into play in postulating and evaluating hypotheses. The patterns of reasoning yield conclusions that go beyond the logical entailments of deductive logic. These patterns are 'different systems of rules of the scientific game' and deductive logic is merely one pattern of rationality (p 117)[11].

It must be admitted that logical positivism did instil in society a particular perception as the ultimate form of inquiry into truth. That this perspective has undergone substantial change and is subject to criticism does not mean that science has lost its bearings. Science today is not in epistemological chaos. The original claim of the logical positivists was only one way of viewing the world and now we know and accept that it is not the only one. Science is no longer perceived in terms of a single fixed methodology focussed on a specific view of the world. It is more an expanded rationality for problem-identifying, structuring and solving activities. Science may therefore be characterized as a means to solve conceptual and practical problems through a variety of reasoning or inquiry systems.

It is suggested that this expanded perception of 'scientific' inquiry systems provides a more even view of the current debate of the relationships among science, technology and design. Furthermore, it is almost axiomatic, in the light of this expanded view, that 'knowing that' and 'knowing how' is the domain of both science and technology, and that science and technology should not be considered as separate entities. In addition, a philosophy of technology would be nonsense if it were not integrated with the philosophy of science, and in the same way, an understanding of the activity of design would be nonsense if it were not embraced by a scientific/technological epistemology.

DESIGN AS SCIENTIFIC AND TECHNOLOGICAL PRAXIS

Whilst one may agree that philosophy of technology is not a particularly well developed sphere of study, it is perhaps possible to say that, owing to the current movement of scientific epistemology towards a more expansive form of historical realism, philosophies of technology will begin to emerge in the very near future[27–29]. It can also be said, without fear of criticism, that science is no longer viewed as being above or superior to technology, and it is in this spirit that we should discuss the future roles of science and technology.

Let us begin this discussion by returning to the paper by Cross *et al*[2]. These authors go to some length to explain that, in their view, designing is not the same kind of activity as doing science, and as a consequence, design should only be viewed within a technological framework. This is, however, after a little reflection, an argument based on too simplistic a view of science and technology. To label science as being only 'knowing that' (i.e. explicit knowledge leading to 'organized rules of conventional wisdom' and competence) and technology as being only 'knowing how' (i.e. tacit knowledge leading to 'standards of performance' and quality) implies a particularly fractionated view of both science and technology. This kind of 'decoupage' is the direct result of expressing scientific epistemology in the form of the received view of science and ignoring, to some extent, the spectrum of scientific inquiry systems and the structural changes that have occurred in scientific epistemology this century, especially over the past thirty years. Consequently, the advice given by these authors, i.e. to put science to one side because it is perceived to be in epistemological chaos and in its place embrace technology as the sole harbinger of the design template, is considered the result of a hasty interpretation of the contemporary scientific/technological paradigm.

If an 'implicit model of designing' is to be accepted[2], which manifests mainly what is called 'craft knowledge' or 'tacit knowledge', then we may fall into the trap of deifying, at the expense of all others, what Von Gottle-Ottlilienfeld[30] called the intuitive sense of technology or the 'art of the right way to end'. One simply knows how to do things without explanation. That's all very well, but what of the nonintuitive side of technology, the side dealing with the procedures, the organization, and the instruments of human activity? What of the critical reflective modalities which are required to understand technological *epistémé* as a human activity; what of our understanding of technology as a cohort of science, industry and society; what of technology as one of the motors of culture; as artistic expression; as an indicator of epistemological incertitude? How do we begin to explain these issues if we lock the design process into the single mould of intuitive 'know-how'?

To being responding to some of these questions consider the following definition of technology by Tucher as cited by Rapp[28]:

> Technology is the general term for all objects, procedures and systems, which on the basis of creative construction are produced for the fulfilment of individual and social needs; which through defined functions serve certain purposes; and in their totality change the world (p 35).

The definition is interesting in that it involves not only epistemological, ontological and methodological notions, but in addition also the teleological issues of purposefulness and change. Rapp[28] also reminds us that every human action that is 'consciously and purposefully executed follows a methodological pattern' and as such all individual and social purposeful action is technological. Purposefulness or intentionality is thus a key concept in human action and fundamental to our understanding of technological and scientific activities. Tucher's definition also highlights another important concept. Technology is both material and non-material. Non-material technology i.e. procedures, would include, for example, the construction of legal systems, political plans, medical treatment and the design of therapy programmes. Thus to obtain an expanded insight into what technology means, it is imperative that we critically study the concepts, methodologies, implicit epistemologies, ontologies and teleologies of technological praxis[16,29] and not simply ground our understanding on the unidimensional notion of 'know-how'.

Unidimensional representations of technology are a direct result of the nomological view of science. In this context, for example, we often find technology cast in the role of 'practice' in the traditional 'theory–practice' duality. Ihde[29] makes an insightful interpretation of this duality by equating it with the Cartesian 'mind–body' distinction. Theory is usually perceived as the product of mind, while practice is usually related to the product of body. We all know that the 'mind–body' duality has largely, over the past decades, been leached from scientific and technological epistemology and replaced with what can loosely be called the identity theory. It is, therefore, time to treat the science–technology analogy in the same way, and promote a monistic system without distinctions. We thus join the concepts of Morin[1] in which science, technology, industry and society are an inseparable entity.

If we maintain the traditional view of technology, i.e. as equivalent to 'practice', there is also the tendency to equate technology with the concept of 'how things work' or 'making things work' and confine these activities to the special domain of engineers, inventors, designers, architects, doctors, etc. These concepts match the intuitive notion of 'know-how' which is used to illustrate the implicit model of design. What is lost, however, in the emphasis on knowing how to do things and making things, is the equally important concept of how we use things; how we use technology. Can it be that this notion has not been viewed by technologists as important because it entails an interpretation of the 'meaning' of technology? In this regard one should not confuse 'meaning of technology' with what some call 'technology assessment' or 'impact studies'. These are *post-facto* understandings of technology. We end up studying the symptoms of technology and not its origins or its meaning.

To reach for the meaning of technology is to begin to understand how and why technology is part of the fabric of human activities. Once technologies are made and used, alterations in the patterns of human activity or forms of life take place[27]. To understand the meaning of technology is to understand the significance of technological praxis. To understand the 'use' of technology is to be able to identify what kinds of ideas (political, cultural, economic, psychological, etc.) are embodied in technological designs[27]. Knowing what the epistemological, methodological, ontological and teleological embodiments of technological praxis are will allow the process of decision making, at all levels of society, to be more clearly defined. In this way, the criteria of analysis, interpretation and judgement of a particular technological template on forms of life would move far beyond the traditional criteria of evaluation based simply on the practical notions of precision, objectivity, standardization, efficiency, utility, economy, convenience and competitiveness[27,31].

It would seem rather ironic then, in terms of the expanded view of science and technology, if the design disciplines should think of abandoning their ties to science. By all means, think of design as technological praxis but take care not to fall back into a mode of explanation dominated by an implicit or mystical process of 'the art of the right way to end'. This kind of unidimensionalism can only lead to a severe lack of knowledge-constitutive interests[32]. As an alternative, we should rather begin to think of increasing our sophistication of knowledge of the human inquiry paradigm to improve our overall understanding of the values of the design activity and the integration of science, technology, industry and society.

CONCLUSION

The ontogenic characteristics of scientific inquiry over the last sixty years have undergone five distinct Kuhnian-like tranformations. Science has moved from formal theories based on syntactic logic, through formal and nonformal semantic approaches, to arrive at the contemporary structural formulations of human inquiry. This structural transformation of science has created an epistemological framework which allows alternative contexts for treating and understanding the world. In addition, strong ideological transformations are implicit in these changes. We are witnessing, today, shifts in the fundamental belief structures of Western society. In this sense we are not looking at 'simple' changes in the methodological practice of scientific inquiry, but rather we are looking at complex and fundamental paradigmatic transformations. As a result, the philosophy of technology, a field long neglected by mainstream philosophy, is beginning to emerge, not as a separate philosophy but one intimately tied to science and society. These changes in ideology and approach will be reflected in the development of a new 'cadre epistémique' and a new 'paradigme epistémique'[33]. This new epistemic paradigm will address the enfolded[34] view of science, technology and society.

The traditional science–technology distinction stemming from the mind–body analogy is being replaced by a single integrated entity. Material and nonmaterial technology is viewed as the accepted template of social evolution. The process of designing, the praxis of the enfolded science/technology paradigm is fundamental to our understanding of new forms of human activity and as such is influential in the developmnt of all inquiry systems and epistemological paradigms. Thus design, a fundamental human activity, must be viewed as both scientific and technological, both implicit and explicit and contain both the practical process of making and the critical reflection inherent in the studies of meaning. Design must involve not only knowing what and knowing how, but also knowing why and knowing its purpose. And finally, to reflect the current trends in epistemological development, design should be expressed in terms of the embodiment of four principles: the princple of art, the principle of *techne*, the principle of critical scientific reflection and, perhaps the most important, the principle of epistemological incertitude.

REFERENCES

1 **Morin, E** 'Epistemologie de la technologie' in *Science avec Conscience* Fayard, Paris (1982)

2 **Cross, N, Naughton, J and Walker, D** 'Design method and scientific method' *Design Studies* Vol 2 No 4 (October, 1981) pp 198–201

3 **Morin, E** 'Theorie et méthode' in *Science avec Conscience* Fayard, Paris (1982)

4 **Varela, F J** *Principles of Biological Autonomy* North Holland, New York (1979)

5 **Kuhn, T S** *The Structure of Scientific Revolutions* University of Chicago Press, Chicago (1962)

6 **Lakatos, I** 'Falsification and the methodology of scientific research programmes' in **Lakatos, I and Musgrave, A (Eds)** *Criticism and the Growth of Knowledge: Proceedings of the International Colloquium in the Philosophy of Science, London, 1965* Vol 4, Cambridge University Press, Cambridge (1970)

7 **Popper, K** *Objective Knowledge: An Evolutionary Approach* 5th rev. ed. Clarendon Press, Oxford (1979)

8 **Polanyi, M** *Personal Knowledge: Towards a Post Critical Philosophy* University of Chicago Press, Chicago (1958)

9 **Toulmin S** *The Philosophy of Science: An Introduction* Hutchinson, London (1953)

10 **Feyerabend, P** *Against Method* New Left Books, London (1975)

11 **Polkinghorne, D** *Methodology for the Human Sciences* State University of New York Press, Albany (1983)

12 **Bertalanffy, L von** *General System Theory: Foundations, Development, Applications* George Braziller, New York (1968)

13 **Laszlo, E** *Introduction to Systems Philosophy: Toward a New Paradigm of Contemporary Thought* Gordon and Breach, New York (1972)

14 **Sutherland, J W** *A General Systems Philosophy for the Social and Behavioral Sciences*, George Braziller, New York (1973)

15 **Schedrovitsky, G P** 'Methodological organization of system-structural research and development: principles and general framework' *General Systems* Vol XXVII (1982) pp 75–96

16 **Levy, R, Forget, A and Laporte, I** 'Vers une paradigme systemique de la readaptation' in *Proc. of Conference Internationale sur la science des systèmes dans le domaine de services sociosanitaires pour les personnes âgées ou handicapées* Montréal (1983)

17 **Hempel, C G** 'Fundamentals of Concept Formation in Empirical Science' *International Encyclopedia of Unified Science* Vol 2 No 7 University of Chicago Press, Chicago (1952)

18 **Dray, W** *Laws and Explanation in History* Oxford University Press, Oxford (1957)

19 **Taylor, C** *The Explanation of Behavior* Routledge and Kegan Paul, London (1964)

20 **Quine, W** 'Two dogmas of empiricism' in *From a Logical Point of View* Harvard University Press, Cambridge, Massachusetts (1953)

21 **Desmarais, G** 'Pour une définition de la méthodologie hypothetico-deductive' unpublished paper, Faculté de l'Aménagement, Université de Montréal, Montréal (1984)

22 **Suppe, F** 'Afterwards–1977' in **Suppe, F (Ed)** *The Structure of Scientific Theories* 2nd ed., University of Illinois Press, Urbana (1977)

23 **Shapere, D** 'Scientific theories and their domains' in **Suppe, F (Ed)** *The Structure of Scientific Theories* 2nd ed., University of Illinois Press, Urbana (1977)

24 **Radnitzky, G** *Contemporary Schools of Metascience* Henry Regnery, Chicago (1973)

25 **Laudan, L** *Progress and Its Problems: Towards a Theory of*

Scientific Growth University of California Press, Berkeley (1977)

26 **Toulmin, S** *Human Understanding: The Collective Use and Evolution of Concepts* Princeton University Press, Princeton, N J (1972)

27 **Winner, L** 'Technologies as forms of life' in **Cohen, R S and Wartofsky, M W (Eds)** *Epistemology, Methodology and the Social Sciences* D Reidel Publishing Company, Dordrecht (1983)

28 **Rapp, F** *Analytical Philosophy of Technology* D Reidel Publishing Company, Dordrecht (1981)

29 **Ihde, D** *Technics and Praxis* D Reidel Publishing Company, Dordrecht (1979)

30 **Von Gottl-Ottlilienfeld, F** *Wirtschaft und Technik* Grundriss der Sozialökonomik, Section 5, Tübingen (1914) in Rapp *op. cit.*

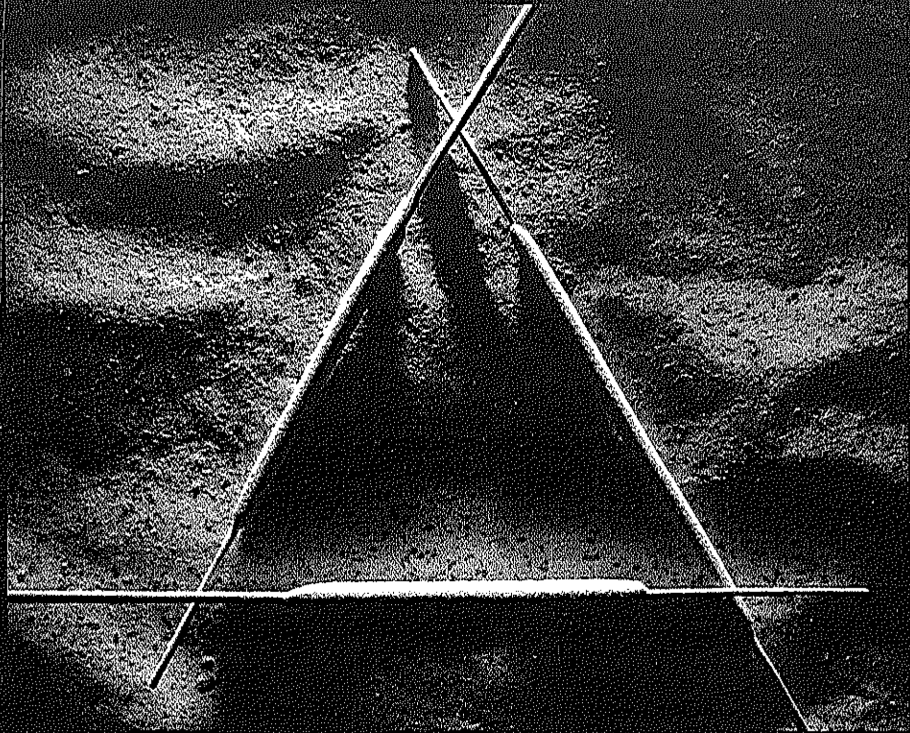
31 **McWhinney, I R** 'Medical knowledge and the rise of technology' *The Journal of Medicine and Philosophy* Vol 3 No 4 (1978) pp 293–304

32 **Habermas, J** *Knowledge and Human Interests* Beacon Press, Boston (1971)

33 **Piaget, J and Garcia, R** *Psychogenèse et Histoire des Sciences* Flammarion, Paris (1983)

34 **Bohm, D** *Wholeness and the Implicate Order* Routledge and Kegan Paul, London (1980)

Herbert A. Simon



The Sciences of the Artificial

Third Edition

1

Understanding the Natural and the Artificial Worlds

About three centuries after Newton we are thoroughly familiar with the concept of natural science—most unequivocally with physical and biological science. A natural science is a body of knowledge about some class of things—objects or phenomena—in the world: about the characteristics and properties that they have; about how they behave and interact with each other.

The central task of a natural science is to make the wonderful commonplace: to show that complexity, correctly viewed, is only a mask for simplicity; to find pattern hidden in apparent chaos. The early Dutch physicist Simon Stevin, showed by an elegant drawing (figure 1) that the law of the inclined plane follows in "self-evident fashion" from the *impossibility of perpetual motion*, for experience and reason tell us that the chain of balls in the figure would rotate neither to right nor to left but would remain at rest. (Since rotation changes nothing in the figure, if the chain moved at all, it would move perpetually.) Since the pendant part of the chain hangs symmetrically, we can snip it off without disturbing the equilibrium. But now the balls on the long side of the plane balance those on the shorter, steeper side, and their relative numbers are in inverse ratio ro the sines of the angles at which the planes are inclined.

Stevin was so pleased with his construction that he incorporated it into a vignette, inscribing above it

Wonder, en is gheen wonder

that is to say: "Wonderful, but not incomprehensible."

This is the task of natural science: to show that the wonderful is not incomprehensible, to show how it can be comprehended—but not to

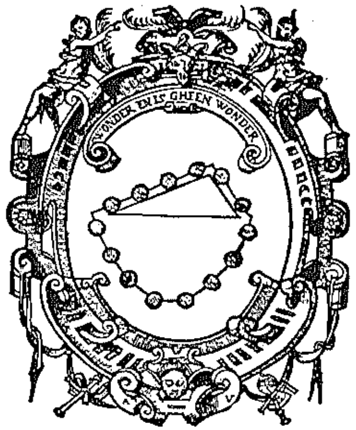


Figure 1
The vignette devised by Simon Stevin to illustrate his derivation of the law of the inclined plane

destroy wonder. For when we have explained the wonderful, unmasked the hidden pattern, a new wonder arises at how complexity was woven out of simplicity. The aesthetics of natural science and mathematics is at one with the aesthetics of music and painting—both inhere in the discovery of a partially concealed pattern.

The world we live in today is much more a man-made,[1] or artificial, world than it is a natural world. Almost every element in our environment shows evidence of human artifice. The temperature in which we spend most of our hours is kept artificially at 20 degrees Celsius; the humidity is added to or taken from the air we breathe; and the impurities we inhale are largely produced (and filtered) by man.

Moreover for most of us—the white-collared ones—the significant part of the environment consists mostly of strings of artifacts called "symbols" that we receive through eyes and ears in the form of written and spoken language and that we pour out into the environment—as I am now doing—by mouth or hand. The laws that govern these strings of

1. I will occasionally use "man" as an androgynous noun, encompassing both sexes, and "he," "his," and "him" as androgynous pronouns including women and men equally in their scope.

symbols, the laws that govern the occasions on which we emit and receive them, the determinants of their content are all consequences of our collective artifice.

One may object that I exaggerate the artificiality of our world. Man must obey the law of gravity as surely as does a stone, and as a living organism man must depend for food, and in many other ways, on the world of biological phenomena. I shall plead guilty to overstatement, while protesting that the exaggeration is slight. To say that an astronaut, or even an airplane pilot, is obeying the law of gravity, hence is a perfectly natural phenomenon, is true, but its truth calls for some sophistication in what we mean by "obeying" a natural law. Aristotle did not think it natural for heavy things to rise or light ones to fall (*Physics*, Book IV); but presumably we have a deeper understanding of "natural" than he did.

So too we must be careful about equating "biological" with "natural." A forest may be a phenomenon of nature; a farm certainly is not. The very species upon which we depend for our food—our corn and our cattle—are artifacts of our ingenuity. A plowed field is no more part of nature than an asphalted street—and no less.

These examples set the terms of our problem, for those things we call artifacts are not apart from nature. They have no dispensation to ignore or violate natural law. At the same time they are adapted to human goals and purposes. They are what they are in order to satisfy our desire to fly or to eat well. As our aims change, so too do our artifacts—and vice versa.

If science is to encompass these objects and phenomena in which human purpose as well as natural law are embodied, it must have means for relating these two disparate components. The character of these means and their implications for certain areas of knowledge—economics, psychology, and design in particular—are the central concern of this book.

The Artificial

Natural science is knowledge about natural objects and phenomena. We ask whether there cannot also be "artificial" science—knowledge about artificial objects and phenomena. Unfortunately the term "artificial" has a pejorative air about it that we must dispel before we can proceed.

My dictionary defines "artificial" as, "Produced by art rather than by nature; not genuine or natural; affected; not pertaining to the essence of the matter." It proposes, as synonyms: affected, factitious, manufactured, pretended, sham, simulated, spurious, trumped up, unnatural. As antonyms, it lists: actual, genuine, honest, natural, real, truthful, unaffected. Our language seems to reflect man's deep distrust of his own products. I shall not try to assess the validity of that evaluation or explore its possible psychological roots. But you will have to understand me as using "artificial" in as neutral a sense as possible, as meaning man-made as opposed to natural.[2]

In some contexts we make a distinction between "artificial" and "synthetic." For example, a gem made of glass colored to resemble sapphire would be called artificial, while a man-made gem chemically indistinguishable from sapphire would be called synthetic. A similar distinction is often made between "artificial" and "synthetic" rubber. Thus some artificial things are imitations of things in nature, and the imitation may use either the same basic materials as those in the natural object or quite different materials.

As soon as we introduce "synthesis" as well as "artifice," we enter the realm of engineering. For "synthetic" is often used in the broader sense of "designed" or "composed." We speak of engineering as concerned with "synthesis," while science is concerned with "analysis." Synthetic or artificial objects—and more specifically prospective artificial objects having desired properties—are the central objective of engineering activity and skill. The engineer, and more generally the designer, is concerned with how things *ought* to be—how they ought to be in order to *attain goals*,

2. I shall disclaim responsibility for this particular choice of terms. The phrase "artificial intelligence," which led me to it, was coined, I think, right on the Charles River, at MIT. Our own research group at Rand and Carnegie Mellon University have preferred phrases like "complex information processing" and "simulation of cognitive processes." But then we run into new terminological difficulties, for the dictionary also says that "to simulate" means "to assume or have the mere appearance or form of, without the reality; imitate; counterfeit; pretend." At any rate, "artificial intelligence" seems to be here to stay, and it may prove easier to cleanse the phrase than to dispense with it. In time it will become sufficiently idiomatic that it will no longer be the target of cheap rhetoric.

and to *function*. Hence a science of the artificial will be closely akin to a science of engineering—but very different, as we shall see in my fifth chapter, from what goes currently by the name of "engineering science."

With goals and "oughts" we also introduce into the picture the dichotomy between normative and descriptive. Natural science has found a way to exclude the normative and to concern itself solely with how things are. Can or should we maintain this exclusion when we move from natural to artificial phenomena, from analysis to synthesis?[3]

We have now identified four indicia that distinguish the artificial from the natural; hence we can set the boundaries for sciences of the artificial:

1. Artificial things are synthesized (though not always or usually with full forethought) by human beings.
2. Artificial things may imitate appearances in natural things while lacking, in one or many respects, the reality of the latter.
3. Artificial things can be characterized in terms of functions, goals, adaptation.
4. Artificial things are often discussed, particularly when they are being designed, in terms of imperatives as well as descriptives.

The Environment as Mold

Let us look a little more closely at the functional or purposeful aspect of artificial things. Fulfillment of purpose or adaptation to a goal involves a relation among three terms: the purpose or goal, the character of the artifact, and the environment in which the artifact performs. When we think of a clock, for example, in terms of purpose we may use the child's definition: "a clock is to tell time." When we focus our attention on the clock itself, we may describe it in terms of arrangements of gears and the

3. This issue will also be discussed at length in my fifth chapter. In order not to keep readers in suspense, I may say that I hold to the pristine empiricist's position of the irreducibility of "ought" to "is," as in chapter 3 of my *Administrative Behavior* (New York: Macmillan, 1976). This position is entirely consistent with treating natural or artificial goal-seeking systems as phenomena, without commitment to their goals. *Ibid.*, appendix. See also the well-known paper by A. Rosenbluth, N. Wiener, and J. Bigelow, "Behavior, Purpose, and Teleology," *Philosophy of Science, 10* (1943):18–24.

application of the forces of springs or gravity operating on a weight or pendulum.

But we may also consider clocks in relation to the environment in which they are to be used. Sundials perform as clocks *in sunny climates*—they are more useful in Phoenix than in Boston and of no use at all during the Arctic winter. Devising a clock that would tell time on a rolling and pitching ship, with sufficient accuracy to determine longitude, was one of the great adventures of eighteenth-century science and technology. To perform in this difficult environment, the clock had to be endowed with many delicate properties, some of them largely or totally irrelevant to the performance of a landlubber's clock.

Natural science impinges on an artifact through two of the three terms of the relation that characterizes it: the structure of the artifact itself and the environment in which it performs. Whether a clock will in fact tell time depends on its internal construction and where it is placed. Whether a knife will cut depends on the material of its blade and the hardness of the substance to which it is applied.

The Artifact as "Interface"

We can view the matter quite symmetrically. An artifact can be thought of as a meeting point—an "interface" in today's terms—between an "inner" environment, the substance and organization of the artifact itself, and an "outer" environment, the surroundings in which it operates. If the inner environment is appropriate to the outer environment, or vice versa, the artifact will serve its intended purpose. Thus, if the clock is immune to buffeting, it will serve as a ship's chronometer. (And conversely, if it isn't, we may salvage it by mounting it on the mantel at home.)

Notice that this way of viewing artifacts applies equally well to many things that are not man-made—to all things in fact that can be regarded as adapted to some situation; and in particular it applies to the living systems that have evolved through the forces of organic evolution. A theory of the airplane draws on natural science for an explanation of its inner environment (the power plant, for example), its outer environment (the character of the atmosphere at different altitudes), and the relation between its inner and outer environments (the movement of an airfoil

through a gas). But a theory of the bird can be divided up in exactly the same way.[4]

Given an airplane, or *given* a bird, we can analyze them by the methods of natural science without any particular attention to purpose or adaptation, without reference to the interface between what I have called the inner and outer environments. After all, their behavior is governed by natural law just as fully as the behavior of anything else (or at least we all believe this about the airplane, and most of us believe it about the bird).

Functional Explanation

On the other hand, if the division between inner and outer environment is not necessary to the analysis of an airplane or a bird, it turns out at least to be highly convenient. There are several reasons for this, which will become evident from examples.

Many animals in the Arctic have white fur. We usually explain this by saying that white is the best color for the Arctic environment, for white creatures escape detection more easily than do others. This is not of course a natural science explanation; it is an explanation by reference to purpose or function. It simply says that these are the kinds of creatures that will "work," that is, survive, in this kind of environment. To turn the statement into an explanation, we must add to it a notion of natural selection, or some equivalent mechanism.

An important fact about this kind of explanation is that it demands an understanding mainly of the outer environment. Looking at our snowy surroundings, we can predict the predominant color of the creatures we are likely to encounter; we need know little about the biology of the creatures themselves, beyond the facts that they are often mutually hostile, use visual clues to guide their behavior, and are adaptive (through selection or some other mechanism).

4. A generalization of the argument made here for the separability of "outer" from "inner" environment shows that we should expect to find this separability, to a greater or lesser degree, in *all* large and complex systems, whether they are artificial or natural. In its generalized form it is an argument that all nature will be organized in "levels." My essay "The Architecture of Complexity," included in this volume as chapter 8, develops the more general argument in some detail.

Analogous to the role played by natural selection in evolutionary biology is the role played by rationality in the sciences of human behavior. If we know of a business organization only that it is a profit-maximizing system, we can often predict how its behavior will change if we change its environment—how it will alter its prices if a sales tax is levied on its products. We can sometimes make this prediction—and economists do make it repeatedly—without detailed assumptions about the adaptive mechanism, the decision-making apparatus that constitutes the inner environment of the business firm.

Thus the first advantage of dividing outer from inner environment in studying an adaptive or artificial system is that we can often predict behavior from knowledge of the system's goals and its outer environment, with only minimal assumptions about the inner environment. An instant corollary is that we often find quite different inner environments accomplishing identical or similar goals in identical or similar outer environments—airplanes and birds, dolphins and tunafish, weight-driven clocks and battery-driven clocks, electrical relays and transistors.

There is often a corresponding advantage in the division from the standpoint of the inner environment. In very many cases whether a particular system will achieve a particular goal or adaptation depends on only a few characteristics of the outer environment and not at all on the detail of that environment. Biologists are familiar with this property of adaptive systems under the label of homeostasis. It is an important property of most good designs, whether biological or artifactual. In one way or another the designer insulates the inner system from the environment, so that an invariant relation is maintained between inner system and goal, independent of variations over a wide range in most parameters that characterize the outer environment. The ship's chronometer reacts to the pitching of the ship only in the negative sense of maintaining an invariant relation of the hands on its dial to the real time, independently of the ship's motions.

Quasi independence from the outer environment may be maintained by various forms of passive insulation, by reactive negative feedback (the most frequently discussed form of insulation), by predictive adaptation, or by various combinations of these.

Functional Description and Synthesis

In the best of all possible worlds—at least for a designer—we might even hope to combine the two sets of advantages we have described that derive from factoring an adaptive system into goals, outer environment, and inner environment. We might hope to be able to characterize the main properties of the system and its behavior without elaborating the detail of *either* the outer or inner environments. We might look toward a science of the artificial that would depend on the relative simplicity of the interface as its primary source of abstraction and generality.

Consider the design of a physical device to serve as a counter. If we want the device to be able to count up to one thousand, say, it must be capable of assuming any one of at least a thousand states, of maintaining itself in any given state, and of shifting from any state to the "next" state. There are dozens of different inner environments that might be used (and have been used) for such a device. A wheel notched at each twenty minutes of arc, and with a ratchet device to turn and hold it, would do the trick. So would a string of ten electrical switches properly connected to represent binary numbers. Today instead of switches we are likely to use transistors or other solid-state devices.[5]

Our counter would be activated by some kind of pulse, mechanical or electrical, as appropriate, from the outer environment. But by building an appropriate transducer between the two environments, the physical character of the interior pulse could again be made independent of the physical character of the exterior pulse—the counter could be made to count anything.

Description of an artifice in terms of its organization and functioning—its interface between inner and outer environments—is a major objective of invention and design activity. Engineers will find familiar the language of the following claim quoted from a 1919 patent on an improved motor controller:

> What I claim as new and desire to secure by Letters Patent is:
> 1 In a motor controller, in combination, reversing means, normally effective field-weakening means and means associated with said reversing means for

5. The theory of functional equivalence of computing machines has had considerable development in recent years. See Marvin L. Minsky, *Computation: Finite and Infinite Machines* (Englewood Cliffs, N.J.: Prentice-Hall, 1967), chapters 1–4.

rendering said field-weakening means ineffective during motor starting and thereafter effective to different degrees determinable by the setting of said reversing means . . . [6]

Apart from the fact that we know the invention relates to control of an electric motor, there is almost no reference here to specific, concrete objects or phenomena. There is reference rather to "reversing means" and "field-weakening means," whose further purpose is made clear in a paragraph preceding the patent claims:

The advantages of the special type of motor illustrated and the control thereof will be readily understood by those skilled in the art. Among such advantages may be mentioned the provision of a high starting torque and the provision for quick reversals of the motor.[7]

Now let us suppose that the motor in question is incorporated in a planing machine (see figure 2). The inventor describes its behavior thus:

Referring now to [figure 2], the controller is illustrated in outline connection with a planer (100) operated by a motor M, the controller being adapted to govern the motor M and to be automatically operated by the reciprocating bed (101) of the planer. The master shaft of the controller is provided with a lever (102) connected by a link (103) to a lever (104) mounted upon the planer frame and projecting into the path of lugs (105) and (106) on the planer bed. As will be understood, the arrangement is such that reverse movements of the planer bed will, through the connections described, throw the master shaft of the controller back and forth between its extreme positions and in consequence effect selective operation of the reversing switches (1) and (2) and automatic operation of the other switches in the manner above set forth.[8]

In this manner the properties with which the inner environment has been endowed are placed at the service of the goals in the context of the outer environment. The motor will reverse periodically under the control of the position of the planer bed. The "shape" of its behavior—the time path, say, of a variable associated with the motor—will be a function of the "shape" of the external environment—the distance, in this case, between the lugs on the planer bed.

The device we have just described illustrates in microcosm the nature of artifacts. Central to their description are the goals that link the inner

6. U.S. Patent 1,307,836, granted to Arthur Simon, June 24, 1919.

7. Ibid.

8. Ibid.

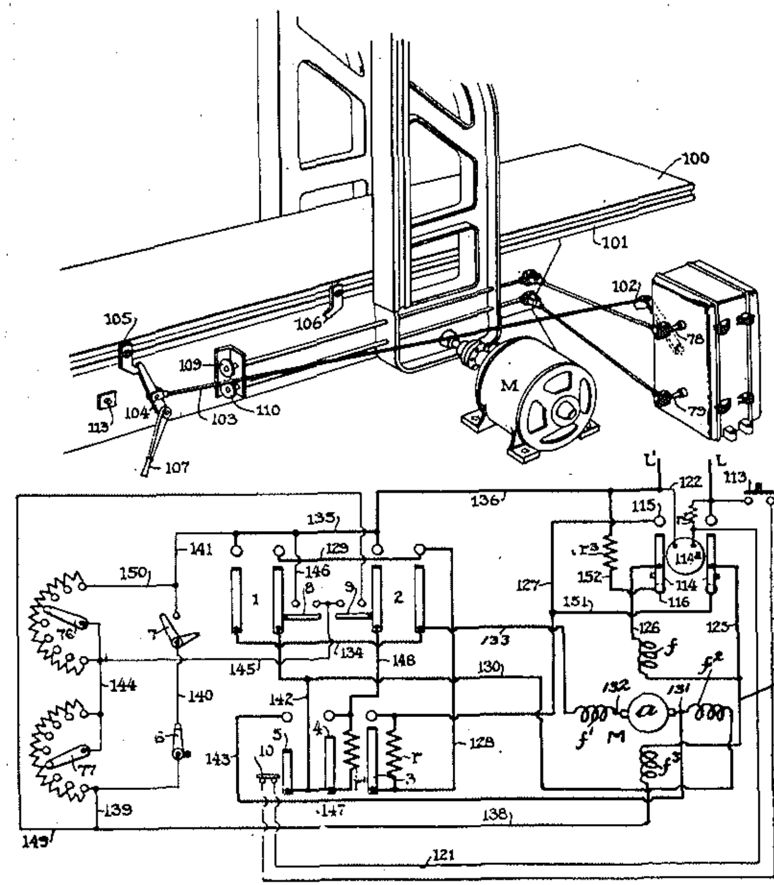


Figure 2
Illustrations from a patent for a motor controller

to the outer system. The inner system is an organization of natural phenomena capable of attaining the goals in some range of environments, but ordinarily there will be many functionally equivalent natural systems capable of doing this.

The outer environment determines the conditions for goal attainment. If the inner system is properly designed, it will be adapted to the outer environment, so that its behavior will be determined in large part by the behavior of the latter, exactly as in the case of "economic man." To predict how it will behave, we need only ask, "How would a rationally designed system behave under these circumstances?" The behavior takes on the shape of the task environment.[9]

Limits of Adaptation

But matters must be just a little more complicated than this account suggests. "If wishes were horses, all beggars would ride." And if we could always specify a protean inner system that would take on exactly the shape of the task environment, designing would be synonymous with wishing. "Means for scratching diamonds" defines a design objective, an objective that *might* be attained with the use of many different substances. But the design has not been achieved until we have discovered at least one realizable inner system obeying the ordinary natural laws—one material, in this case, hard enough to scratch diamonds.

Often we shall have to be satisfied with meeting the design objectives only approximately. Then the properties of the inner system will "show through." That is, the behavior of the system will only partly respond to the task environment; partly, it will respond to the limiting properties of the inner system.

Thus the motor controls described earlier are aimed at providing for "quick" reversal of the motor. But the motor must obey electromagnetic and mechanical laws, and we could easily confront the system with a task where the environment called for quicker reversal than the motor was capable of. In a benign environment we would learn from the motor only what it had been called upon to do; in a taxing environment we would learn something about its internal structure—specifically about those aspects of the internal structure that were chiefly instrumental in limiting performance.[10]

9. On the crucial role of adaptation or rationality—and their limits—for economics and organization theory, see the introduction to part IV, "Rationality and Administrative Decision Making," of my *Models of Man* (New York: Wiley, 1957); pp. 38–41, 80–81, and 240–244 of *Administrative Behavior;* and chapter 2 of this book.

10. Compare the corresponding proposition on the design of administrative organizations: "Rationality, then, does not determine behavior. Within the area of rationality behavior is perfectly flexible and adaptable to abilities, goals, and knowledge. Instead, behavior is determined by the irrational and nonrational elements that bound the area of rationality . . . administrative theory must be concerned with the limits of rationality, and the manner in which organization affects these limits for the person making a decision." *Administrative Behavior,* p. 241. For a discussion of the same issue as it arises in psychology, see my "Cognitive Architectures and Rational Analysis: Comment," in Kurt VanLehn (ed.), *Architectures for Intelligence* (Hillsdale, NJ: Erlbaum, 1991).

A bridge, under its usual conditions of service, behaves simply as a relatively smooth level surface on which vehicles can move. Only when it has been overloaded do we learn the physical properties of the materials from which it is built.

Understanding by Simulating

Artificiality connotes perceptual similarity but essential difference, resemblance from without rather than within. In the terms of the previous section we may say that the artificial object imitates the real by turning the same face to the outer system, by adapting, relative to the same goals, to comparable ranges of external tasks. Imitation is possible because distinct physical systems can be organized to exhibit nearly identical behavior. The damped spring and the damped circuit obey the same second-order linear differential equation; hence we may use either one to imitate the other.

Techniques of Simulation

Because of its abstract character and its symbol manipulating generality, the digital computer has greatly extended the range of systems whose behavior can be imitated. Generally we now call the imitation "simulation," and we try to understand the imitated system by testing the simulation in a variety of simulated, or imitated, environments.

Simulation, as a technique for achieving understanding and predicting the behavior of systems, predates of course the digital computer. The model basin and the wind tunnel are valued means for studying the behavior of large systems by modeling them in the small, and it is quite certain that Ohm's law was suggested to its discoverer by its analogy with simple hydraulic phenomena.

Simulation may even take the form of a thought experiment, never actually implemented dynamically. One of my vivid memories of the Great Depression is of a large multicolored chart in my father's study that represented a hydraulic model of an economic system (with different fluids for money and goods). The chart was devised by a technocratically inclined engineer named Dahlberg. The model never got beyond the pen-and-paint stage at that time, but it could be used to trace through the imputed consequences of particular economic measures or events—provided the theory was right![11]

As my formal education in economics progressed, I acquired a disdain for that naive simulation, only to discover after World War II that a distinguished economist, Professor A. W. Phillips had actually built the Moniac, a hydraulic model that simulated a Keynesian economy.[12] Of course Professor Phillips's simulation incorporated a more nearly correct theory than the earlier one and was actually constructed and operated—two points in its favor. However, the Moniac, while useful as a teaching tool, told us nothing that could not be extracted readily from simple mathematical versions of Keynesian theory and was soon priced out of the market by the growing number of computer simulations of the economy.

Simulation as a Source of New Knowledge

This brings me to the crucial question about simulation: *How can a simulation ever tell us anything that we do not already know?* The usual implication of the question is that it can't. As a matter of fact, there is an interesting parallelism, which I shall exploit presently, between two assertions about computers and simulation that one hears frequently:

1. A simulation is no better than the assumptions built into it.
2. A computer can do only what it is programmed to do.

I shall not deny either assertion, for both seem to me to be true. But despite both assertions simulation can tell us things we do not already know.

11. For some published versions of this model, see A. O. Dahlberg, *National Income Visualized* (N.Y.: Columbia University Press, 1956).

12. A. W. Phillips, "Mechanical Models in Economic Dynamics," *Economica*, New Series, *17* (1950):283–305.

There are two related ways in which simulation can provide new knowledge—one of them obvious, the other perhaps a bit subtle. The obvious point is that, even when we have correct premises, it may be very difficult to discover what they imply. All correct reasoning is a grand system of tautologies, but only God can make direct use of that fact. The rest of us must painstakingly and fallibly tease out the consequences of our assumptions.

Thus we might expect simulation to be a powerful technique for deriving, from our knowledge of the mechanisms governing the behavior of gases, a theory of the weather and a means of weather prediction. Indeed, as many people are aware, attempts have been under way for some years to apply this technique. Greatly oversimplified, the idea is that we already know the correct basic assumptions, the local atmospheric equations, but we need the computer to work out the implications of the interactions of vast numbers of variables starting from complicated initial conditions. This is simply an extrapolation to the scale of modern computers of the idea we use when we solve two simultaneous equations by algebra.

This approach to simulation has numerous applications to engineering design. For it is typical of many kinds of design problems that the inner system consists of components whose fundamental laws of behavior—mechanical, electrical, or chemical—are well known. The difficulty of the design problem often resides in predicting how an assemblage of such components will behave.

Simulation of Poorly Understood Systems

The more interesting and subtle question is whether simulation can be of any help to us when we do not know very much initially about the natural laws that govern the behavior of the inner system. Let me show why this question must also be answered in the affirmative.

First, I shall make a preliminary comment that simplifies matters: we are seldom interested in explaining or predicting phenomena in all their particularity; we are usually interested only in a few properties abstracted from the complex reality. Thus, a NASA-launched satellite is surely an artificial object, but we usually do not think of it as "simulating" the moon or a planet. It simply obeys the same laws of physics, which relate only to its inertial and gravitational mass, abstracted from most of its other properties. It *is* a moon. Similarly electric energy that entered my house from the early atomic generating station at Shippingport did not "simulate" energy generated by means of a coal plant or a windmill. Maxwell's equations hold for both.

The more we are willing to abstract from the detail of a set of phenomena, the easier it becomes to simulate the phenomena. Moreover we do not have to know, or guess at, all the internal structure of the system but only that part of it that is crucial to the abstraction.

It is fortunate that this is so, for if it were not, the topdown strategy that built the natural sciences over the past three centuries would have been infeasible. We knew a great deal about the gross physical and chemical behavior of matter before we had a knowledge of molecules, a great deal about molecular chemistry before we had an atomic theory, and a great deal about atoms before we had any theory of elementary particles—if indeed we have such a theory today.

This skyhook-skyscraper construction of science from the roof down to the yet unconstructed foundations was possible because the behavior of the system at each level depended on only a very approximate, simplified, abstracted characterization of the system at the level next beneath.[13] This is lucky, else the safety of bridges and airplanes might depend on the correctness of the "Eightfold Way" of looking at elementary particles.

Artificial systems and adaptive systems have properties that make them particularly susceptible to simulation via simplified models. The characterization of such systems in the previous section of this chapter

13. This point is developed more fully in "The Architecture of Complexity," chapter 8 in this volume. More than fifty years ago, Bertrand Russell made the same point about the architecture of mathematics. See the "Preface" to *Principia Mathematica*: ". . . the chief reason in favour of any theory on the principles of mathematics must always be inductive, i.e., it must lie in the fact that the theory in question enables us to deduce ordinary mathematics. In mathematics, the greatest degree of self-evidence is usually not to be found quite at the beginning, but at some later point; hence the early deductions, until they reach this point, give reasons rather for believing the premises because true consequences follow from them, than for believing the consequences because they follow from the premises." Contemporary preferences for deductive formalisms frequently blind us to this important fact, which is no less true today than it was in 1910.

explains why. Resemblance in behavior of systems without identity of the inner systems is particularly feasible if the aspects in which we are interested arise out of the *organization* of the parts, independently of all but a few properties of the individual components. Thus for many purposes we may be interested in only such characteristics of a material as its tensile and compressive strength. We may be profoundly unconcerned about its chemical properties, or even whether it is wood or iron.

The motor control patent cited earlier illustrates this abstraction to organizational properties. The invention consisted of a "combination" of "reversing means," of "field weakening means," that is to say, of components specified in terms of their functioning in the organized whole. How many ways are there of reversing a motor, or of weakening its field strength? We can simulate the system described in the patent claims in many ways without reproducing even approximately the actual physical device that is depicted. With a small additional step of abstraction, the patent claims could be restated to encompass mechanical as well as electrical devices. I suppose that any undergraduate engineer at Berkeley, Carnegie Mellon University, or MIT could design a mechanical system embodying reversibility and variable starting torque so as to simulate the system of the patent.

The Computer as Artifact

No artifact devised by man is so convenient for this kind of functional description as a digital computer. It is truly protean, for almost the only ones of its properties that are detectable in its behavior (when it is operating properly!) are the organizational properties. The speed with which it performs it basic operations may allow us to infer a little about its physical components and their natural laws; speed data, for example, would allow us to rule out certain kinds of "slow" components. For the rest, almost no interesting statement that one can make about an operating computer bears any particular relation to the specific nature of the hardware. A computer is an organization of elementary functional components in which, to a high approximation, only the function performed by those components is relevant to the behavior of the whole system.[14]

Computers as Abstract Objects

This highly abstractive quality of computers makes it easy to introduce mathematics into the study of their theory—and has led some to the erroneous conclusion that, as a computer science emerges, it will necessarily be a mathematical rather than an empirical science. Let me take up these two points in turn: the relevance of mathematics to computers and the possibility of studying computers empirically.

Some important theorizing, initiated by John von Neumann, has been done on the topic of computer reliability. The question is how to build a reliable system from unreliable parts. Notice that this is not posed as a question of physics or physical engineering. The components engineer is assumed to have done his best, but the parts are still unreliable! We can cope with the unreliability only by our manner of organizing them.

To turn this into a meaningful problem, we have to say a little more about the nature of the unreliable parts. Here we are aided by the knowledge that *any* computer can be assembled out of a small array of simple, basic elements. For instance, we may take as our primitives the so-called Pitts-McCulloch neurons. As their name implies, these components were devised in analogy to the supposed anatomical and functional characteristics of neurons in the brain, but they are highly abstracted. They are formally isomorphic with the simplest kinds of switching circuits—"and," "or," and "not" circuits. We postulate, now, that we are to build a system from such elements and that each elementary part has a specified probability of functioning correctly. The problem is to arrange the elements and their interconnections in such a way that the complete system will perform reliably.

The important point for our present discussion is that the parts could as well be neurons as relays, as well relays as transistors. The natural laws governing relays are very well known, while the natural laws governing

14. On the subject of this and the following paragraphs, see M. L. Minsky, *op. cit.;* then John von Neumann, "Probabilistic Logics and the Synthesis of Reliable Organisms from Unreliable Components," in C. E. Shannon and J. McCarthy (eds.), *Automata Studies* (Princeton: Princeton University Press, 1956).

neurons are known most imperfectly. But that does not matter, for all that is relevant for the theory is that the components have the specified level of unreliability and be interconnected in the specified way.

This example shows that the possibility of building a mathematical theory of a system or of simulating that system does not depend on having an adequate microtheory of the natural laws that govern the system components. Such a microtheory might indeed be simply irrelevant.

Computers as Empirical Objects

We turn next to the feasibility of an *empirical* science of computers—as distinct from the solid-state physics or physiology of their componentry.[15] As a matter of empirical fact almost all of the computers that have been designed have certain common organizational features. They almost all can be decomposed into an active processor (Babbage's "Mill") and a memory (Babbage's "Store") in combination with input and output devices. (Some of the larger systems, somewhat in the manner of colonial algae, are assemblages of smaller systems having some or all of these components. But perhaps I may oversimplify for the moment.) They are all capable of storing symbols (program) that can be interpreted by a program-control component and executed. Almost all have exceedingly limited capacity for simultaneous, parallel activity—they are basically one-thing-at-a-time systems. Symbols generally have to be moved from the larger memory components into the central processor before they can be acted upon. The systems are capable of only simple basic actions: recoding symbols, storing symbols, copying symbols, moving symbols, erasing symbols, and comparing symbols.

Since there are now many such devices in the world, and since the properties that describe them also appear to be shared by the human central nervous system, nothing prevents us from developing a natural history of them. We can study them as we would rabbits or chipmunks and discover how they behave under different patterns of environmental stimulation. Insofar as their behavior reflects largely the broad functional

15. A. Newell and H. A. Simon, "Computer Science as Empirical Inquiry," *Communications of the ACM, 19*(March 1976):113–126. See also H. A. Simon, "Artificial Intelligence: An Empirical Science," *Artificial Intelligence,* 77(1995): 95–127.

characteristics we have described, and is independent of details of their hardware, we can build a general—but empirical—theory of them.

The research that was done to design computer time-sharing systems is a good example of the study of computer behavior as an empirical phenomenon. Only fragments of theory were available to guide the design of a time-sharing system or to predict how a system of a specified design would actually behave in an environment of users who placed their several demands upon it. Most actual designs turned out initially to exhibit serious deficiencies, and most predictions of performance were startlingly inaccurate.

Under these circumstances the main route open to the development and improvement of time-sharing systems was to build them and see how they behaved. And this is what was done. They were built, modified, and improved in successive stages. Perhaps theory could have anticipated these experiments and made them unnecessary. In fact it didn't, and I don't know anyone intimately acquainted with these exceedingly complex systems who has very specific ideas as to how it might have done so. To understand them, the systems had to be constructed, and their behavior observed.[16]

In a similar vein computer programs designed to play games or to discover proofs for mathematical theorems spend their lives in exceedingly large and complex task environments. Even when the programs themselves are only moderately large and intricate (compared, say, with the monitor and operating systems of large computers), too little is known about their task environments to permit accurate prediction of how well they will perform, how selectively they will be able to search for problem solutions.

Here again theoretical analysis must be accompanied by large amounts of experimental work. A growing literature reporting these experiments is beginning to give us precise knowledge about the degree of heuristic power of particular heuristic devices in reducing the size of the problem spaces that must be searched. In theorem proving, for example, there has

16. The empirical, exploratory flavor of computer research is nicely captured by the account of Maurice V. Wilkes in his 1967 Turing Lecture, "Computers Then and Now," *Journal of the Association for Computing Machinery,* 15(January 1968):1–7.

been a whole series of advances in heuristic power based on and guided by empirical exploration: the use of the Herbrand theorem, the resolution principle, the set-of-support principle, and so on.[17]

Computers and Thought

As we succeed in broadening and deepening our knowledge—theoretical and empirical—about computers, we discover that in large part their behavior is governed by simple general laws, that what appeared as complexity in the computer program was to a considerable extent complexity of the environment to which the program was seeking to adapt its behavior.

This relation of program to environment opened up an exceedingly important role for computer simulation as a tool for achieving a deeper understanding of human behavior. For if it is the organization of components, and not their physical properties, that largely determines behavior, and if computers are organized somewhat in the image of man, then the computer becomes an obvious device for exploring the consequences of alternative organizational assumptions for human behavior. Psychology could move forward without awaiting the solutions by neurology of the problems of component design—however interesting and significant these components turn out to be.

Symbol Systems: Rational Artifacts

The computer is a member of an important family of artifacts called symbol systems, or more explicitly, physical symbol systems.[18] Another important member of the family (some of us think, anthropomorphically, it is the *most* important) is the human mind and brain. It is with this family

17. Note, for example, the empirical data in Lawrence Wos, George A. Robinson, Daniel F. Carson, and Leon Shalla, "The Concept of Demodulation in Theorem Proving," *Journal of the Association for Computing Machinery,* 14(October 1967):698–709, and in several of the earlier papers referenced there. See also the collection of programs in Edward Feigenbaum and Julian Feldman (eds.), *Computers and Thought* (New York: McGraw-Hill, 1963). It is common practice in the field to title papers about heuristic programs, "Experiments with an XYZ Program."

18. In the literature the phrase *information-processing system* is used more frequently than symbol system. I will use the two terms as synonyms.

of artifacts, and particularly the human version of it, that we will be primarily concerned in this book. Symbol systems are almost the quintessential artifacts, for adaptivity to an environment is their whole *raison d'être.* They are goal-seeking, information-processing systems, usually enlisted in the service of the larger systems in which they are incorporated.

Basic Capabilities of Symbol Systems

A physical symbol system holds a set of entities, called symbols. These are physical patterns (e.g., chalk marks on a blackboard) that can occur as components of symbol structures (sometimes called "expressions"). As I have already pointed out in the case of computers, a symbol system also possesses a number of simple processes that operate upon symbol structures—processes that create, modify, copy, and destroy symbols. A physical symbol system is a machine that, as it moves through time, produces an evolving collection of symbol structures.[19] Symbol structures can, and commonly do, serve as internal representations (e.g., "mental images") of the environments to which the symbol system is seeking to adapt. They allow it to model that environment with greater or less veridicality and in greater or less detail, and consequently to reason about it. Of course, for this capability to be of any use to the symbol system, it must have windows on the world and hands, too. It must have means for acquiring information from the external environment that can be encoded into internal symbols, as well as means for producing symbols that initiate action upon the environment. Thus it must use symbols to *designate* objects and relations and actions in the world external to the system.

Symbols may also designate processes that the symbol system can interpret and execute. Hence the programs that govern the behavior of a symbol system can be stored, along with other symbol structures, in the system's own memory, and executed when activated.

Symbol systems are called "physical" to remind the reader that they exist as real-world devices, fabricated of glass and metal (computers) or flesh and blood (brains). In the past we have been more accustomed to thinking of the symbol systems of mathematics and logic as abstract and disembodied, leaving out of account the paper and pencil and human minds that were required actually to bring them to life. Computers have

19. Newell and Simon, "Computer Science as Empirical Inquiry," p. 116.

transported symbol systems from the platonic heaven of ideas to the empirical world of actual processes carried out by machines or brains, or by the two of them working together.

Intelligence as Computation

The three chapters that follow rest squarely on the hypothesis that intelligence is the work of symbol systems. Stated a little more formally, the hypothesis is that a physical symbol system of the sort I have just described has the necessary and sufficient means for general intelligent action.

The hypothesis is clearly an empirical one, to be judged true or false on the basis of evidence. One task of chapters 3 and 4 will be to review some of the evidence, which is of two basic kinds. On the one hand, by constructing computer programs that are demonstrably capable of intelligent action, we provide evidence on the sufficiency side of the hypothesis. On the other hand, by collecting experimental data on human thinking that tend to show that the human brain operates as a symbol system, we add plausibility to the claims for necessity, for such data imply that all known intelligent systems (brains and computers) are symbol systems.

Economics: Abstract Rationality

As prelude to our consideration of human intelligence as the work of a physical symbol system, chapter 2 introduces a heroic abstraction and idealization—the idealization of human rationality which is enshrined in modern economic theories, particularly those called neoclassical. These theories are an idealization because they direct their attention primarily to the external environment of human thought, to decisions that are optimal for realizing the adaptive system's goals (maximization of utility or profit). They seek to define the decisions that would be substantively rational in the circumstances defined by the outer environment.

Economic theory's treatment of the limits of rationality imposed by the inner environment—by the characteristics of the physical symbol system—tends to be pragmatic, and sometimes even opportunistic. In the more formal treatments of general equilibrium and in the so-called "rational expectations" approach to adaptation, the possibilities that an information-processing system may have a very limited capability for adaptation are almost ignored. On the other hand, in discussions of the rationale for market mechanisms and in many theories of decision making under uncertainty, the procedural aspects of rationality receive more serious treatment.

In chapter 2 we will see examples both of neglect for and concern with the limits of rationality. From the idealizations of economics (and some criticisms of these idealizations) we will move, in chapters 3 and 4, to a more systematic study of the inner environment of thought—of thought processes as they actually occur within the constraints imposed by the parameters of a physical symbol system like the brain.

2

Economic Rationality: Adaptive Artifice

Because scarcity is a central fact of life—land, money, fuel, time, attention, and many other things are scarce—it is a task of rationality to allocate scarce things. Performing that task is the focal concern of economics.

Economics exhibits in purest form the artificial component in human behavior, in individual actors, business firms, markets, and the entire economy. The outer environment is defined by the behavior of other individuals, firms, markets, or economies. The inner environment is defined by an individual's, firm's, market's, or economy's goals and capabilities for rational, adaptive behavior. Economics illustrates well how outer and inner environment interact and, in particular, how an intelligent system's adjustment to its outer environment (its *substantive rationality*) is limited by its ability, through knowledge and computation, to discover appropriate adaptive behavior (its *procedural rationality*).

The Economic Actor

In the textbook theory of the business firm, an "entrepreneur" aims at maximizing profit, and in such simple circumstances that the computational ability to find the maximum is not in question. A cost curve relates dollar expenditures to amount of product manufactured, and a revenue curve relates income to amount of product sold. The goal (maximizing the difference between income and expenditure) fully defines the firm's inner environment. The cost and revenue curves define the outer environment.[1] Elementary calculus shows how to find the profit-maximizing

1. I am drawing the line between outer and inner environment not at the firm's boundary but at the skin of the entrepreneur, so that the factory is part of the external technology; the brain, perhaps assisted by computers, is the internal.

Human beings, viewed as behaving systems, are quite simple. The apparent complexity of our behavior over time is largely a reflection of the complexity of the environment in which we find ourselves . . .

provided that we include in what we call the human environment the cocoon of information, stored in books and in long-term memory, that we spin about ourselves.

That information, stored both as data and as procedures and richly indexed for access in the presence of appropriate stimuli, enables the simple basic information processes to draw upon a very large repertory of information and strategies, and accounts for the appearance of complexity in their behavior. The inner environment, the hardware, is simple. Complexity emerges from the richness of the outer environment, both the world apprehended through the senses and the information about the world stored in long-term memory.

A scientific account of human cognition describes it in terms of several sets of invariants. First, there are the parameters of the inner environment. Then, there are the general control and search-guiding mechanisms that are used over and over again in all task domains. Finally, there are the learning and discovery mechanisms that permit the system to adapt with gradually increasing effectiveness to the particular environment in which it finds itself. The adaptiveness of the human organism, the facility with which it acquires new representations and strategies and becomes adept in dealing with highly specialized environments, makes it an elusive and fascinating target of our scientific inquiries—and the very prototype of the artificial.

5

The Science of Design: Creating the Artificial

Historically and traditionally, it has been the task of the science disciplines to teach about natural things: how they are and how they work. It has been the task of engineering schools to teach about artificial things: how to make artifacts that have desired properties and how to design.

Engineers are not the only professional designers. Everyone designs who devises courses of action aimed at changing existing situations into preferred ones. The intellectual activity that produces material artifacts is no different fundamentally from the one that prescribes remedies for a sick patient or the one that devises a new sales plan for a company or a social welfare policy for a state. Design, so construed, is the core of all professional training; it is the principal mark that distinguishes the professions from the sciences. Schools of engineering, as well as schools of architecture, business, education, law, and medicine, are all centrally concerned with the process of design.

In view of the key role of design in professional activity, it is ironic that in this century the natural sciences almost drove the sciences of the artificial from professional school curricula, a development that peaked about two or three decades after the Second World War. Engineering schools gradually became schools of physics and mathematics; medical schools became schools of biological science; business schools became schools of finite mathematics. The use of adjectives like "applied" concealed, but did not change, the fact. It simply meant that in the professional schools those topics were selected from mathematics and the natural sciences for emphasis which were thought to be most nearly relevant to professional practice. It did not mean that design continued to be taught, as distinguished from analysis.

The movement toward natural science and away from the sciences of the artificial proceeded further and faster in engineering, business, and medicine than in the other professional fields I have mentioned, though it has by no means been absent from schools of law, journalism, and library science. The stronger universities were more deeply affected than the weaker, and the graduate programs more than the undergraduate. During that time few doctoral dissertations in first-rate professional schools dealt with genuine design problems, as distinguished from problems in solid-state physics or stochastic processes. I have to make partial exceptions—for reasons I shall mention—of dissertations in computer science and management science, and there were undoubtedly some others, for example, in chemical engineering.

Such a universal phenomenon must have had a basic cause. It did have a very obvious one. As professional schools, including the independent engineering schools, were more and more absorbed into the general culture of the university, they hankered after academic respectability. In terms of the prevailing norms, academic respectability calls for subject matter that is intellectually tough, analytic, formalizable, and teachable. In the past much, if not most, of what we knew about design and about the artificial sciences was intellectually soft, intuitive, informal, and cookbooky. Why would anyone in a university stoop to teach or learn about designing machines or planning market strategies when he could concern himself with solid-state physics? The answer has been clear: he usually wouldn't.

The damage to professional competence caused by the loss of design from professional curricula gradually gained recognition in engineering and medicine and to a lesser extent in business. Some schools did not think it a problem (and a few still do not), because they regarded schools of applied science as a superior alternative to the trade schools of the past. If that were the choice, we could agree.[1] But neither alternative is satisfactory. The older kind of professional school did not know how to educate for professional design at an intellectual level appropriate to a university; the newer kind of school nearly abdicated responsibility for training in the core professional skill. Thus we were faced with a problem of devising a professional school that could attain two objectives simultaneously: education in both artificial and natural science at a high intellectual level. This too is a problem of design—organizational design.

The kernel of the problem lies in the phrase "artificial science." The previous chapters have shown that a science of artificial phenomena is always in imminent danger of dissolving and vanishing. The peculiar properties of the artifact lie on the thin interface between the natural laws within it and the natural laws without. What can we say about it? What is there to study besides the boundary sciences—those that govern the means and the task environment?

The artificial world is centered precisely on this interface between the inner and outer environments; it is concerned with attaining goals by adapting the former to the latter. The proper study of those who are concerned with the artificial is the way in which that adaptation of means to environments is brought about—and central to that is the process of design itself. The professional schools can reassume their professional responsibilities just to the degree that they discover and teach a science of design, a body of intellectually tough, analytic, partly formalizable, partly empirical, teachable doctrine about the design process.

It is the thesis of this chapter that such a science of design not only is possible but also has been emerging since the mid-1970s. In fact, it is fair to say that the first edition of this book, published in 1969, was influential in its development, serving as a call to action and outlining the form that the action could take. At Carnegie Mellon University, one of the first engineering schools to move toward research on the process of design, the

1. That was in fact the choice in our engineering schools a generation ago. The schools needed to be purged of vocationalism; and a genuine science of design did not exist even in a rudimentary form as an alternative. Hence, introducing more fundamental science was the road forward. This was a main theme in Karl Taylor Compton's presidential inaugural address at MIT in 1930:

> I hope . . . that increasing attention in the Institute may be given to the fundamental sciences; that they may achieve as never before the spirit and results of research; that all courses of instruction may be examined carefully to see where training in details has been unduly emphasized at the expense of the more powerful training in all-embracing fundamental principles.

Notice that President Compton's emphasis was on "fundamental," an emphasis as sound today as it was in 1930. What is called for is not a departure from the fundamental but an inclusion in the curriculum of the fundamental in engineering along with the fundamental in natural science. That was not possible in 1930; but it is possible today.

step was to form a Design Research Center, about 1975. The Cen-
nce 1985 called the "Engineering Design Research Center") facili-
collaboration among the faculty and students undertaking research
: science and practice of design and developed elements of a theory
ign that found their way back into the undergraduate and graduate
ula. The Center continues to play an important role in the modern-
a and strengthening of education and research in design at Carnegie
n and elsewhere in the United States.

ubstantial part, design theory is aimed at broadening the capabili-
computers to aid design, drawing upon the tools of artificial intel-
e and operations research. Hence, research on many aspects of
uter-aided design is being pursued with growing intensity in com-
science, engineering and architecture departments, and in opera-
research groups in business schools. The need to make design theory
it and precise in order to introduce computers into the process has
the key to establishing its academic acceptability—its appropriate-
or a university. In the remainder of this chapter I will take up some
: topics that need to be incorporated in a theory of design and in
ction in design.

ogic of Design: Fixed Alternatives

iust start with some questions of logic.[2] The natural sciences are
rned with how things are. Ordinary systems of logic—the standard
sitional and predicate calculi, say—serve these sciences well. Since
oncern of standard logic is with declarative statements, it is well
l for assertions about the world and for inferences from those
ions.

sign, on the other hand, is concerned with how things ought to be,
devising artifacts to attain goals. We might question whether the

ave treated the question of logical formalism for design at greater length
earlier papers: "The Logic of Rational Decision," *British Journal for the
ophy of Science*, 16(1965):169–186; and "The Logic of Heuristic Decision
ig," in Nicholas Rescher (ed.), *The Logic of Decision and Action* (Pitts-
: University of Pittsburgh Press, 1967), pp. 1–35. The present discussion is
on these two papers, which have been reprinted as chapters 3.1 and 3.2 in
odels of Discovery (Dordrecht: D. Reidel Pub. Co., 1977).

forms of reasoning that are appropriate to natural science are suitable also for design. One might well suppose that introduction of the verb "should" may require additional rules of inference, or modification of the rules already imbedded in declarative logic.

Paradoxes of Imperative Logic

Various "paradoxes" have been constructed to demonstrate the need for a distinct logic of imperatives, or a normative, deontic logic. In ordinary logic from "Dogs are pets" and "Cats are pets," one can infer "Dogs and cats are pets." But from "Dogs are pets," "Cats are pets," and "You should keep pets," can one infer "You should keep cats and dogs"? And from "Give me needle and thread!" can one deduce, in analogy with declarative logic, "Give me needle or thread!"? Easily frustrated people would perhaps rather have neither needle nor thread than one without the other, and peace-loving people, neither cats nor dogs, rather than both.

As a response to these challenges of apparent paradox, there have been developed a number of constructions of modal logic for handling "shoulds," "shalts," and "oughts" of various kinds. I think it is fair to say that none of these systems has been sufficiently developed or sufficiently widely applied to demonstrate that it is adequate to handle the logical requirements of the process of design.

Fortunately, such a demonstration is really not essential, for it can be shown that the requirements of design can be met fully by a modest adaptation of ordinary declarative logic. Thus a special logic of imperatives is unnecessary.

I should like to underline the word "unnecessary," which does not mean "impossible." Modal logics can be shown to exist in the same way that giraffes can—namely, by exhibiting some of them. The question is not whether they exist, but whether they are needed for, or even useful for, design.

Reduction to Declarative Logic

The easiest way to discover what kinds of logic are needed for design is to examine what kinds of logic designers use when they are being careful about their reasoning. Now there would be no point in doing this if designers were always sloppy fellows who reasoned loosely, vaguely, and intuitively. Then we might say that whatever logic they used was not the logic they *should* use.

However, there exists a considerable area of design practice where standards of rigor in inference are as high as one could wish. I refer to the domain of so-called "optimization methods," most highly developed in statistical decision theory and management science but acquiring growing importance also in engineering design theory. The theories of probability and utility, and their intersection, have received the painstaking attention not only of practical designers and decision makers but also of a considerable number of the most distinguished logicians and mathematicians of recent generations. F. P. Ramsey, B. de Finetti, A. Wald, J. von Neumann, J. Neyman, K. Arrow, and L. J. Savage are examples.

The logic of optimization methods can be sketched as follows: The "inner environment" of the design problem is represented by a set of given alternatives of action. The alternatives may be given *in extenso:* more commonly they are specified in terms of *command variables* that have defined domains. The "outer environment" is represented by a set of parameters, which may be known with certainty or only in terms of a probability distribution. The goals for adaptation of inner to outer environment are defined by a utility function—a function, usually scalar, of the command variables and environmental parameters—perhaps supplemented by a number of constraints (inequalities, say, between functions of the command variables and environmental parameters). The optimization problem is to find an admissible set of values of the command variables, compatible with the constraints, that maximize the utility function for the given values of the environmental parameters. (In the probabilistic case we might say, "maximize the expected value of the utility function," for instance, instead of "maximize the utility function.")

A stock application of this paradigm is the so-called "diet problem" shown in figure 6. A list of foods is provided, the command variables being quantities of the various foods to be included in the diet. The environmental parameters are the prices and nutritional contents (calories, vitamins, minerals, and so on) of each of the foods. The utility function is the cost (with a minus sign attached) of the diet, subject to the constraints, say, that it not contain more than 2,000 calories per day, that it

Logical Terms		*Example:* *The diet problem*
Command variables	("Means")	Quantities of foods
Fixed parameters	("Laws")	Prices of foods Nutritional contents
Constraints Utility function	("Ends")	Nutritional requirements —Cost of diet

Constraints characterize the inner environment; parameters characterize the outer environment.

Problem: Given the constraints and fixed parameters, find values of the command variables that maximize utility.

Figure 6
The paradigm for imperative logic

meet specified minimum needs for vitamins and minerals, and that rutabaga not be eaten more than once a week. The constraints may be viewed as characterizing the inner environment. The problem is to select the quantities of foods that will meet the nutritional requirements and side conditions at the given prices for the lowest cost.

The diet problem is a simple example of a class of problems that are readily handled, even when the number of variables is exceedingly large, by the mathematical formalism known as linear programming. I shall come back to the technique a little later. My present concern is with the logic of the matter.

Since the optimization problem, once formalized, is a standard mathematical problem—to maximize a function subject to constraints—it is evident that the logic used to deduce the answer is the standard logic of the predicate calculus on which mathematics rests. How does the formalism avoid making use of a special logic of imperatives? It does so by dealing with sets of *possible worlds:* First consider all the possible worlds that meet the constraints of the outer environment; then find the particular world in the set that meets the remaining constraints of the goal and maximizes the utility function. The logic is exactly the same as if we were to adjoin the goal constraints and the maximization requirement, as new "natural laws," to the existing natural laws embodied in the environmental conditions.[3] We simply ask what values the command variables *would* have in a world meeting all these conditions and conclude that these are the values the command variables *should* have.

Computing the Optimum

Our discussion thus far has already provided us with two central topics for the curriculum in the science of design:

1. *Utility theory and statistical decision theory as a logical framework for rational choice among given alternatives.*
2. *The body of techniques for actually deducing which of the available alternatives is the optimum.*

Only in trivial cases is the computation of the optimum alternative an easy matter (Recall Chapter 2). If utility theory is to have application to real-life design problems, it must be accompanied by tools for actually making the computations. The dilemma of the rational chess player is familiar to all. The optimal strategy in chess is easily demonstrated: simply assign a value of +1 to a win, 0 to a draw, −1 to a loss; consider all possible courses of play; minimax backward from the outcome of each, assuming each player will take the most favorable move at any given point. This procedure will determine what move to make now. The only trouble is that the computations required are astronomical (the number 10^{120} is often mentioned in this context) and hence cannot be carried out—not by humans, not by existing computers, not by prospective computers.

A theory of design as applied to the game of chess would encompass not only the utopian minimax principle but also some practicable procedures for finding good moves in actual board positions in real time, within the computational capacities of real human beings or real computers. The best procedures of this kind that exist today are still those stored in the memories of grandmasters, having the characteristics I described in chapters 3 and 4. But there are now several computer programs that can rather regularly defeat all but a few of the strongest human grandmasters. Even these programs do not possess anything like the chess knowledge of human masters, but succeed by a combination of brute-force computation (sometimes hundreds of millions of variations are analysed) with a good deal of "book" knowledge of opening variations and a reasonably sophisticated criterion function for evaluating positions.

The second topic then for the curriculum in the science of design consists in the efficient computational techniques that are available for actually finding optimum courses of action in real situations, or reasonable approximations to real situations. As I mentioned in chapter 2, that topic has a number of important components today, most of them developed—at least to the level of practical application—within the past years. These include linear programming theory, dynamic programming, geometric programming, queuing theory, and control theory.

Finding Satisfactory Actions

The subject of computational techniques need not be limited to optimization. Traditional engineering design methods make much more use of inequalities—specifications of satisfactory performance—than of maxima and minima. So-called "figures of merit" permit comparison between designs in terms of "better" and "worse" but seldom provide a judgment of "best." For example, I may cite the root-locus methods employed in the design of control systems.

Since there did not seem to be any word in English for decision methods that look for good or satisfactory solutions instead of optimal ones, some years ago I introduced the term "satisficing" to refer to such procedures. Now no one in his right mind will satisfice if he can equally well optimize; no one will settle for good or better if he can have best. But that is not the way the problem usually poses itself in actual design situations.

In chapter 2 I argued that in the real world we usually do not have a choice between satisfactory and optimal solutions, for we only rarely have

3. The use of the notion of "possible worlds" to embed the logic of imperatives in declarative logic goes back at least to Jørgen Jørgensen, "Imperatives and Logic," *Erkenntnis*, 7(1937-1938):288–296. See also my *Administrative Behavior* (New York: Macmillan, 1947), chapter 3. Typed logics can be used to distinguish, as belonging to different types, statements that are true under different conditions (i.e., in different possible worlds), but, as my example shows, even this device is not usually needed. Each new equation or constraint we introduce into a system reduces the set of possible states to a subset of those previously possible.

a method of finding the optimum. Consider, for example, the well-known combinatorial problem called the traveling salesman problem: given the geographical locations of a set of cities, find the routing that will take a salesman to all the cities with the shortest mileage.[4] For this problem there is a straightforward optimizing algorithm (analogous to the minimax algorithm for chess): try all possible routings, and pick the shortest. But for any considerable number of cities, the algorithm is computationally infeasible (the number of routes through N cities will be $N!$). Although some ways have been found for cutting down the length of the search, no algorithm has been discovered sufficiently powerful to solve the traveling salesman problem with a tolerable amount of computing for a set of, say, fifty cities.

Rather than keep our salesman at home, we shall prefer of course to find a satisfactory, if not optimal, routing for him. Under most circumstances, common sense will probably arrive at a fairly good route, but an even better one can often be found by one or another of several heuristic methods.

An earmark of all these situations where we satisfice for inability to optimize is that, although the set of available alternatives is "given" in a certain abstract sense (we can define a generator guaranteed to generate all of them eventually), it is not "given" in the only sense that is practically relevant. We cannot within practicable computational limits generate all the admissible alternatives and compare their respective merits. Nor can we recognize the best alternative, even if we are fortunate enough to generate it early, until we have seen all of them. We satisfice by looking for alternatives in such a way that we can generally find an acceptable one after only moderate search.

Now in many satisficing situations, the expected length of search for an alternative meeting specified standards of acceptability depends on how high the standards are set, but it depends hardly at all on the total size of the universe to be searched. The time required for a search through a haystack for a needle sharp enough to sew with depends on the density of distribution of sharp needles but not on the total size of the stack.

4. "The traveling salesman problem" and a number of closely analogous combinatorial problems—such as the "warehouse location problem"—have considerable practical importance, for instance, in siting central power stations for an interconnected grid.

Hence, when we use satisficing methods, it often does not matter whether or not the total set of admissible alternatives is "given" by a formal but impracticable algorithm. It often does not even matter how big that set is. For this reason satisficing methods may be extendable to design problems in that broad range where the set of alternatives is not "given" even in the quixotic sense that it is "given" for the traveling salesman problem. Our next task is to examine this possibility.

The Logic of Design: Finding Alternatives

When we take up the case where the design alternatives are not given in any constructive sense but must be synthesized, we must ask once more whether any new forms of reasoning are involved in the synthesis, or whether again the standard logic of declarative statements is all we need.

In the case of optimization we asked: "Of all possible worlds (those attainable for some admissible values of the action variables), which is the best (yields the highest value of the criterion function)?" As we saw, this is a purely empirical question, calling only for facts and ordinary declarative reasoning to answer it.

In this case, where we are seeking a satisfactory alternative, once we have found a candidate we can ask: "Does this alternative satisfy all the design criteria?" Clearly this is also a factual question and raises no new issues of logic. But how about the process of *searching* for candidates? What kind of logic is needed for the search?

Means-Ends Analysis

The condition of any goal-seeking system is that it is connected to the outside environment through two kinds of channels: the afferent, or sensory, channels through which it receives information about the environment and the efferent, or motor, channels through which it acts on the environment.[5] The system must have some means of storing in its memory information about states of the world—afferent, or sensory, information—

5. Notice that we are not saying that the two kinds of channels operate independently of each other, since they surely do not in living organisms, but that we can distinguish conceptually, and to some extent neurologically, between the incoming and outgoing flows.

and information about actions—efferent, or motor, information. Ability to attain goals depends on building up associations, which may be simple or very complex, between particular changes in states of the world and particular actions that will (reliably or not) bring these changes about. In chapter 4 we described these associations as productions.

Except for a few built-in reflexes, an infant has no basis for correlating its sensory information with its actions. A very important part of its early learning is that particular actions or sequences of actions will bring about particular changes in the state of the sensed world. Until the infant builds up this knowledge, the world of sense and the motor world are two entirely separate, entirely unrelated worlds. Only as it begins to acquire experience as to how elements of the one relate to elements of the other can it act purposefully on the world.

The computer problem-solving program called GPS, designed to model some of the main features of human problem solving, exhibits in stark form how goal-directed action depends on building this kind of bridge between the afferent and the efferent worlds. On the afferent, or sensory, side, GPS must be able to represent desired situations or desired objects as well as the present situation. It must be able also to represent *differences* between the desired and the present. On the efferent side, GPS must be able to represent *actions* that change objects or situations. To behave purposefully, GPS must be able to select from time to time those particular actions that are likely to remove the particular differences between desired and present states that the system detects. In the machinery of GPS, this selection is achieved through a *table of connections*, which associates with each kind of detectable difference those actions that are relevant to reducing that difference. These are its associations, in the form of productions, which relate the afferent to the efferent world. Since reaching a goal generally requires a sequence of actions, and since some attempts may be ineffective, GPS must also have means for detecting the progress it is making (the changes in the differences between the actual and the desired) and for trying alternate paths.

The Logic of Search

GPS then is a system that searches selectively through a (possibly large) environment in order to discover and assemble sequences of actions that will lead it from a given situation to a desired situation. What are the rules of logic that govern such a search? Is anything more than standard logic involved? Do we require a modal logic to rationalize the process?

Standard logic would seem to suffice. To represent the relation between the afferent and the efferent worlds, we conceive GPS as moving through a large maze. The nodes of the maze represent situations, described afferently; the paths joining one node to another are the actions, described as motor sequences, that will transform the one situation into the other. At any given moment GPS is always faced with a single question: "What action shall I try next?" Since GPS has some imperfect knowledge about the relations of actions to changes in the situation, this becomes a question of choice under uncertainty of a kind already discussed in a previous section.

It is characteristic of the search for alternatives that the solution, the complete action that constitutes the final design, is built from a sequence of component actions. The enormous size of the space of alternatives arises out of the innumerable ways in which the component actions, which need not be very numerous, can be combined into sequences.

Much is gained by considering the component actions in place of the sequences that constitute complete actions, because the situation when viewed afferently usually factors into components that match at least approximately the component actions derived from an efferent factorization. The reasoning implicit in GPS is that, if a desired situation differs from a present situation by differences $D_1, D_2, \ldots, D_n$, and if action A_1 removes differences of type D_1, action A_2 removes differences of type D_2, and so on, then the present situation can be transformed into the desired situation by performing the sequence of actions $A_1A_2 \ldots A_n$.

This reasoning is by no means valid in terms of the rules of standard logic in all possible worlds. Its validity requires some rather strong assumptions about the independence of the effects of the several actions on the several differences. One might say that the reasoning is valid in worlds that are "additive" or "factorable" in a certain sense. (The air of paradox about the cat-dog and needle-thread examples cited earlier arises precisely from the nonadditivity of the actions in these two cases. The first is, in economists' language, a case of decreasing returns; the second, a case of increasing returns.)

Now the real worlds to which problem solvers and designers address themselves are seldom completely additive in this sense. Actions have side consequences (may create new differences) and sometimes can only be taken when certain side conditions are satisfied (call for removal of other differences before they become applicable). Under these circumstances one can never be certain that a partial sequence of actions that accomplishes *certain* goals can be augmented to provide a solution that satisfies *all* the conditions and attains *all* the goals (even though they be satisficing goals) of the problem.

For this reason problem-solving systems and design procedures in the real world do not merely *assemble* problem solutions from components but must *search* for appropriate assemblies. In carrying out such a search, it is often efficient to divide one's eggs among a number of baskets—that is, not to follow out one line until it succeeds completely or fails definitely but to begin to explore several tentative paths, continuing to pursue a few that look most promising at a given moment. If one of the active paths begins to look less promising, it may be replaced by another that had previously been assigned a lower priority.

Our discussion of design when the alternatives are not given has yielded at least three additional topics for instruction in the science of design:

3. *Adaptation of standard logic to the search for alternatives.* Design solutions are sequences of actions that lead to possible worlds satisfying specified constraints. With satisficing goals the sought-for possible worlds are seldom unique; the search is for *sufficient*, not *necessary*, actions for attaining goals.

4. *The exploitation of parallel, or near-parallel, factorizations of differences.* Means-end analysis is an example of a broadly applicable problem-solving technique that exploits this factorization.

5. *The allocation of resources for search to alternative, partly explored action sequences.* I should like to elaborate somewhat on this last-mentioned topic.

Design as Resource Allocation

There are two ways in which design processes are concerned with the allocation of resources. First, conservation of scarce resources may be one of the criteria for a satisfactory design. Second, the design process itself involves management of the resources of the designer, so that his efforts will not be dissipated unnecessarily in following lines of inquiry that prove fruitless.

There is nothing special that needs to be said here about resource conservation—cost minimization, for example, as a design criterion. Cost minimization has always been an implicit consideration in the design of engineering structures, but until a few years ago it generally *was* only implicit, rather than explicit. More and more cost calculations have been brought explicitly into the design procedure, and a strong case can be made today for training design engineers in that body of technique and theory that economists know as "cost-benefit analysis."

An Example from Highway Design

The notion that the costs of designing must themselves be considered in guiding the design process began to take root only as formal design procedures have developed, and it still is not universally applied. An early example, but still a very good one, of incorporating design costs in the design process is the procedure, developed by Marvin L. Manheim as a doctoral thesis at MIT, for solving highway location problems.[6]

Manheim's procedure incorporates two main notions: first, the idea of specifying a design progressively from the level of very general plans down to determining the actual construction; second, the idea of attaching values to plans at the higher levels as a basis for deciding which plans to pursue to levels of greater specificity.

In the case of highway design the higher-level search is directed toward discovering "bands of interest" within which the prospects of finding a good specific route are promising. Within each band of interest one or more locations is selected for closer examination. Specific designs are then developed for particular locations. The scheme is not limited of course to this specific three-level division, but it can be generalized as appropriate.

Manheim's scheme for deciding which alternatives to pursue from one level to the next is based on assigning costs to each of the design activities and estimating highway costs for each of the higher-level plans. The

6. Marvin L. Manheim, *Hierarchical Structure: A Model of Design and Planning Processes* (Cambridge: The MIT Press, 1966).

highway cost associated with a plan is a prediction of what the cost would be for the actual route if that plan were particularized through subsequent design activity. In other words, it is a measure of how "promising" a plan is. Those plans are then pursued to completion that look most promising after the prospective design costs have been offset against them.

In the particular method that Manheim describes, the "promise" of a plan is represented by a probability distribution of outcomes that would ensue if it were pursued to completion. The distribution must be estimated by the engineer—a serious weakness of the method—but, once estimated, it can be used within the framework of Bayesian decision theory. The particular probability model used is not the important thing about the method; other methods of valuation without the Bayesian superstructure might be just as satisfactory.

In the highway location procedure the evaluation of higher-level plans performs two functions. First, it answers the question, "Where shall I search next?" Second, it answers the question, "When shall I stop the search and accept a solution as satisfactory?" Thus it is both a steering mechanism for the search and a satisficing criterion for terminating the search.

Schemes for Guiding Search

Let us generalize the notion of schemes for guiding search activity beyond Manheim's specific application to a highway location problem and beyond his specific guidance scheme based on Bayesian decision theory. Consider the typical structure of a problem-solving program. The program begins to search along possible paths, storing in memory a "tree" of the paths it has explored. Attached to the end of each branch—each partial path—is a number that is supposed to express the "value" of that path.

But the term "value" is really a misnomer. A partial path is not a solution of the problem, and a path has a "true" value of zero unless it leads toward a solution. Hence it is more useful to think of the values as estimates of the gain to be expected from further search along the path than to think of them as "values" in any more direct sense. For example, it may be desirable to attach a relatively high value to a partial exploration that *may* lead to a very good solution but with a low probability. If the prospect fades on further exploration, only the cost of the search has been lost. The disappointing outcome need not be accepted, but an alternative

path may be taken instead. Thus the scheme for attaching values to partial paths may be quite different from the evaluation function for proposed complete solutions.[7]

When we recognize that the purpose of assigning values to incomplete paths is to guide the choice of the next point for exploration, it is natural to generalize even further. All kinds of information gathered in the course of search may be of value in selecting the next step in search. We need not limit ourselves to valuations of partial search paths.

For example, in a chess-playing program an exploration may generate a continuation move different from any that was proposed by the initial move generator. Whatever the context—the branch of the search tree—on which the move was actually generated, it can now be removed from the context and considered in the context of other move sequences. Such a scheme was added on a limited basis by Baylor to MATER, a program for discovering check-mating combinations in chess, and it proved to enhance the program's power significantly.[8]

Thus search processes may be viewed—as they have been in most discussions of problem solving—as processes for seeking a problem solution. But they can be viewed more generally as processes for gathering information about problem structure that will ultimately be valuable in discovering a problem solution. The latter viewpoint is more general than the former in a significant sense, in that it suggests that information obtained along any particular branch of a search tree may be used in many contexts besides the one in which it was generated. Only a few problem-solving programs exist today that can be regarded as moving even a modest distance from the earlier, more limited viewpoint to the newer one.[9]

7. That this point is not obvious can be seen from the fact that most chess-playing programs have used similar or identical evaluation procedures both to guide search and to evaluate the positions reached at the ends of paths.

8. George W. Baylor and Herbert A. Simon, "A Chess Mating Combinations Program," *Proceedings of the Spring Joint Computer Conference*, Boston, April 26–28, (1966):431–447 (Washington: Spartan Books, 1966), reprinted in *Models of Thought*, chapter 4.3.

9. A formal theory of the optimal choice of search paths can be found in H. A. Simon and J. B. Kadane, "Optimal Problem-Solving Search: All-or-none Solutions," *Artificial Intelligence*, 6(1975):235–247.

The Shape of the Design: Hierarchy

In my first chapter I gave some reasons why complex systems might be expected to be constructed in a hierarchy of levels, or in a boxes-within-boxes form. The basic idea is that the several components in any complex system will perform particular subfunctions that contribute to the over-all function. Just as the "inner environment" of the whole system may be defined by describing its functions, without detailed specification of its mechanisms, so the "inner environment" of each of the subsystems may be defined by describing the functions of that subsystem, without detailed specification of *its* submechanisms.[10]

To design such a complex structure, one powerful technique is to discover viable ways of decomposing it into semi-independent components corresponding to its many functional parts. The design of each component can then be carried out with some degree of independence of the design of others, since each will affect the others largely through its function and independently of the details of the mechanisms that accomplish the function.[11]

There is no reason to expect that the decomposition of the complete design into functional components will be unique. In important instances there may exist alternative feasible decompositions of radically different kinds. This possibility is well known to designers of administrative organizations, where work can be divided up by subfunctions, by subprocesses, by subareas, and in other ways. Much of classical organization theory in fact was concerned precisely with this issue of alternative decompositions of a collection of interrelated tasks.

The Generator-Test Cycle

One way of considering the decomposition, but acknowledging that the interrelations among the components cannot be ignored completely, is to think of the design process as involving, first, the generation of alternatives and, then, the testing of these alternatives against a whole array of requirements and constraints. There need not be merely a single generate-test cycle, but there can be a whole nested series of such cycles. The generators implicitly define the decomposition of the design problem, and the tests guarantee that important indirect consequences will be noticed and weighed. Alternative decompositions correspond to different ways of dividing the responsibilities for the final design between generators and tests.

To take a greatly oversimplified example, a series of generators may generate one or more possible outlines and schemes of fenestration for a building, while tests may be applied to determine whether needs for particular kinds of rooms can be met within the outlines generated. Alternatively the generators may be used to evolve the structure of rooms, while tests are applied to see whether they are consistent with an acceptable over-all shape and design. The house can be designed from the outside in or from the inside out.[12]

Alternatives are also open, in organizing the design process, as to how far development of possible subsystems will be carried before the over-all coordinating design is developed in detail, or vice-versa, how far the over-all design should be carried before various components, or possible components, are developed. These alternatives of design are familiar to architects. They are familiar also to composers, who must decide how far the architectonics of a musical structure will be evolved before some of the component musical themes and other elements have been invented. Computer programmers face the same choices, between working downward from executive routines to subroutines or upward from component subroutines to a coordinating executive.

A theory of design will include principles for deciding such questions of precedence and sequence in the design process. As examples, the approach to designing computer programs called structured programming is concerned in considerable part with attending to design subproblems

10. I have developed this argument at greater length in my essay "The Architecture of Complexity," chapter 8.

11. For a recent discussion of functional analysis in design, see Clive L. Dym, *Engineering Design* (New York, NY: Cambridge University Press, 1994), pp. 134–139.

12. I am indebted to John Grason for many ideas on the topic of this section. J. Grason, "Fundamental Description of a Floor Plan Design Program," EDRA1, *Proceedings of the First Environmental Design Association Conference*, H. Sanoff and S. Cohn (eds.), North Carolina State University, 1970.

in the proper order (usually top-down); and much instruction in schools of architecture focuses on the same concerns.

Process as a Determinant of Style

When we recall that the process will generally be concerned with finding a satisfactory design, rather than an optimum design, we see that sequence and the division of labor between generators and tests can affect not only the efficiency with which resources for designing are used but also the nature of the final design as well. What we ordinarily call "style" may stem just as much from these decisions about the design process as from alternative emphases on the goals to be realized through the final design.[13] An architect who designs buildings from the outside in will arrive at quite different buildings from one who designs from the inside out, even though both of them might agree on the characteristics that a satisfactory building should possess.

When we come to the design of systems as complex as cities, or buildings, or economies, we must give up the aim of creating systems that will optimize some hypothesized utility function, and we must consider whether differences in style of the sort I have just been describing do not represent highly desirable variants in the design process rather than alternatives to be evaluated as "better" or "worse." Variety, within the limits of satisfactory constraints, may be a desirable end in itself, among other reasons, because it permits us to attach value to the search as well as its outcome—to regard the design process as itself a valued activity for those who participate in it.

We have usually thought of city planning as a means whereby the planner's creative activity could build a system that would satisfy the needs of a populace. Perhaps we should think of city planning as a valuable creative activity in which many members of a community can have the opportunity of participating—if we have wits to organize the process that way. I shall have more to say on these topics in the next chapter.

However that may be, I hope I have illustrated sufficiently that both the shape of the design and the shape and organization of the design process

13. H. A. Simon, "Style in Design," *Proceedings of the 2nd Annual Conference of the Environmental Design Research Association,* Pittsburgh, PA: Carnegie Mellon University (1971), pp. 1–10.

are essential components of a theory of design. These topics constitute the sixth item in my proposed curriculum in design:

6. *The organization of complex structures and its implication for the organization of design processes.*

Representation of the Design

I have by no means surveyed all facets of the emerging science of design. In particular I have said little about the influence of problem representation on design. Although the importance of the question is recognized today, we are still far from a systematic theory of the subject—in particular, a theory that would tell us how to generate effective problem representations.[14] I shall cite one example, to make clear what I mean by "representation."

Here are the rules of a game, which I shall call number scrabble. The game is played by two people with nine cards—let us say the ace through the nine of hearts. The cards are placed in a row, face up, between the two players. The players draw alternately, one at a time, selecting any one of the cards that remain in the center. The aim of the game is for a player to make up a "book," that is, a set of exactly three cards whose spots add to 15, before his opponent can do so. The first player who makes a book wins; if all nine cards have been drawn without either player making a book, the game is a draw.

What is a good strategy in this game? How would you go about finding one? If the reader has not already discovered it for himself, let me show how a change in representation will make it easy to play the game well. The magic square here, which I introduced in the third chapter, is made up of the numerals from 1 through 9.

4	9	2
3	5	7
8	1	6

14. As examples of current thinking about representation see chapters 5 ("Representing Designed Artifacts") and 6 ("Representing Design Processes") in C. L. Dym, *op. cit.,* and chapter 6 ("Representation in Design") in Ömer Akin, *op. cit.* For a more general theoretical discussion, see R. E. Korf, "Toward a Model of Representational Changes," *Artificial Intelligence,* 14(1980):41–78.

Each row, column, or diagonal adds to 15, and every triple of these numerals that add to 15 is a row, column, or diagonal of the magic square. From this, it is obvious that "making a book" in number scrabble is equivalent to getting "three in a row" in the game of tic-tac-toe. But most people know how to play tic-tac-toe well, hence can simply transfer their usual strategy to number scrabble.[15]

Problem Solving as Change in Representation

That representation makes a difference is a long-familiar point. We all believe that arithmetic has become easier since Arabic numerals and place notation replaced Roman numerals, although I know of no theoretic treatment that explains why.

That representation makes a difference is evident for a different reason. All mathematics exhibits in its conclusions only what is already implicit in its premises, as I mentioned in a previous chapter. Hence all mathematical derivation can be viewed simply as change in representation, making evident what was previously true but obscure.

This view can be extended to all of problem solving—solving a problem simply means representing it so as to make the solution transparent.[16] If the problem solving could actually be organized in these terms, the issue of representation would indeed become central. But even if it cannot—if this is too exaggerated a view—a deeper understanding of how representations are created and how they contribute to the solution of problems will become an essential component in the future theory of design.

Spatial Representation

Since much of design, particularly architectural and engineering design, is concerned with objects or arrangements in real Euclidean two-

15. Number scrabble is not the only isomorph of tic-tac-toe. John A. Michon has described another, JAM, which is the dual of tic-tac-toe in the sense of projective geometry. That is, the rows, columns, and diagonals of tic-tac-toe become points in JAM, and the squares of the former become line segments joining the points. The game is won by "jamming" all the segments through a point—a move consists of seizing or jamming a single segment. Other isomorphs of tic-tac-toe are known as well.

16. Saul Amarel, "On the Mechanization of Creative Processes," IEEE *Spectrum* 3(April 1966):112–114.

dimensional or three-dimensional space, the representation of space and of things in space will necessarily be a central topic in a science of design. From our previous discussion of visual perception, it should be clear that "space" inside the head of the designer or the memory of a computer may have very different properties from a picture on paper or a three-dimensional model.

These representational issues have already attracted the attention of those concerned with computer-aided design—the cooperation of human and computer in the design process. As a single example, I may mention Ivan Sutherland's pioneering SKETCHPAD program which allowed geometric shapes to be represented and conditions to be placed on these shapes in terms of constraints, to which they then conformed.[17]

Geometric considerations are also prominent in the attempts to automate completely the design, say, of printed or etched circuits, or of buildings. Grason, for example, in a system for designing house floor plans, constructs an internal representation of the layout that helps one decide whether a proposed set of connections among rooms, selected to meet design criteria for communication, and so on, can be realized in a plane.[18]

The Taxonomy of Representation

An early step toward understanding any set of phenomena is to learn what kinds of things there are in the set—to develop a taxonomy. This step has not yet been taken with respect to representations. We have only a sketchy and incomplete knowledge of the different ways in which problems can be represented and much less knowledge of the significance of the differences.

In a completely pragmatic vein we know that problems can be described verbally, in natural language. They often can be described mathematically, using standard formalisms of algebra, geometry, set theory, analysis, or topology. If the problems relate to physical objects, they (or their solutions) can be represented by floor plans, engineering drawings,

17. I. E. Sutherland, "SKETCHPAD, A Man-Machine Graphical Communication System," *Proceedings, AFIPS Spring Joint Computer Conference, 1963* (Baltimore: Spartan Books), pp. 329–346.

18. See also C. E. Pfefferkorn, "The Design Problem Solver: A System for Designing Equipment or Furniture Layours," in C. M. Eastman (ed.), *Spatial Synthesis in Computer-Aided Building Design* (London: Applied Science Publishers, 1975).

renderings, or three-dimensional models. Problems that have to do with actions can be attacked with flow charts and programs.

Other items most likely will need to be added to the list, and there may exist more fundamental and significant ways of classifying its members. But even though our classification is incomplete, we are beginning to build a theory of the properties of these representations. The growing theories of computer architectures and programming languages—for example, the work on functional languages and object-oriented languages—illustrate some of the directions that a theory of representations can take. There has also been closely parallel progress, some of it reviewed in chapters 3 and 4, toward understanding the human use of representations in thinking. These topics begin to provide substance for the final subject in our program on the theory of design:

7. *Alternative representations for design problems.*

Summary—Topics in The Theory of Design

My main goal in this chapter has been to show that there already exist today a number of components of a theory of design and a substantial body of knowledge, theoretical and empirical, relating to each. As we draw up our curriculum in design—in the science of the artificial—to take its place by the side of natural science in the whole engineering curriculum, it includes at least the following topics:

THE EVALUATION OF DESIGNS

1. Theory of evaluation: utility theory, statistical decision theory
2. Computational methods:
   a. Algorithms for choosing *optimal* alternatives such as linear programming computations, control theory, dynamic programming
   b. Algorithms and heuristics for choosing *satisfactory* alternatives
3. THE FORMAL LOGIC OF DESIGN: imperative and declarative logics

THE SEARCH FOR ALTERNATIVES

4. Heuristic search: factorization and means-ends analysis
5. Allocation of resources for search
6. THEORY OF STRUCTURE AND DESIGN ORGANIZATION: hierarchic systems
7. REPRESENTATION OF DESIGN PROBLEMS

In small segments of the curriculum—the theory of evaluation, for example, and the formal logic of design—it is already possible to organize the instruction within a framework of systematic, formal theory. In many other segments the treatment would be more pragmatic, more empirical.

But nowhere do we need to return or retreat to the methods of the cookbook that originally put design into disrepute and drove it from the engineering curriculum. For there exist today a considerable number of examples of actual design processes, of many different kinds, that have been defined fully and cast in the metal, so to speak, in the form of running computer programs: optimizing algorithms, search procedures, and special-purpose programs for designing motors, balancing assembly lines, selecting investment portfolios, locating warehouses, designing highways, diagnosing and treating diseases, and so forth.[19]

Because these computer programs describe complex design processes in complete, painstaking detail, they are open to full inspection and analysis, or to trial by simulation. They constitute a body of empirical phenomena to which the student of design can address himself and which he can seek to understand. There is no question, since these programs exist, of the design process hiding behind the cloak of "judgment" or "experience." Whatever judgment or experience was used in creating the programs must now be incorporated in them and hence be observable. The programs are the tangible record of the variety of schemes that man has devised to explore his complex outer environment and to discover in that environment the paths to his goals.

Role of Design in the Life of the Mind

I have called my topic "the theory of design" and my curriculum a "program in design." I have emphasized its role as complement to the natural

19. A number of these programs are described in Dym, *op. cit.*, and others are discussed in a forthcoming book on *Engineering Design in the Large*, written by faculty associated with the Engineering Design Research Center at Carnegie Mellon University. Dym concludes each chapter of his book with a commentary on other relevant publications. Dym's book has a bibliography of more than 200 items, a majority of them referring to specific design projects and systems; its extent gives some indication of the rate at which the science of design is now progressing.

science curriculum in the total training of a professional engineer—or of any professional whose task is to solve problems, to choose, to synthesize, to decide.

But there is another way in which the theory of design may be viewed in relation to other knowledge. My third and fourth chapters were chapters on psychology—specifically on man's relation to his biological inner environment. The present chapter may also be construed as a chapter on psychology: on man's relation to the complex outer environment in which he seeks to survive and achieve.

All three chapters, so construed, have import that goes beyond the professional work of the person we have called the "designer." Many of us have been unhappy about the fragmentation of our society into two cultures. Some of us even think there are not just two cultures but a large number of cultures. If we regret that fragmentation, then we must look for a common core of knowledge that can be shared by the members of all cultures—a core that includes more significant topics than the weather, sports, automobiles, the care and feeding of children, or perhaps even politics. A common understanding of our relation to the inner and outer environments that define the space in which we live and choose can provide at least part of that significant core.

This may seem an extravagant claim. Let me use the realm of music to illustrate what I mean. Music is one of the most ancient of the sciences of the artificial, and was so recognized by the Greeks. Anything I have said about the artificial would apply as well to music, its composition or its enjoyment, as to the engineering topics I have used for most of my illustrations.

Music involves a formal pattern. It has few (but important) contacts with the inner environment; that is, it is capable of evoking strong emotions, its patterns are detectable by human listeners, and some of its harmonic relations can be given physical and physiological interpretations (though the aesthetic import of these is debatable). As for the outer environment, when we view composition as a problem in design, we encounter just the same tasks of evaluation, of search for alternatives, and of representation that we do in any other design problem. If it pleases us, we can even apply to music some of the same techniques of automatic design by computer that have been used in other fields of design. If

computer-composed music has not yet reached notable heights of aesthetic excellence, it deserves, and has already received, serious attention from professional composers and analysts, who do not find it written in tongues alien to them.[20]

Undoubtedly there are tone-deaf engineers, just as there are mathematically ignorant composers. Few engineers and composers, whether deaf, ignorant, or not, can carry on a mutually rewarding conversation about the content of each other's professional work. What I am suggesting is that they *can* carry on such a conversation about design, can begin to perceive the common creative activity in which they are both engaged, can begin to share their experiences of the creative, professional design process.

Those of us who have lived close to the development of the modern computer through gestation and infancy have been drawn from a wide variety of professional fields, music being one of them. We have noticed the growing communication among intellectual disciplines that takes place around the computer. We have welcomed it, because it has brought us into contact with new worlds of knowledge—has helped us combat our own multiple-cultures isolation. This breakdown of old disciplinary boundaries has been much commented upon, and its connection with computers and the information sciences often noted.

But surely the computer, as a piece of hardware, or even as a piece of programmed software, has nothing to do directly with the matter. I have already suggested a different explanation. The ability to communicate across fields—the common ground—comes from the fact that all who use computers in complex ways are using computers to design or to participate in the process of design. Consequently we as designers, or as designers of design processes, have had to be explicit as never before about what is involved in creating a design and what takes place while the creation is going on.

The real subjects of the new intellectual free trade among the many cultures are our own thought processes, our processes of judging, deciding,

20. L. A. Hillier and L. M. Isaacson's *Experimental Music* (New York: McGraw-Hill, 1959), reporting experiments begun more than four decades ago, still provides a good introduction to the subject of musical composition, viewed as design. See also Walter R. Reitman, *Cognition and Thought* (New York: Wiley, 1965), chapter 6, "Creative Problem Solving: Notes from the Autobiography of a Fugue."

choosing, and creating. We are importing and exporting from one intellectual discipline to another ideas about how a serially organized information-processing system like a human being—or a computer, or a complex of men and women and computers in organized cooperation—solves problems and achieves goals in outer environments of great complexity.

The proper study of mankind has been said to be man. But I have argued that people—or at least their intellective component—may be relatively simple, that most of the complexity of their behavior may be drawn from their environment, from their search for good designs. If I have made my case, then we can conclude that, in large part, the proper study of mankind is the science of design, not only as the professional component of a technical education but as a core discipline for every liberally educated person.

6

Social Planning: Designing the Evolving Artifact

In chapter 5 I surveyed some of the modern tools of design that are used by planners and artificers. Even before most of these tools were available to them, ambitious planners often took whole societies and their environments as systems to be refashioned. Some recorded their utopias in books—Plato, Sir Thomas More, Marx. Others sought to realize their plans by social revolution in America, France, Russia, China. Many or most of the large-scale designs have centered on political and economic arrangements, but others have focused on the physical environment—river development plans, for example, reaching from ancient Egypt to the Tennessee Valley to the Indus and back to today's Nile.

As we look back on such design efforts and their implementation, and as we contemplate the tasks of design that are posed in the world today, our feelings are very mixed. We are energized by the great power our technological knowledge bestows on us. We are intimidated by the magnitude of the problems it creates or alerts us to. We are sobered by the very limited success—and sometimes disastrous failure—of past efforts to design on the scale of whole societies. We ask, "If we can go to the Moon, why can't we . . . ?"—not expecting an answer, for we know that going to the Moon was a simple task indeed, compared with some others we have set for ourselves, such as creating a humane society or a peaceful world. Wherein lies the difference?

Going to the Moon was a complex matter along only one dimension: it challenged our technological capabilities. Though it was no mean accomplishment, it was achieved in an exceedingly cooperative environment, employing a single new organization, NASA, that was charged with a single, highly operational goal. With enormous resources provided to

Bartneck, C. (2007). Design Methodology is not Design Science. Proceedings of the CHI 2007 Workshop: Converging on a "Science of Design" through the Synthesis of Design Methodologies, San Jose.

Design Methodology is not Design Science

Christoph Bartneck
Department of Industrial Design
Eindhoven University of Technology
Den Dolech 2, 5600MB Eindhoven
The Netherlands
christoph@bartneck.de

ABSTRACT
This paper argues that design methodology cannot become the science of design. A method does not constitute a science. 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.

Author Keywords
Design, science, methodology

ACM Classification Keywords
H5.0. Information interfaces and presentation (e.g., HCI): General.

INTRODUCTION

It is custom to submit papers to workshops that support the fundamental ideas of the workshop. When I read about the "Converging on a Science of Design through the Synthesis of Design Methodologies" workshop I felt obliged to do the opposite. In this paper I will challenge the goal of this workshop to converge on a science of the design through converging of design methodologies. This will probably raise the eyebrows of the organizers and maybe also of the workshop participants. However, it is the nature of science that truth remains truth, independently of what people think of it. This quest for truth is fueled through dialectic discussion and I hope that this manuscript will spark an open dialogue about the goal and status of design in the HCI community.

DISCUSSION

Besides the workshop title, the description also states that the workshop will focus on design methodology and that it will "make a contribution to the establishment of design as a science." While the definition of a design science is a noble goal, the method chosen appears flawed. Science consists of a method to observer and abstract reality into models that are then used to explain and predict reality (see Figure 1). 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.

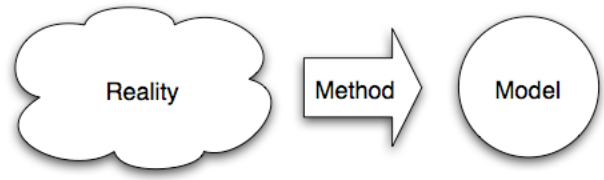


Figure 1: scientific process

The method of science is to some degree universal and is often referred to as the 'scientific method'. The scientific method is a body of techniques for investigating phenomena and acquiring new knowledge, as well as for correcting and integrating previous knowledge. It is based on gathering observable, empirical, measurable evidence, subject to the principles of reasoning. Chalmer (1999) provides a fair discussion of the scientific method. However, a methodology in itself can never constitute a science. Lets take the example of the dissection method. Biologists may use dissection to analyze animals, but also butchers use it to cut steaks. The method is the same, but one results in scientific knowledge, while 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. Even converging on a specific design method cannot overcome this conceptual limitation. Again: a method does not constitute a science and design methodology cannot be the phenomena of design science. The goal of the workshop to create a design science cannot be achieved by converging on a design method.

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. The problem becomes clearer when we consider that design's prime objective lays in the intersection between artifacts and users (see Figure 2). Designers contribute to the creation of artifacts that interact with humans.

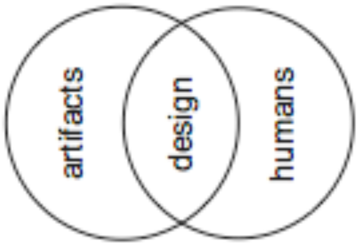


Figure 2: Framework of Design

Everything there is to know about the artifact (left side Figure 2) is available from its manufacturer. All its dimensions, material properties and functions are known. The artifacts are therefore not good phenomena to investigate. The creation of new materials and operational principles has also already been claimed by engineering and physics. Engineers also discussed rational design methodology that heavily relies on mathematics (Alexander, 1964; Simon, 1996; Vincenti, 1990). Interestingly, these rational design methodologies have not been included in the description of the workshop even though the 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 into the direction of generizability.

On the other side (right side Figure 2), understanding humans is the prime objective of medicine, anthropology and psychology. Design science would have difficulties competing. Even "Design methodology", or to be more general, "human problem solving", has already been treated as a phenomena investigated by psychologist (Dorfman, Shames, & Kihlstrom, 1996; Feist, 1994).

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.

When we take a look the body of scientific knowledge, it has been engineers again that attempted to create a consistent and logical body of knowledge (Hubka & Eder, 1996; Vincenti, 1990). As we can see, the arena of design science is filled with actors and it one may ask then why the designers in the HCI community are so keen on turning design into a science? Design has been criticized by the academic section of the HCI community to be non-scientific. An example of this conflict occurred at the 2005 SIGCHI membership meeting. The organization of the CHI2006 was discussed, which ignited a shouting match between academics and practitioners (Arnowitz & Dykstra-Erickson, 2005). Both groups defended their access to the conference through the different publication formats, such as papers sessions, panels, and case studies. At the conference itself the conflict reoccurred 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 provided to designers (Dourish, 2006). Next, Tracee Verring Wolf and Jennifer Rode defended creative design against the scientific criticism by referring to design rigor (Wolf , Rode, Sussman, & Kellogg, 2006). Both groups felt the need to defend themselves, which indicated that both had the feeling of being under attack. Trying to defend design by claiming that it is scientific may appear to be a good respond to the academic criticism, and designers are naturally attracted by the quality label of science. Chalmer (1999) pointed out that:

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" id done in a way that is intended to imply some kind or merit or special kind of reliability.

It is a noble goal to create good and reliable design, but this may not be achieved by using the scientific method and neither may the claim of a design science be a good response to the academic criticism. 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 play in the HCI community. Discussions on design methodology are a good step forward to further improve design practice. A CHI workshop is a good forum for such a discussion. However, for reasons explained above, it may not be wise to claim that this would lead to a design science. A possible better name for the workshop might have been "Converging on Good Design through the Synthesis of Design Methodologies".

REFERENCES

Alexander, C. (1964). Notes on the synthesis of form. Cambridge,: Harvard University Press.

Arnowitz, J., & Dykstra-Erickson, E. (2005). CHI and the Practitioner Dilemma. Interactions, 12(4), 5-9.

Chalmers, A. F. (1999). What is this thing called science? (3rd ed.). Indianapolis: Hackett.

Dorfman, J., Shames, V. A., & Kihlstrom, J. F. (1996). Intuition, incubation, and insight: implicit cognition in problem solving. In G. Underwood (Ed.), Implicit Cognition. Oxford, UK: Oxford University Press.

Dourish, P. (2006). Implications for design. Paper presented at the Conference Name|. Retrieved Access Date|. from URL|.

Feist, G. (1994). The affective consequences of artistic and scientific problem solving. Cognition and emotion, 8, 489-502.

Hubka, V., & Eder, W. E. (1996). Design science : introduction to needs, scope and organization of engineering design knowledge. Berlin ; New York: Springer.

Simon, H. A. (1996). The sciences of the artificial (3rd ed.). Cambridge, Mass.: MIT Press.

Vincenti, W. G. (1990). What engineers know and how they know it : analytical studies from aeronautical history. Baltimore: Johns Hopkins University Press.

Wolf , T. V., Rode, J., A. , Sussman, J., & Kellogg, W., A. (2006). Dispelling "design" as the black art of CHI. Paper presented at the Proceedings of the SIGCHI conference on Human Factors in computing systems, Montreal, Quebec, Canada.

What Engineers Know

Joseph C. Pitt
Virginia Tech

To say that what engineers know constitutes engineering knowledge, just as what scientists know constitutes scientific knowledge, is a misleading way of expressing what ought to be a truism. For surely what constitutes scientific knowledge exceeds not only what one scientist knows but even the sum total of what all scientists know – since there are scientific truths that no scientists may remember at any given time. Thus, Mendel's laws were forgotten until they were "rediscovered". On the other hand, it may be the case that the total of scientific knowledge is less than the sum of what all scientists know since what scientists know is not uniformly consistent. That is, what some scientists know is sometimes at odds with what other scientists know – perhaps even contradictory – hence a reduction in total knowledge.

Interestingly, the sum total of engineering knowledge does not seem to suffer from this problem. Contradictions do not seem to appear within the confines of the epistemology of engineering. There may be disagreements among engineers as to what is the most efficient solution to a problem but – given certain assumptions about the contingencies involved – it is not the case that two engineers similarly educated and experienced could be armed with sufficiently different perspectives that they would flat out contradict each other.

In this paper I examine some aspects of engineering knowledge in order to determine what it is that engineers know. A lot will depend on how we construe "knowledge". I will argue for a pragmatic account of knowledge, in which based*on the very grounds on which the claim of superiority is made for scientific knowledge*, engineering knowledge is shown to be far more reliable than scientific knowledge – thereby exposing the lie in the traditional view that science is our best and most successful means of producing knowledge. I will begin with a quick sketch of a pragmatic theory of knowledge, followed by a look at scientific knowledge before turning to engineering knowledge. I conclude with a look at the fate of some tradition philosophical problems.

A Pragmatic Theory of Knowledge

Epistemology is an old topic and it remains stuck-in-a-rut. Since at least Plato, theories of knowledge have concentrated on one crucial factor – the inner mental state of a single individual. Prior to the work of David Hume that mental state was *certainty.* After Hume, empiricists abandoned certainty for some modified form of *justified true belief.* Nevertheless, the stress is still on what a single person knows. The view I am urging was first expressed in the work of Charles Saunders Peirce. The tradition Peirce founded extends through William James, John Dewey, C. I. Lewis, Nelson Goodman, W. V. O. Quine, Nicholas Rescher and, of course, Wilfrid Sellars, just to name a few. The simple idea they endorse in one form or another, is that to qualify as knowledge a proposition or set of propositions must be endorsed by an appropriate community. In *Thinking About Technology* (1999), I put it roughly this way: Individuals produce candidate claims for knowledge, and these candidates become knowledge once they are endorsed by the appropriate community using agreed upon standards. This gives nothing to the Strong Programme sociologists, nor to the relativists – after all, Peirce was a realist. But it does relieve us from the fruitless tedium of devising doomed criteria by which we can determine whether an individual uttering a proposition with X, Y, and Z properties can be said to know something. The criteria are doomed because they ignore contingency, historical and otherwise. A pragmatic account, on the other hand, shifts the emphasis to, for example, the criteria that the scientific community has devised. But, even here, the criteria must meet some bottom line condition. For the pragmatist the bottom line is *successful action*. According to C.I. Lewis, "the utility of knowledge lies in the control it gives us, through appropriate action, over the quality of our future experience" (Lewis 1962, p. 4).

Scientific Knowledge

The nature, structure, and justification of scientific knowledge have been topics of central importance for most of the twentieth century. While it is still not clear that there is complete consensus on the criteria for scientific knowledge, nor should there be since science is an evolving activity, several key features have emerged from the discussion. These have grown out of a reassessment of criteria initially proposed for scientific knowledge in the course of the Scientific Revolution, when the kind of knowledge the New Science was proposed to deliver was alleged to differ fundamentally in kind from what had been previously accepted as knowledge - Aristotelian in character, proceeding from esoteric definitions of fundamental concept.

From the New Science tradition, there are several treasured characteristics of scientific knowledge that recent discussions have forced us to abandon or significantly modify. Given the New Science's emphasis on the role of mathematics, scientific knowledge was described as "universal," "true," and "certain." As the special features of the different sciences – most notably the social sciences – became more pronounced, however, the universality claim had to be modified and carefully bracketed. In the social sciences the development of social relativism made this inevitable. Scientific claims to "truth" and "certainity" suffered a similar fate. But in these cases the problems were not due to specific aspects of the individual sciences. Rather, they resulted from the difficulty of demonstrating the truth of scientific claims in a non question-begging manner – on the one hand – and – on the other hand – from the recognition of the fundamentally underdetermined nature of the relation between any scientific claim and its evidence.

Given emendations formulated in light of criticism, which arose in response to these newly reconstituted problems traditional account offers some features that remain viable. For example, it characterizes scientific knowledge as produced by researchers exploring the domain of a *theory* who aim to provide an account of the relations among the objects and processes of that domain, an account which provides the basis for an explanation of phenomena generally observed or detected in another domain. If I were tempted to isolate one crucial characteristic of scientific knowledge, it would be this: Scientific claims derive their meaning from the theories within which they are associated, hence, *scientific knowledge is theory-bound.*[1]

The theory-bound nature of scientific knowledge presents additional problems beyond those noted above for some traditional assumptions about

---

1 This is not the place to explore the intriguing question of the relation between theory and the technological infrastructure of science, but it should be noted that there is a complex interaction between the theories scientists employ and that infrastructure. For example, sometimes the examination of the kinds of objects which populate a given domain may be made possible by new instruments, e.g., Galileo's telescope revealing the existence of the moons of Jupiter. Likewise, certain theories may require increasing sophistication in their supporting instrumentation. Referring again to the history of astronomy, once viewing the heavens through the telescope was possible, questions concerning the size of the universe forced the modification of the telescope by incorporating a micrometer for the purposes of making such measurements, which required the development of a theory of measurement and distance, etc. (See Pitt 1994).

scientific knowledge – in particular the view that scientific knowledge, if true, is true for all time. If scientific knowledge is theory-bound, and if – as we know from the history of science – theories change, then scientific knowledge changes. Hence, what is accepted as scientific knowledge is not true for all time, at least not all of it, not yet.[2] But this should not be a startling claim. The development of human knowledge is a process of continuous exploration in which we re-evaluate what we know in the course of new findings, and we jettison that which no longer remains consistent with the latest body of information.

We should note further that the tentative nature of scientific knowledge does not mean that knowledge is merely relative – especially in any sense that gives comfort to those opposed to the epistemic priority we traditionally give to scientific claims. The dynamic process in which scientists continuously revise what they are willing to endorse – and by which they examine their assumptions and their methods – is at the very heart of the strength of the sciences. Thus, despite the theory-bound nature of scientific knowledge, the self-critical process of scientific inquiry insures that the knowledge it claims is the best available at that time insofar as it is judged "best" according to community standards.

The ultimate aim of scientific inquiry is explanation. Thus, in the context of a pragmatic account, the ultimate success of the use of scientific knowledge is explanation. We use a theory to explore a domain of objects, sorting out their various relations for the purpose of explaining what can't be explained otherwise by appeal to the activities of the objects in that domain. Why is a tabletop hard? To answer that question we have found that we need to appeal to a scientific theory which proposes that there is a domain of smaller objects which are held together by a series of forces and that it is because of the forces and objects in that micro-domain that our phenomenological report of a hard table is possible. The aim of science is to help us understand the way the world appears to us, and it accomplishes this aim by constructing and testing theories which appeal to features of the world which are not immediately obvious.[3]

---

2 I am referring specifically here to the history of the acceptance of theories and not to the question of their truth.

3 This account of scientific explanation appears to endorse a form of scientific realism. The theory of explanation on which this view rests was developed by Wilfrid Sellars, and he was a scientific realist – he accepted the view that the ultimately real constituents of the world are the theoretical

There are other aspects of scientific knowledge which are essential to its vitality, but they need not be of concern here. In order to have a fruitful starting point to investigate the nature of engineering knowledge we need only concentrate on these two features; (1) Scientific knowledge is theory bound, and (2) scientific knowledge is developed to explain the way the world works. Unfortunately, while the process of trial and error and reappraisal characteristic of scientific activity seems to reveal its strength, this process also serves to undermine its claim of epistemic superiority over engineering knowledge. Likewise – as we shall see – the theory-bound nature of scientific knowledge creates a number of problems that do not plague engineering knowledge.

Engineering Knowledge

In *What Engineers Know and How they Know it* **(**1988), Walter Vincenti identifies and develops a theme first introduced by Edwin Layton in his landmark paper "Technology as Knowledge." Vincenti provides an account of engineering knowledge from the point of view of a practicing and deeply reflective engineer. Both Layton and Vincenti endorse the view that engineering knowledge – and technological knowledge in general – constitutes a discrete form of knowledge that is different from scientific knowledge. In a later piece, his classic 1987 Society for the History of Technology Presidential Address, "Through the Looking Glass or News from Lake Mirror Image," Layton endorses the findings of A.R. Hall, and claims that "technological knowledge is knowledge of how to do or make things, whereas the basic sciences have a more general form of knowing." (Layton 1987, p. 603) Vincenti echoes this, invoking Gilbert Ryle's famous distinction between knowing how (technology) and knowing that (science).

---

entities posited by our best confirmed theories. I accept the structure of Sellars' theory of explanation and replace Sellarsian scientific realism with Sicilian Realism. Sicilian Realism rests on two points: (1) accepting the position that the entities postulated by the current set of accepted scientific theories are *all* real, the world being a very complicated place, and (2) rejecting the principle of reduction, by which the entities of one domain are said to be nothing other than compositions of the entities of the domain of this or that scientific theory, e.g., tables are nothing more than collections of molecules. Sicilian Realism is realism with a vengeance.

Both Layton and Vincenti are concerned to defend the view that – while both science and technology may borrow from or rely on each other in various ways – they constitute two distinct forms of knowledge since they aim at different ends. Science aims to explain and technology/engineering aims to create artifices. Vincenti puts it this way, "technology, though it may *apply* science, is not the same as or entirely *applied* science" (Vincenti 1990, p. 4). He defends this claim in part with an intriguing and highly suggestive proposal. As he sees it, if we start with the proposition that technology is applied science, then there is no possibility of considering the view that technology could involve an autonomous form of knowledge that could account for those technological achievements which are science independent – such as the pyramids of Egypt and the roads of ancient Rome. Given the existence of highly visible science-independent technologies, we have good reasons to believe that we should not characterize technology as merely applied science. It is does not follow from the fact that science and technology each has occasion to rely on the other, nor that one is a subset of the other. Assuming is quasi-autonomous from, what can we say about the distinctive nature of engineering knowledge as a specific form of technological knowledge?

Starting from a wonderfully succinct definition of "engineering" by G.F.C. Rogers – which is highly reminiscent of Emmanuel Mesthene's definition of "technology" (Mesthene 1970, pg. 25). – Vincenti identifies three main components of engineering and then concentrates on the notion of design. According to Rogers (as quoted by Vincenti and augmented somewhat by me),

> Engineering refers to the practice of organizing the deign and construction (and I (Vincenti) would add operation) of any artifice which transforms the physical (and, I (Pitt) would add, social) world around us to meet some recognized need (Vincenti 1990, p.6).

One of the commendable aspects of Rogers' definition is his characterization of engineering as a practice. That is, engineering – like science – is an *activity* with specific objectives. Given Rogers' insight and Mesthene's definition of "technology" as "the organization of knowledge for the achieving of practical purposes" – by a series of substitutions we see that, appropriately enough, *engineering knowledge concerns the design, construction, and operation of artifices for the purpose of manipulating the human environment.* Vincenti proceeds to further narrow the focus of engineering knowledge to the topic of "design knowledge," by concentrating on design. It is worth quoting Vincenti's description of the design process at length because it immediately introduces an important distinction between the design as a set of plans and the design process.

> "Design", of course, denotes both the content of a set of plans (as in "the design for a new airplane") and the process by which those plans are produced. In the latter meaning, it typically involves tentative layout (or layouts) of the arrangement and dimensions of the artifice, checking of the candidate device by mathematical analysis or experimental test to see if it does the required job, *and modification when (as commonly happens at first) it does not. Such procedure usually requires several iterations before finally dimensioned plans can be released for production. Events in the doing are also more complicated than such a brief outline suggests. Numerous difficult trade-offs may be required, calling for decisions on the basis of incomplete or uncertain knowledge. If available knowledge is inadequate, special research may have to be undertaken* (Vincenti 1990, p. 7 - emphasis added).

The process Vincenti describes is "task specific" and essentially characterized by trial and error, but that still doesn't reveal the general nature of the contents of design knowledge. This is case because to capture the nature of the knowledge required for any kind of task, Vincenti must invoke a detailed model which breaks that process up into both vertical and horizontal components, thereby allowing for a precise identification of what is needed when and where in the total design process. This schema is proposed for what Vincenti, calls normal design, as opposed to radical design.[4] Normal design has five divisions beginning with the crucial aspect of any problem-solving process, the identification of the problem. Vincenti, an aeronautical engineer, draws from his own discipline for appropriate examples, but the schema is general enough to encompass a large number of design processes. For example, the design of an architectural project including sighting of the building, electrical systems, plumbing, etc., or the design of a space-based, orbiting telescope.

1. Project definition - translation of some usually ill-defined military or commercial requirement into a concrete technical problem for level 2.

---

4 Following E. Constant in his **The Origins of the Turbojet Revolution**, Johns Hopkins Press 1980)

2. Overall design - layout of arrangement and proportions of the airplane to meet project definition.

3. Major-component design - division of project into wing design, fuselage design, landing-gear design, electrical-system design, etc.

4. Subdivision of areas of component design from level 3 according to engineering discipline required (e.g., aerodynamic wing design, structural wing design, mechanical wing design).

5. Further division of categories in level 4 into highly specific problems (e.g., aerodynamic wing design into problems of platform, airfoil section, and high-life devices. (Vincenti 1990, p. 9)

The process Vincenti outlines appears simple enough. One defines the problem, breaks it into components, and subdivides the areas by problem and specialty required, as needed. What is not obvious at first glance is the way in which the levels interact. Upon further reflection, one can see that what happens at level three will have ramifications for the overall design and visa versa, but recognizing this requires some work. In short, any design project must allow for a good deal of give and take throughout the process. In this respect, if one focuses only on the give and take, the design process sounds reminiscent of the scientific process. But there is a more and it clearly marks out a crucial difference between the process of scientific inquiry and engineering design. As Vincenti says it, "Such successive division resolves the airplane problem into smaller manageable subproblems, each of which can be attacked in semi-isolation. *The complete design process then goes on iteratively, up and down and horizontally through the hierarchy*." (Vincenti 1990, p.9, emphasis added) If – by way of example – we apply this way of thinking to an architectural problem, we can easily determine what kind of a building to design (level 1), e.g., specific or multi-purpose, as opposed to the kinds of bathroom fixtures to have (level 4), although the one will ultimately bear on the other.

At this point we can pause and take stock of this comparison of scientific and engineering knowledge. First, the characterization of scientific knowledge as theory-bound and aiming at explanation appears to be in sharp contrast to the kind of knowledge Vincenti seeks. Engineering knowledge is task-specific and aims at the production of an artifact to serve a predetermined purpose.

There is a second important difference between the two forms of knowledge thai is revealed by Vincenti's account of engineering knowledge. With engineering cast as a problem-solving activity (not in itself a characteristic which distinguishes it from other activities such as biology or even philosophy), the manner in which engineers solve their problems does have a distinctive aspect. The solution to specific *kinds* of problems ends up catalogued and recorded in the form of reference works which can be employed across engineering areas. For example, measuring material stress has been systematized to a great extent. Depending on the material, how to do it can be found in an appropriate book. This gives rise to the idea that much engineering is "cookbook engineering," but what is forgotten in this caricature is that another part of the necessary knowledge is knowing what book to look for. This a unique form of knowledge that engineers bring to problem solving. But there is more: We read the phrase cookbook engineering usually in a derogatory way. But what is wrong with it? If the knowledge in the book represents information we can use in a variety of circumstances, nay, in circumstances wherever certain contingencies hold – then isn't this knowledge that comes close to being universal, certain and, must we say it? – true? Could it be said that those who refer to engineering knowledge as stored in books as cookbook knowledge are employing a bit of rhetoric, in order to hide the inadequacies of scientific knowledge?

Contrast this cookbook knowledge with theory bound knowledge. When the theory is shown in some way or other to be flawed fundamentally, it is replaced. That means that what we thought we knew to be the case, isn't – which hardly sounds like knowledge to me. However, a good cookbook providing stress calculations can be used anywhere, anytime, as long as you factor in the appropriate contingencies. Just reflect on the basis of the metaphor – a good cookbook makes it possible for anyone to prepare a good meal.

Let's go one step further and contrast Vincenti's account of the engineering design process with the activity of science. I think it has been shown in sufficient detail in a number of places, by a number of people, that there is no such thing as *the* scientific method, i.e., that there exists one method which insures objectivity and guarantees the production of universal, certain and true knowledge. One appeal to the theory-based nature of scientific work should dispel any lingering illusions. In light of the fact that a scientist working within a theory is exploring the domain circumscribed by that theory, the direction of his or her research, i.e., the kind of research he or she will undertake, will be theory–determined. On the other hand, while the domain of the theory is necessarily where the research will be directed, there is no guide supplied by the theory as to what should be investigated and how. Further, there is no one method that works for all sciences. Consider Astronomy. Given the kind of one time only observations that we find in astronomy – replication, traditionally a cornerstone of scientific method, at least in principle, is impossible. Does this make astronomy not a science, hardly. On the other hand, Vincenti's account of the engineering design process provides specific and definite structure to the process of proceeding through the design process.

We can also go beyond Vincenti and look at the work of Larry Bucciarelli (Designing Engineers, Cambridge: MIT Press), who denies that there is one single design process in engineering. Bucciarelli observes that no single unique design is dictated by the nature of the object being designed or the problem to be solved. But his objection stems not from the denial of design in engineering, but rather from a fine-grained understanding of the nature of the contingencies associated. That is, with Bucciarelli, we can find processes whereby the give and flow of ideas and the importation of the relevant contingencies follow the kind of pattern that Vincenti suggests, only in a more complicated way, when you consider the different types of communities interacting. The important point here is that in engineering design, there is at least a beginning point, for Vincenti, it is the problem, for Bucciarelli it is the object. Both see that whatever processes are at work are dynamic and interactive, but they have a task-oriented beginning point, but no such beginning point is given for scientific research.

Philosophical Problems

Two possible consequences of the cookbook nature of engineering knowledge are: (1) That such knowledge can be transported across fields and (2) it can be used anywhere – the fundamentals of dam building do not change – the contingencies of the particular circumstances may dictate one approach over another, but the basics will remain solid. In contrast, scientific knowledge is not clearly "transportable" across fields in the same way as engineering knowledge. One crucial obstacle presents itself: The problem of incommensurability.

The problem of incommensurability is a philosophical problem that came to the forefront in large part with Kuhn's characterization of the nature of scientific change. For Kuhn, fundamental change in science occurs through paradigm replacement, with his view of incommensurability applying, primarily, across paradigms. A paradigm for Kuhn is many things. However, for the process of this discussion let us consider it as a complete system of thought, including methodological rules, metaphysical assumptions, practices, and linguistic conventions. Two paradigms are incommensurable, it is alleged, because claims in different paradigms cannot be compared so as to determine which claim from which paradigm is true.

For this view to be plausible, a particular theory of meaning must be assumed and a very dubious meta-linguistic assumption must be activated. First, let us look at the theory of meaning. Basically, the theory of meaning, behind the assumption of incommensurability, presumes that expressions receive their meaning contextually, within systems, i.e., paradigms, governed by unique sets of rules. This by itself is not so troublesome. The difficult part comes through the meta-linguistic assumption that there is no point of view common to both paradigms from which it would then be possible to compare claims from different paradigms. Such a common neutral point of view is necessary, it is argued, since the meanings of expressions are governed by the rules of the paradigm. If we shift an expression from one paradigm to another, its meaning will change since it will be determined according to different rules.

Among other difficult problems to sort through here is the apparently unjustified two-fold assumption that there is *one* fundamental theory of meaning which applies to all paradigms, i.e., the meanings of expressions within any particular paradigm are determined by the ***rules*** of the paradigm, but, by contrast there is no single theory of meaning that allows for comparison of expressions across paradigms. However, if we can assert that all paradigms provide meanings for the terms which occur in that paradigm through the specification of rules, then why can we not, in the same meta-language in which we pronounce this dictum, then create another paradigm with the express purpose of allowing for the comparison of expressions? It is, for example, not at all obvious that the ways by which terms are made meaningful is through the specification of rules. That is, however, the account we are considering, and it is the source of Kuhn’s problem of incommensurability. That much has been stipulated through Kuhn's account of a paradigm. But, unless something further prohibits us from doing so, surely we can say something like this: for the purpose of comparing two expressions, each drawn from a different paradigm: If the results of applying those expressions in the *meta-language*, according to the rules of the *meta-language*, is the same, in the meta-language**,** *then for all accounts and purposes those two expressions mean the same thing*. In short, if two expressions drawn from two different scientific theories yield the same result when transported into a third theory, then they can be said to make the same claim.

The solution is based on our account of engineering knowledge. If something formulated in the context of one paradigm can be used successfully in another, then deep philosophical problems about obscure theories of meaning recede. To treat the problem of incommensurability this way is not to solve it as much as to ignore it. This too may not be a bad thing. There are many philosophical problems still around to which we no longer pay attention since they seem beside the point, for example consider the pseudo problem of how many angels can dance on the head of a pin? It is not clear that this problem was ever solved, but who cares? And so to the problem of incommensurability. If the problem as stated was never solved it appears not to matter. This lack of concern is a function of having *shifted our ground* from worrying about providing an abstract philosophical justification for something that only philosophers worry about to a pragmatic condition of success: Consider the consequences of using this claim from this theory in this context.[5] If it solves our problem, then does it matter if we fail to have a philosophical justification for using it? To adopt this attitude is to reject the primary approach to philosophical analysis of science of the major part of the twentieth century, logical positivism, and to embrace pragmatism. This is a good thing to do, especially when we are concerned with technologies that have real world effects.

Finally, I noted that engineering knowledge was transportable, not just across fields but throughout the world (and perhaps beyond). Anticipating an objection from my colleagues concerned with various manifestations of cultural imperialism – let me attempt to forestall such issues. I am not saying that we *should* transport such knowledge. The appropriateness of such activities is a matter for policy considerations. That is not what I am talking about here.

Returning now to the issue I proposed at the beginning – that engineering knowledge is a more secure form of knowledge than scientific knowledge is, on the very grounds by which it is alleged that scientific knowledge is our best form

[5] (c.f. Richard Rorty. The Linquistic Turn, Chicago, U. Of Chicago Press, 1967, p. 39.)

of knowledge. However briefly, we have noted that scientific knowledge is transitory – that it changes as theories change. We have also noted that scientific method is likewise not only transitory, but unstable, depending on the area of science being discussed, not only is there no method that will work across the sciences, within a science, the nature of the domain of objects being investigated may suggest different methods; compare biochemistry with botany. Finally if scientific knowledge is to be appraised through a pragmatic theory of knowledge, and given that the objective is explanation, then as theories change, explanations fail. The history of science then becomes the history of failed theories and unsuccessful explanations.

In contrast we have engineering knowledge, which is task oriented. If the application of engineering knowledge, consisting of information in books and task specific methods and techniques results in the production of objects and the solutions of problems which meet the criteria of those for who the jobs are done, then it is successful. Because it is task oriented, and because real world tasks have a variety of contingencies to meet – e.g., materials, time frame, budget, etc., we know when an engineering project is successful or not. Further, those cookbooks represent the accumulated knowledge of what works. It is universal, certain and, if it works, must be true in some sense of "true". So, on the criteria we advocate for science, engineering knowledge seems more secure, more trustworthy, with longevity. What engineers know, therefore, is how to get the job done – primarily because they know what the job is.

References

Bucciarelli, Larry. *Designing Engineers.* Cambrdge, MIT Press, 1996.

Layton, Edward. *Technology as Knowledge.* Technology and Culture Vol. 15 1974.

_____. "Through the Looking Glass or News from Lake Mirror Image," Society for the History of Technology, Presidential Address, 1987.

Lewis, C.I. *An Analysis of Knowledge and Valuation.* La Salle: Open Court Press, 1962.

Mesthene, Emmanuel. *Technological Change: It's Impact on Man and Society*. Cambridge: Harvard University Press, 1970.

Vincenti, Walter. *What Engineers Know and How They Know It.* Baltimore: Johns Hopkins Press, 1990.

NOTES ON THE SYNTHESIS OF FORM

CHRISTOPHER ALEXANDER

$6.95

ALEXANDER NOTES ON THE SYNTHESIS OF FORM

"First, the taking in of scattered particulars under one Idea, so that everyone understands what is being talked about . . . Second, the separation of the Idea into parts, by dividing it at the joints, as nature directs, not breaking any limb in half as a bad carver might."

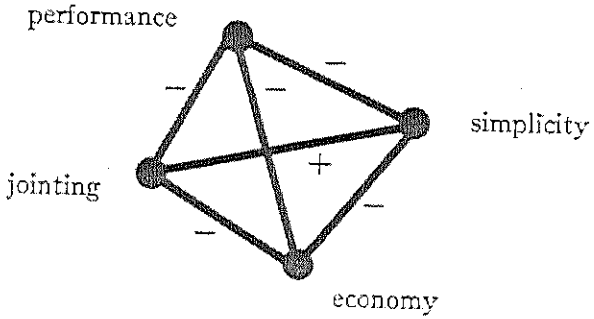
Plato, *Phaedrus*, 265D

1 / INTRODUCTION: THE NEED FOR RATIONALITY

These notes are about the process of design; the process of inventing physical things which display new physical order, organization, form, in response to function.

Today functional problems are becoming less simple all the time. But designers rarely confess their inability to solve them. Instead, when a designer does not understand a problem clearly enough to find the order it really calls for, he falls back on some arbitrarily chosen formal order. The problem, because of its complexity, remains unsolved.

Consider a simple example of a design problem, the choice of the materials to be used in the mass production of any simple household object like a vacuum cleaner. Time and motion studies show that the fewer different kinds of materials there are, the more efficient factory assembly is — and therefore demand a certain simplicity in the variety of materials used. This need for simplicity conflicts with the fact that the form will function better if we choose the best material for each separate purpose separately. But then, on the other hand, functional diversity of materials makes for expensive and complicated joints between components, which is liable to make maintenance less easy. Further still, all three issues, simplicity, performance, and jointing, are at odds with our desire to minimize the cost of the materials. For if we choose the cheapest material for each separate task, we shall not necessarily have simplicity, nor optimum performance, nor materials which can be cleanly jointed. Writing a minus sign beside a line for conflict, and a plus beside a line for positive agreement, we see that even this simple problem has the five-way conflict pictured below.



This is a typical design problem; it has requirements which have to be met; and there are interactions between the requirements, which makes the requirements hard to meet. This problem is simple to solve. It falls easily within the compass of a single man's intuition. But what about a more complicated problem?

Consider the task of designing a complete environment for a million people. The ecological balance of human and animal and plant life must be correctly adjusted both internally and to the given exterior physical conditions. People must be able to lead the individual lives they wish for. The social conditions induced must not lead to gross ill-health or to gross personal misery, and must not cause criminal delinquency. The cyclical intake of food and goods must not interfere with the regular movements of the inhabitants. The economic forces which

develop must not lead to real-estate speculation which destroys the functional relation between residential areas and areas supporting heavy goods. The transportation system must not be organized so that it creates a demand that aggravates its own congestion. People must somehow be able to live in close cooperation and yet pursue the most enormous variety of interests. The physical layout must be compatible with foreseeable future regional developments. The conflict between population growth and diminishing water resources, energy resources, parklands, must somehow be taken care of. The environment must be organized so that its own regeneration and reconstruction does not constantly disrupt its performance.

As in the simpler example, each of these issues interacts with several of the others. But in this case each issue is itself a vast problem; and the pattern of interactions is vastly complicated. The difference between these two cases is really like the difference between the problem of adding two and two, and the problem of calculating the seventh root of a fifty digit number. In the first case we can quite easily do it in our heads. In the second case, the complexity of the problem will defeat us unless we find a simple way of writing it down, which lets us break it into smaller problems.

Today more and more design problems are reaching insoluble levels of complexity. This is true not only of moon bases, factories, and radio receivers, whose complexity is internal, but even of villages and teakettles. In spite of their superficial simplicity, even these problems have a background of needs and activities which is becoming too complex to grasp intuitively.

To match the growing complexity of problems, there is a growing body of information and specialist experience. This information is hard to handle; it is widespread, diffuse, unorganized.[1] Moreover, not only is the quantity of information itself by now beyond the reach of single designers, but the various specialists who retail it are narrow and unfamiliar with the form-makers' peculiar problems, so that it is never clear quite how the designer should best consult them.[2] As a result, although ideally a form should reflect all the known facts relevant to its design, in fact the average designer scans whatever information he happens on, consults a consultant now and then when faced by extra-special difficulties, and introduces this randomly selected information into forms otherwise dreamt up in the artist's studio of his mind. The technical difficulties of grasping all the information needed for the construction of such a form are out of hand — and well beyond the fingers of a single individual.[3]

At the same time that the problems increase in quantity, complexity, and difficulty, they also change faster than before. New materials are developed all the time, social patterns alter quickly, the culture itself is changing faster than it has ever changed before. In the past — even after the intellectual upheaval of the Renaissance — the individual designer would stand to *some* extent upon the shoulders of his predecessors. And although he was expected to make more and more of his own decisions as traditions gradually dissolved, there was always still some body of tradition which made his decisions easier. Now the last shreds of tradition are being torn from him. Since cultural pressures change so fast, any slow development of form becomes impossible. Bewildered, the form-maker stands alone. He has to make clearly conceived forms without the possibility of trial and error over time. He has

to be encouraged now to think his task through from the beginning, and to "create" the form he is concerned with, for what once took many generations of gradual development is now attempted by a single individual.[4] But the burden of a thousand years falls heavily on one man's shoulders, and this burden has not yet materially been lightened. The intuitive resolution of contemporary design problems simply lies beyond a single individual's integrative grasp.

Of course there are no definite limits to this grasp (especially in view of the rare cases where an exceptional talent breaks all bounds). But if we look at the lack of organization and lack of clarity of the forms around us, it is plain that their design has often taxed their designer's cognitive capacity well beyond the limit. The idea that the capacity of man's invention is limited is not so surprising, after all. In other areas it has been shown, and we admit readily enough, that there are bounds to man's cognitive and creative capacity. There are limits to the difficulty of a laboratory problem which he can solve;[5] to the number of issues he can consider simultaneously;[6] to the complexity of a decision he can handle wisely.[7] There are no absolute limits in any of these cases (or usually even any scale on which such limits could be specified); yet in practice it is clear that there are limits of some sort. Similarly, the very frequent failure of individual designers to produce well organized forms suggests strongly that there are limits to the individual designer's capacity.

We know that there are similar limits to an individual's capacity for mental arithmetic. To solve a sticky arithmetical problem, we need a way of setting out the problem which makes it perspicuous. Ordinary arithmetic convention gives us such a way. Two minutes with a pencil on the back of an envelope lets us solve problems which we could not do in our heads if we tried for a hundred years. But at present we have no corresponding way of simplifying design problems for ourselves. These notes describe a way of representing design problems which does make them easier to solve. It is a way of reducing the gap between the designer's small capacity and the great size of his task.

Part One contains a general account of the nature of design problems. It describes the way such problems have been solved in the past: first, in cultures where new problems are so rare that there are no actual designers; and then, by contrast, in cultures where new problems occur all the time, so that they have to be solved consciously by designers. From the contrast between the two, we shall learn how to represent a design problem so that it can be solved. Part Two describes the representation itself, and the kind of analysis the representation allows. Appendix 1 shows by example how the method works in practice.

The analysis of design problems is by no means obviously possible. There is a good deal of superstition among designers as to the deathly effect of analysis on their intuitions — with the unfortunate result that very few designers have tried to understand the process of design analytically. So that we get off to a fair start, let us try first to lay the ghosts which beset designers and make them believe that analysis is somehow at odds with the real problem of design.

It is not hard to see why the introduction of mathematics into design is likely to make designers nervous. Mathematics, in the popular view, deals with magnitude. Designers recognize, correctly, that calculations of magnitude only have

strictly limited usefulness in the invention of form, and are therefore naturally rather skeptical about the possibility of basing design on mathematical methods.[8] What they do not realize, however, is that modern mathematics deals at least as much with questions of order and relation as with questions of magnitude. And though even this kind of mathematics may be a poor tool if used to prescribe the physical nature of forms, it can become a very powerful tool indeed if it is used to explore the conceptual order and pattern which a problem presents to its designer.

Logic, like mathematics, is regarded by many designers with suspicion. Much of it is based on various superstitions about the kind of force logic has in telling us what to do. First of all, the word "logic" has some currency among designers as a reference to a particularly unpleasing and functionally unprofitable kind of formalism.[9] The so-called logic of Jacques François Blondel or Vignola, for instance, referred to rules according to which the elements of architectural style could be combined.[10] As rules they may be logical. But this gives them no special force unless there is also a legitimate relation between the system of logic and the needs and forces we accept in the real world. Again, the cold visual "logic" of the steel-skeleton office building seems horribly constrained, and if we take it seriously as an intimation of what logic is likely to do, it is certain to frighten us away from analytical methods.[11] But no one shape can any more be a consequence of the use of logic than any other, and it is nonsense to blame rigid physical form on the rigidity of logic. It is not possible to set up premises, trace through a series of deductions, and arrive at a form which is logically determined by the premises, unless the premises already have the seeds of a particular us such a way. Two minutes with a pencil on the back of an envelope lets us solve problems which we could not do in our heads if we tried for a hundred years. But at present we have no corresponding way of simplifying design problems for ourselves. These notes describe a way of representing design problems which does make them easier to solve. It is a way of reducing the gap between the designer's small capacity and the great size of his task.

Part One contains a general account of the nature of design problems. It describes the way such problems have been solved in the past: first, in cultures where new problems are so rare that there are no actual designers; and then, by contrast, in cultures where new problems occur all the time, so that they have to be solved consciously by designers. From the contrast between the two, we shall learn how to represent a design problem so that it can be solved. Part Two describes the representation itself, and the kind of analysis the representation allows. Appendix 1 shows by example how the method works in practice.

The analysis of design problems is by no means obviously possible. There is a good deal of superstition among designers as to the deathly effect of analysis on their intuitions — with the unfortunate result that very few designers have tried to understand the process of design analytically. So that we get off to a fair start, let us try first to lay the ghosts which beset designers and make them believe that analysis is somehow at odds with the real problem of design.

It is not hard to see why the introduction of mathematics into design is likely to make designers nervous. Mathematics, in the popular view, deals with magnitude. Designers recognize, correctly, that calculations of magnitude only have

strictly limited usefulness in the invention of form, and are therefore naturally rather skeptical about the possibility of basing design on mathematical methods.[8] What they do not realize, however, is that modern mathematics deals at least as much with questions of order and relation as with questions of magnitude. And though even this kind of mathematics may be a poor tool if used to prescribe the physical nature of forms, it can become a very powerful tool indeed if it is used to explore the conceptual order and pattern which a problem presents to its designer.

Logic, like mathematics, is regarded by many designers with suspicion. Much of it is based on various superstitions about the kind of force logic has in telling us what to do. First of all, the word "logic" has some currency among designers as a reference to a particularly unpleasing and functionally unprofitable kind of formalism.[9] The so-called logic of Jacques François Blondel or Vignola, for instance, referred to rules according to which the elements of architectural style could be combined.[10] As rules they may be logical. But this gives them no special force unless there is also a legitimate relation between the system of logic and the needs and forces we accept in the real world. Again, the cold visual "logic" of the steel-skeleton office building seems horribly constrained, and if we take it seriously as an intimation of what logic is likely to do, it is certain to frighten us away from analytical methods.[11] But no one shape can any more be a consequence of the use of logic than any other, and it is nonsense to blame rigid physical form on the rigidity of logic. It is not possible to set up premises, trace through a series of deductions, and arrive at a form which is logically determined by the premises, unless the premises already have the seeds of a particular plastic emphasis built into them. There is no legitimate sense in which deductive logic can prescribe physical form for us.

But, in speaking of logic, we do not need to be concerned with processes of inference at all. While it is true that a great deal of what is generally understood to be logic is concerned with deduction, logic, in the widest sense, refers to something far more general. It is concerned with the form of abstract structures, and is involved the moment we make pictures of reality and then seek to manipulate these pictures so that we may look further into the reality itself. It is the business of logic to invent purely artificial structures of elements and relations. Sometimes one of these structures is close enough to a real situation to be allowed to represent it. And then, because the logic is so tightly drawn, we gain insight into the reality which was previously withheld from us.[12]

The use of logical structures to represent design problems has an important consequence. It brings with it the loss of innocence. A logical picture is easier to criticize than a vague picture since the assumptions it is based on are brought out into the open. Its increased precision gives us the chance to sharpen our conception of what the design process involves. But once what we do intuitively can be described and compared with nonintuitive ways of doing the same things, we cannot go on accepting the intuitive method innocently. Whether we decide to stand for or against pure intuition as a method, we must do so for reasons which can be discussed.

I wish to state my belief in this loss of innocence very clearly, because there are many designers who are apparently not willing to accept the loss. They insist that design must be

a purely intuitive process: that it is hopeless to try and understand it sensibly because its problems are too deep.

There has already been one loss of innocence in the recent history of design; the discovery of machine tools to replace hand craftsmen. A century ago William Morris, the first man to see that the machines were being misused, also retreated from the loss of innocence. Instead of accepting the machine and trying to understand its implications for design, he went back to making exquisite handmade goods.[13] It was not until Gropius started his Bauhaus that designers came to terms with the machine and the loss of innocence which it entailed.[14]

Now we are at a second watershed. This time the loss of innocence is intellectual rather than mechanical. But again there are people who are trying to pretend that it has not taken place. Enormous resistance to the idea of systematic processes of design is coming from people who recognize correctly the importance of intuition, but then make a fetish of it which excludes the possibility of asking reasonable questions.

It is perhaps worth remembering that the loss of intellectual innocence was put off once before. In the eighteenth century already, certain men, Carlo Lodoli and Francesco Algarotti in Italy and the Abbé Laugier in France, no longer content to accept the formalism of the academies, began to have serious doubts about what they were doing, and raised questions of just the sort that have led, a hundred and fifty years later, to the modern revolutionary ideas on form.[15] Oddly enough, however, though these serious doubts were clearly expressed and widely read, architecture did not develop from them in the direction indicated. The doubts and questions were forgotten. Instead, in late eighteenth century Europe, we find evidence of quite another atmosphere developing, in which architects based their formal invention on the rules provided by a variety of manners and "styles" like neo-Tudor, neoclassicism, chinoiserie, and neo-Gothic.[16]

It is possible to see in this course of events a desperate attempt to ward off the insecurity of selfconsciousness, and to maintain the security of innocence.

Lodoli and Laugier wanted to know what they were doing as makers of form. But the search for this knowledge only made the difficulty of their questions clear. Rather than face the responsibility of these difficult questions, designers turned instead to the authority of resurrected "styles." The architectural decisions made within a style are safe from the nagging difficulty of doubt, for the same reason that decisions are easier to make under tradition and taboo than on one's own responsibility. It is no coincidence, in my opinion, that while the Renaissance had allowed free recombinations of classical elements, the neoclassicism which replaced it stuck as closely as it could to the precise detail of Greece and Rome. By leaning on correctness, it was possible to alleviate the burden of decision. To make the secession from responsibility effective, the copy had to be exact.[17]

Now it looks as though a second secession from responsibility is taking place. It is not possible today to escape the responsibility of considered action by working within academic styles. But the designer who is unequal to his task, and unwilling to face the difficulty, preserves his innocence in other ways. The modern designer relies more and more on his position as an "artist," on catchwords, personal idiom, and intuition — for all these relieve him of some of the burden of decision, and make his cognitive problems manageable. Driven on his own resources, unable to cope with the complicated information he is supposed to organize, he hides his incompetence in a frenzy of artistic individuality. As his capacity to invent clearly conceived, well-fitting forms is exhausted further, the emphasis on intuition and individuality only grows wilder.[18]

In this atmosphere the designer's greatest gift, his intuitive ability to organize physical form, is being reduced to nothing by the size of the tasks in front of him, and mocked by the efforts of the "artists." What is worse, in an era that badly needs designers with a synthetic grasp of the organization of the physical world, the real work has to be done by less gifted engineers, because the designers hide their gift in irresponsible pretension to genius.

We must face the fact that we are on the brink of times when man may be able to magnify his intellectual and inventive capability, just as in the nineteenth century he used machines to magnify his physical capacity.[19] Again, as then, our innocence is lost. And again, of course, the innocence, once lost, cannot be regained. The loss demands attention, not denial.

rant

CHI and the Practitioner Dilemma

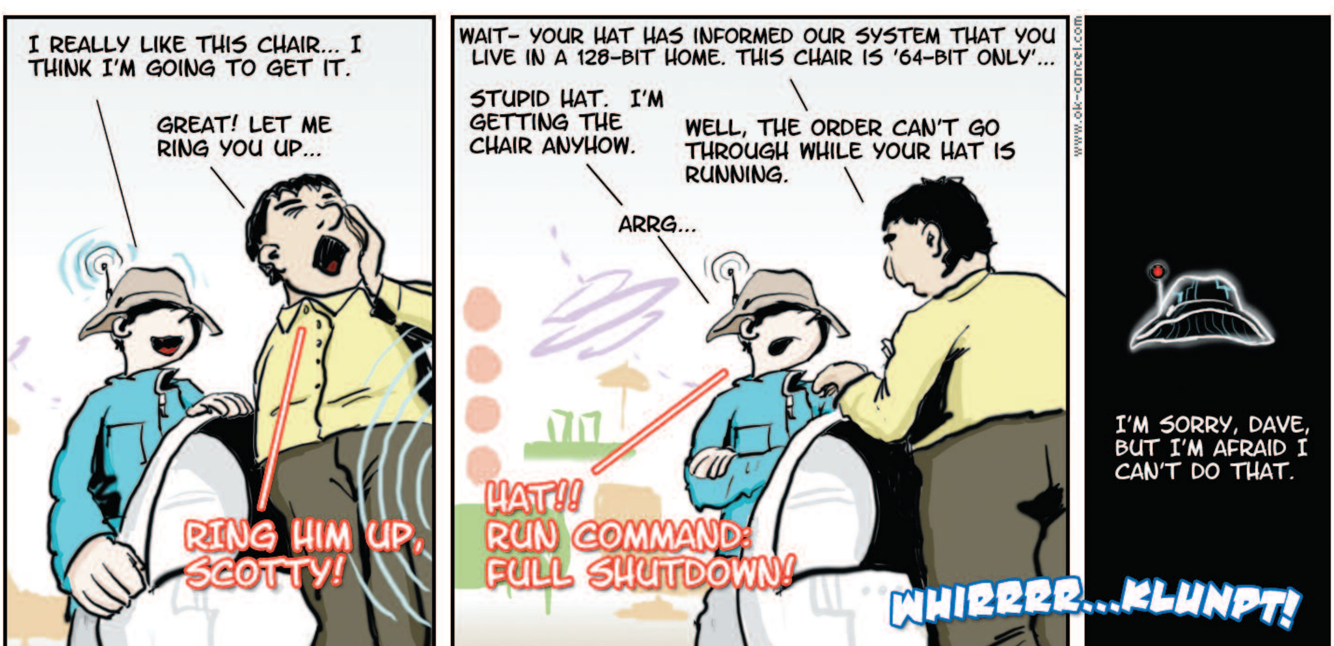
This year's CHI conference was much like any other: the parade of papers, panels, exhibits; entertaining plenary speakers; an overcrowded invigorating reception; and a micro-Design Expo. There were a few new things (like the march to get to the reception) and many of the same traditional things (a second student design competition, a solid program with traction). But there was something else that most of you probably weren't aware of: an outright blow-up between practitioners and academics at no less a venue than the usually staid SIGCHI Membership Meeting.

The meeting started in its traditional stolid and understated manner. Having recited overall accomplishments for the last year at a gallant clip, vice president for membership Julie Jacko and president Joe Konstan arrived at the discussion of CHI 2006. Gary Olson, CHI 2006 chair, was given the floor to discuss a new organization of the conference for next year's event in Montreal. His remarks were met at first with what one might describe as brooding silence. However, a spark was ignited when a couple of CHI big spenders—those that spend tens of thousands of U.S. dollars sponsoring, exhibiting, and sending scores of attendees—stood up and challenged the SIGCHI executive committee and the conference management committee with the following: "Either you make more practitioner-relevant materials available at CHI next year or we will not be coming back."

The well-intentioned challenge became almost surreal when some tenured academics reacted furiously charging that these practitioners were trying to ruin the papers

OK/Cancel

by Tom Chi/Kevin Cheng www.ok-cancel.com



OK/Cancel

<interactions>

is a bimonthly publication of ACM, The Association for Computing Machinery

Director of Publications: Mark Mandelbaum

Publications board: Chair: Robert B. Allen; Gul Agha; Michel Beaudouin-Lafon; Ronald F. Boisvert; Adolfo Guzman-Arenas; Wendy Hall; Carol Hutchins; Mary Jane Irwin; Keith Marzullo; M. Tamer Ozsu; Holly Rushmeier; Mary Lou Soffa.

Subscriptions: Annual cost to ACM members: $49.00, student members $42.00, individual nonmembers: $75.00, nonmember institutions $145.00.

Single copies are $13.00 to members, $20.00 to nonmembers. Please send orders prepaid plus $4.00 for shipping and handling to ACM Order Dept., P.O. Box 11405 Church Street Station, New York, NY 10286-1405 or call +1-212-626-0500. For credit card orders, call +1-800-342-6626. Order personnel available 8:30-4:30 EST. After hours, please leave message and order personnel will return your call.

Change of address: acmcoa@acm.org
For other services, questions, or information: acmhelp@acm.org

fresh

track and destroy the academics' ability to get tenure. It took all of Joe Konstan and Julie Jacko's diplomatic skills to settle things down.

It turns out the academics were truly upset at a proposal by CHI 2006 to radically alter the papers program. Then along came the practitioner's challenge and it was like lighting a match in a dusty coal mine.

Both groups have a serious and legitimate complaint: CHI is not meeting our needs. This is nothing new. But when action suggests that community voices are ignored, those voices grow louder. It's been a long time since we've seen an actual shouting match at CHI. Unfortunately each side of the discussion slightly misunderstood the other's arguments and the context for their frustrations.

The CHI conference needs to address the concerns of these communities. It can't go into black-box mode for solutions. And black box doesn't just mean everything's secret; it also means that news, policies, and the politics of volunteer organizations simply don't reach you unless it is pushed at you. There is the *matter* of communicating, and the *manner*. It's our fault for not getting involved if we don't ask questions and follow through the implications of their answers. We want a process and ultimate content that are

Raise your voice, provide feedback, submit, review, volunteer— for CHI and elsewhere. You have to create the mandate for change.

both more strategic and inclusive. The CHI conference committee(s), the CMC, the publications board, and the executive committee are all well intentioned, but the process by which they work is broken.

Yes, changes need to be made to the CHI conference. By the last week of April, two weeks after CHI 2005 ended, two more championship sponsors added their voices to the earlier mentioned ultimatum.

Here's what we see being done about it: appointing practitioners to the CHI 2006 committee and giving them a sandbox for new submissions. This is a good thing. First year "experiments" often are. But don't count on it developing a "following"—our experience has taught us something about the CHI conference: If you liked what they did in '07, too bad, because in '08 it will disappear. So how do you institute a new thing with any security that it will reappear? Well, one way is to have that new thing keep popping up here and there to build some brand recognition: Enter the case study.

We worked five years on trying to get design case studies canonized into the CHI conference. People like Austin Henderson, Hugh Dubberly, Terry Swack, Ian McClelland, and others from SIGCHI, SIGGRAPH, AIGA, and the Stanford Business School have worked on this case study submission format. Here's its track record to date:

- *CHI 2001*—one Design Expo session, nine presentations, attended by 1300 people in a standing-room-only hall
- *CHI 2002*—Design Expo dropped from the conference program; instead created the CHI2002|AIGA Experience Design FORUM, a two-day co-hosted preconference satellite, well-received and financially successful
- *CHI 2003*—Design Expo dropped again; offered one panel session for four case studies
- *CHI 2004*—three Design Expo sessions; very successful with practitioner-directed marketing and outreach
- *CHI 2005*—two Design Expo sessions, no marketing budget and no outreach; many practitioners did not even realize it was there
- *CHI 2006*—Design Expo dropped from the confer-

ence as of press time, but there will be three new submission formats

This lack of continuity and consistency isn't the volunteers' fault. (Indeed former CMC member Scooter Morris writes a very provocative and encouraging article below.) The real problem, as we see it, is the way these organizations all work together. Or, clearly, don't.

You can guess our *solution* to the problem—we think all the above-named organizations would even agree. Many on the EC and CMC have articulated this already. Practitioners aren't second-class citizens at CHI, they're just a *different* kind of first-class citizen with a literature and a voice to develop:

- Canonize the case-study format and practitioner forums into the core conference.
- Ensure the continuation and consistent naming of CHI conference venues of primary interest to practitioners.
- Give conferences like CHI 2006 the chance to experiment with new ideas, like Experience Reports, CHI Notes, etc., then have a way to review and either iterate at the next conference or drop it and document the reason why with the rationale, to avoid reinventing the square wheel; and please, make a point of communicating the results.
- Assure any proposed change to the community-specific tracks (papers, case studies, student competition, etc.) are done with care and by consulting openly with these communities; then it's their fault if they neglect to speak up.
- Consult with academics and practitioners to make it easier for some research papers to be more practitioner-focused or at least practitioner-friendly; we

Food for Thought: Why Doesn't SIGCHI Eat Its Own Dog Food?

By John "Scooter" Morris > University of California, San Francisco > scooter@cgl.ucsf.edu

Investopedia says a company that "eats its own dog food" sends a message that it considers its own products the best on the market. This slang was popularized during the dot-com craze when companies did not use their own software and thus could not even "eat their own dog food." An example of not "eating your own dog food" would be a software company that creates operating systems but uses its competitor's software on its corporate computers.

As a discipline, we believe in putting the user at the center of our designs—involving them in the process whenever possible, and evaluating the results of our designs and design modifications with real users, preferably during the development stages of the product. Why, then, don't we follow this same approach with respect to the development of the "products" of our organization? Does user-centered design only apply to computer-based products, or should we, as an organization, demonstrate the strength of user-centered design by practicing it for the development of our Web sites, publications, and conferences?

While I believe that we can and should practice user-centered design on all of our products, this rant is going to focus on the annual CHI conference. The CHI 2005 "product" shipped, and we have announced that the overall design for the next "product" in the series, CHI 2006, will incorporate several significant changes. The purpose of this rant is not to discuss or criticize those changes or any of the volunteers who give of their time and energy to benefit the community and the field. This is a rant about our process of design, and why I believe a more user-centered approach is in order.

User-centered design (UCD) is a process by which the user of the product is the focus of the design rather than the product itself being the focus. We believe that involving the user in the design process results in better products that are more usable than products designed without user involvement. We espouse the testing of product designs, features, and interaction techniques with users early in the development process. We encourage the study of users in context before we have firm product directions.

- would greatly appreciate any such effort. (Whatever happened to applied research?)
- Redesign the SIGCHI Web sites to appeal to practitioners. (Well, OK, we just threw that one in but really, don't you think it's about time?)

It has been five CHI conferences since the challenge to attract practitioners was first posed by frustrated designers at CHI 1999. They told us to do something about it. In the resulting everlasting flow of tweaks and changes, the CHI conference at the end of the day looks roughly the same, to the attendee, from year to year.

So we wheel our scope around and aim it at you, dear reader: You have to create the mandate for change. Raise your voice, demand to review formats, provide feedback, submit, review, volunteer—for CHI and elsewhere. We're expecting the solution to come from you. Consider this our call to action: Write to us! Tell us what YOU want! We promise we'll use this magazine to make all opinions count. Write a letter to the editor, write a short article, step forward and volunteer to be a special section editor and collect thoughts from all sides of this debate.

After all, form follows meaning—a perfectly appropriate guiding epithet for CHI 2006 and beyond. —*<eic>*

We caution developers and designers not to make assumptions about users without testing those assumptions.

Let us turn now to the use of UCD for the development of a new release of a successful product. Ideally, we would have two separate pieces of information as input to our design: ongoing studies of our target user community and evaluations of previous versions of our product based on surveys and user testing. Management has set specific goals and constraints for the design, including development time, financial goals, marketing goals, etc. Whenever possible, we would derive metrics to evaluate our product based on our goals and constraints. The design process might include a period of brainstorming to search for innovative new aspects to the product based on our assessment of the user's needs and management goals. Innovations that introduce significant changes to the product would probably be labeled as "risky" and would need to be prototyped and explicitly tested against real users. In the end, the entire product would be tested and evaluated against the metrics and goals established by management.

One of the major products of SIGCHI is the CHI conference. The above outline bears little resemblance to the actual design process for each year's "product." We generally assume we are the user, seldom evaluate, and rarely do formal testing of our prototypes. We have articulated goals for the conference series, but we don't tie those goals to our evaluation of the conference, or our design changes to the conference in any formal manner. We point to conference evaluations and attendance patterns as rationale for changes without testing our assumptions or our proposed designs against our target audience in any formal manner. One of the constraints we have is that this process is managed by volunteers who have real jobs and lives outside of SIGCHI and the CHI conference. All design processes have constraints, but we cannot afford to allow those constraints to force the user out of the design.

User-centered design, or more broad-

Eat Your Own Dog Food, an expression describing the act of a company using its own products for day-to-day operations.

-Answers.com

Research Through Design as a Method for Interaction Design Research in HCI

John Zimmerman, Jodi Forlizzi, Shelley Evenson
Human-Computer Interaction Institute and The School of Design
Carnegie Mellon University
{johnz, forlizzi, evenson}@andrew.cmu.edu

ABSTRACT
For years the HCI community has struggled to integrate design in research and practice. While design has gained a strong foothold in practice, it has had much less impact on the HCI research community. In this paper we propose a new model for interaction design research within HCI. Following a research through design approach, designers produce novel integrations of HCI research in an attempt to make the *right* thing: a product that transforms the world from its current state to a preferred state. This model allows interaction designers to make research contributions based on their strength in addressing under-constrained problems. To formalize this model, we provide a set of four lenses for evaluating the research contribution and a set of three examples to illustrate the benefits of this type of research.

Author Keywords
design, interaction design, interaction design research, HCI research, research through design, wicked problems, design theory, design method

ACM Classification Keywords
H5.2. User Interfaces: Theory and methods.

INTRODUCTION
In recent years we have both witnessed and participated in the struggle as several academic institutions have attempted to integrate design, with technology and behavioral science in support of HCI education and research. While there has been great excitement about the benefits integrating design can bring, we quickly realized that no agreed upon research model existed for interaction designers to make research contributions other than the development and evaluation of new design methods. Over the last two years we have undertaken a research project to (i) understand the nature of the relationship between interaction design and the HCI research community, and (ii) to discover and invent methods for interaction design researchers to more effectively participate in HCI research.

Through our inquiry we learned that many HCI researchers commonly view design as providing surface structure or decoration. In addition, we lack a unified vision of what design researchers can contribute to HCI research. This lack of a vision for interaction design research represents a lost opportunity for the HCI research community to benefit from the added perspective of design thinking in a collaborative research environment. The research community has much to gain from an added design perspective that takes a holistic approach to addressing under-constrained problems.

To address this situation, this paper makes two contributions: (i) a model of interaction design research designed to benefit the HCI research and practice communities, and (ii) a set of criteria for evaluating the quality of an interaction design research contribution. The model is based on Frayling's *research through design* [14], and it stresses how interaction designers can engage "wicked problems" [21]. What is unique to this approach to interaction design research is that it stresses design artifacts as outcomes that can transform the world from its current state to a preferred state. The artifacts produced in this type of research become design exemplars, providing an appropriate conduit for research findings to easily transfer to the HCI research and practice communities. While we in no way intend for this to be the only type of research contribution interaction designers can make, we view it as an important contribution in that it allows designers to employ their strongest skills in making a research contribution and in that it fits well within the current collaborative and interdisciplinary structure of HCI research.

Definitions
As we conducted this inquiry, we quickly realized that within both the HCI and design communities there is an inconsistent and confusing use of the following terms. Therefore, below we provide a set of definitions for these terms with respect to this paper.

Designer. Using such a generic term is a challenge at best. At CHI 2006's SIG: "The CHI Design Community", Bill Buxton sarcastically claimed that if everyone is a designer because they select their own clothes, then everyone is also a mathematician, because we all count our change. His comment captures what a loaded term "designer" is. Within the HCI community, it is quite common for people to use the term design to mean HCI practice and to use the term designer to mean an HCI practitioner. In this case a designer might be an interaction designer, a usability engineer, a software architect, a software developer, etc. However, in the design community, the term designer is generally used to refer to someone who has had training or extensive practical experience in a discipline such as architecture, product design, graphic design, or interaction design. As we use the term designer in this paper, we are following the convention of the design community.

Design research. In the HCI community and in the design practice community, the term design research is generally used to refer to the upfront research practitioners do to ground, inform, and inspire their product development process. However, in the design research community, including institutions such as the Design Research Society, the term design research implies an inquiry focused on producing a contribution of knowledge. This paper follows the convention of the design researchers, and we intend the term design research to mean an intention to produce knowledge and not the work to more immediately inform the development of a commercial product.

Design thinking. This term is often used to describe what designers bring to problem solving and to rationalize why designers need to be included in a project or process; however, it is rarely defined. In some respects its ambiguity is part of its strength, allowing it to be the right thing at the right time. In terms of this paper, we mean the application of a design process that involves grounding—investigation to gain multiple perspectives on a problem; ideation—generation of many possible different solutions; iteration—cyclical process of refining concept with increasing fidelity; and reflection.

In the next section, we provide an overview of our research and methodology in constructing this model. We highlight the findings from our literature review, and detail the evolving history of design in HCI and of interaction design research and its impact on the HCI community in order to situate our contribution within the frameworks of HCI research and design research. We then describe the model, detail how it produces knowledge, and discuss how it produces benefits for both the HCI practice and research communities. We formalize the model by describing four lenses to evaluate the quality of an interaction design research contribution. Finally, we illustrate how three examples of interaction design research can be evaluated by the criteria described here.

METHODOLOGY

Our methodology included a literature review focusing on design in HCI and on models of design research; a workshop on the relationship between design and HCI; semi-structured interviews with leading HCI researchers and leading interaction designers in academia and industry; synthesis of the findings from the literature and interviews, and the construction of a new model of design research; iterative evaluations of this model with leading HCI researchers and designers; and finally, a refinement that produced the current model.

Literature review

We reviewed the design research literature to understand historical and currently proposed models of design research and more specifically, interaction design research. In addition, we reviewed literature from the HCI community discussing the role of design.

CHI 2004 workshop

In 2004 we conducted a workshop at the CHI conference in Vienna, focusing on clarifying the relationship between HCI and design. The workshop, had 22 participants from both academia and industry and from a range of backgrounds including computer science, behavioral science, and interaction design and explored two distinct but complementary tracks: (i) role of design in HCI education and role of HCI in design education, and (ii) the role of interaction design research in HCI. Outcomes from this workshop helped frame our focus on the need to define models of design research in HCI and motivated us to engage the broader HCI practice and research community in a discussion of what these might be.

Interviews

We conducted semi-structured interviews with nine leading academic HCI researchers accompanied by one of their graduate students. During the interview, we asked about their ideas of what design is and what design research is in terms of HCI research. We chose to interview the leaders with a graduate student for two reasons. First, we thought a process of co-discovery would help us elicit better information during the interview. Second, we wanted to see if the students, who were much newer to HCI and were being educated in a multidisciplinary environment that includes behavioral science, computer science, and interaction design, had a substantially different view of design than their advisors, who had all been trained in a single discipline.

We also interviewed six leading interaction designers. Three held senior academic positions, and three held industry positions including head of design at a consumer electronics company, a design researcher at a well-known technology research company, and the principal of a design consultancy. In these interviews, in addition to collecting information on the evolution of their career in HCI, we probed on the nature of the relationship between design and HCI and on what they saw as the important models of design research with respect to the HCI research community.

Synthesis, analysis, and iterative modeling

After generating a preliminary model, we iteratively evaluated the model through presentations and discussions. One included a large group of HCI researchers, none of whom had training in design. In addition, we held four one-on-one presentations of the model along with other research models including Dick Buchanan's model of design research [4] and Daniel Fallman's model of research-oriented design and design-oriented research [12,13]. These one-on-one interviews included a senior HCI practitioner, a leading design researcher in HCI, and two leading HCI researchers from industry. The one-on-one discussions allowed for more free-form feedback on our model and a chance for the interviewee to participate in rapid redesign. The large discussion was particularly beneficial in that it engaged the entire group in a discussion of what design research meant to their specific discipline within HCI and a discussion of what design does and should mean within the HCI research community.

LITERATURE REVIEW

Our literature review, meant to ground our inquiry, focused on the history of the emerging field of design research, the role of design in HCI, and the role of the designed artifact.

Design Research

The emergence of design research as a separate activity from design practice grew out of the need to formally address the increasing complexity of systems designers were being asked to create [3]. The increasing complexity of products such as battleships, airplanes, and rockets created a need for new design methods that were more predictable and more collaborative. The design methods movement grew out of this need, and generated the first cohort of design researchers focusing on the development of knowledge instead of artifacts for consumption.

Within the design research community, there has been an ongoing tension around the relationship between design and science [8]. Motivation for a scientific framing came from sources such as Buckminster Fuller's call "…for a 'design science revolution' based on science, technology, and rationalism…" [8 p.50], and from Herbert Simon's call for the study of science of design to help more liberally educate scientists and engineers in his book *Sciences of the Artificial* [23]. In this case the science can be a scientific study of how designers work or the use of scientific knowledge and methods in a rational practice of design [8].

In adding to the research discussion of design methods, Donald Schön introduced the idea of design as a reflective practice where designers reflect back on the actions taken in order to improve design methodology [22]. While this may seem counter to the science of design, where the practice of design is the focus of a scientific inquiry, several design researchers have argued that reflective practice and a science of design can co-exist in harmony [8, 5].

In reaction to the casting of design as a science and also in response to systems engineers' inability to apply scientific methods to address social problems such as urban crime, Horst Rittel and Melvin Webber proposed the concept of a "Wicked Problem," a problem that because of the conflicting perspectives of the stakeholders cannot be accurately modeled and cannot be addressed using the reductionist approaches of science and engineering [21]. They argued that many problems can never be accurately modeled, thus an engineering approach to addressing them would fail. This work pointed to an opportunity for design research to provide complementary knowledge to the contributions made by scientists and engineers through methods unique to design and design processes.

Design researchers describe their work as "…the study, research, and investigation of the artificial made by human beings, and the way these activities have been directed either in academic studies or manufacturing organizations." [3 p 16.] The focus of this work has been on a study of design in order to improve the process and on the analysis of design artifacts in order to generate theories that unite related methods of addressing design challenges. In general, this design research scholarship has not focused on the outcome (artifacts) of making as a design contribution.

The Role of Design in HCI

In the early days, the term "design" within the HCI community meant usability engineering: "…the process of modeling users and systems and specifying system behavior such that it fitted the users' tasks, was efficient, easy to use and easy to learn." [26 p.1]. Over time, trained designers began working with software developers, bringing skills in visual hierarchy, navigation, color, and typography they had developed designing printed artifacts. Jonas Löwgren labeled the process they brought to interaction design as "creative design" to distinguish it from the engineering approach [16]. In engineering design, developers created software to meet a specification, and in creative design, designers continually reframed the problem, constantly questioning the underlying assumptions during the design process.

Daniel Fallman's work casts HCI as a design discipline [12]. He describes the research performed by engineers and behavioral scientists as "design-oriented research." Researchers engage in designing and making prototypes in order demonstrate a research contribution. In this case, the research community benefits from the processes of design and design thinking because they lead to better research prototypes.

Christopher Alexander's work on Pattern Languages represents an example of how research performed by design researchers on design methods has had an impact on the HCI community. His work asks design researchers to examine the context, system of forces, and solutions used to address repeated design problems in order to extract a set underlying "design patterns", thereby producing a "pattern language" [1]. The HCI community has embraced this approach to address design of web sites [24]. The method turns the work of many designers addressing the same interaction problems into a discourse for the community, allowing interaction designers to more clearly observe the formation of conventions as the technology matures and is reinterpreted by users.

The Artifact as a Part of Interaction Design Research

Daniel Fallman describes the HCI development process used today as *research-oriented design* to describe the research performed to influence the design of commercial products [12]. Brenda Laurel's book, *Design Research: Methods and Perspectives*, also describes how interaction designers can perform research as they practice design to better ground their process and to hopefully increase the chances for success of a product in the marketplace [15]. Through this process, HCI practitioners and interaction designers work together as team members, keeping the needs of the user in focus for the entire development team. While both represent a combination of research and making, the focus is still on design as a practice and not as a research discipline that makes contributions of knowledge.

At last year's CHI conference a paper argued against a commonly held belief in the HCI research community that design is a "Black Art" [25]. The authors argued instead that interaction design performed in a research context employs a set of rational judgments. The case documented in this paper places interaction design in the context of HCI research and interaction designers as collaborators with researchers. However, in this specific case, the designers work in service of research, with the goal of creating a research prototype that more clearly communicates the research contribution. We certainly see this type of collaboration as important to the ongoing relationship between researchers and designers, but push for additional collaborations where designers also participate in research and engage research questions specific to interaction design.

Critical design presents a model of interaction/product design making as a model of research [9]. Unlike design practice, where the making focuses on making a commercially successful product, design researchers engaged in critical design create artifacts intended to be carefully crafted questions. These artifacts stimulate discourse around a topic by challenging the status quo and by placing the design researcher in the role of a critic. The Drift Table offers a well known example of critical design in HCI, where the design of an interactive table that has no intended task for users to perform raises the issue of the community's possibly too narrow focus on successful completion of tasks as a core metric of evaluation and product success [10].

Finally, in their book, *The Design Way*, Harold Nelson and Erik Stolterman frame interaction design—and more generally the practice of design—as a broad culture of inquiry and action. They claim that rather than focusing on problem solving to avoid undesirable states, designers work to frame problems in terms of intentional actions that lead to a desirable and appropriate state of reality. Design is viewed as a unique way to look at the human condition, and is understood through reflective practice, intellectual apperception, and intentional choice. The practice of design is framed as encompassing the real, the true, and the ideal; design research is framed as research on a condition that arises from a number of phenomena in combination, rather than the study of a single phenomenon in isolation. Our model of interaction design research in HCI attempts to formalize many of their ideas in a single method tailored to fit within the context of the HCI research community. We do not view our model as the only way for interaction designers to perform research, but as one of many.

Our model of design research advances the work of the design research community by expanding their focus on methods and analysis of artifacts to include making as a method of inquiry in order to address wicked problems. Our model builds on the current relationship design has with the HCI community by building on Alexander's pattern language model as a method of making research findings actionable by the HCI practice community. Finally our model adds a counterpoint to critical design's focus on design research in the role of critic, by creating a role for the design researcher to be an equal collaborator with HCI engineering and behavioral science researchers.

INTERVIEWS

Interviews were conducted within and outside of the discipline of interaction design to understand how collaborations between design and HCI evolve, and to iteratively test our model in progress.

Interaction designers on design and design research

In our interviews with designers we probed on the value design brings to HCI, and three main themes emerged. First, participants noted that interaction designers brought a process for engaging massively under-constrained problems that were difficult for traditional engineering approaches to address. Second, designers brought a process of integrating ideas from art, design, science, and engineering, in an attempt to make aesthetically functional interfaces. One described this process as similar to composing music or conducting a symphony, where the job is to bring out the richness in a range of voices to make a singular thing. Third, designers brought empathy for users as a part of the process. In addition to considering their needs and desires from an external-observer's perspective, designers worked to also embody the people they made things for.

The designers we spoke with described their early days of collaborating with HCI and software developers, where they were often brought in at the end of the process and asked to make the interface "pretty". In attempting to improve the designs, they were often frustrated that the suggestions they made, which often seemed obvious design improvements, could not be made because they came too late in the development process. However, over time, designers moved from a consultant role at the end of a project to team members working throughout the software design and development cycle.

Our interviewees suggested that recently, the shift from a more narrow focus on work to a broader view of interaction in people's lives has increased the role of designers in HCI research and practice. Industrial designers, communication designers, and newly minted interaction designers all began to play more important roles in the invention and development of radically new artifacts meant to address a broad set of problems and opportunities.

Our interviews with leading interaction designers showed that while they have strong agreement about the role design plays and the benefits it brings to HCI practice, designers lack clarity on what design research is or should be with respect to HCI research. In all cases our questions about design research performed outside of a specific design case caught our participants off guard. However, through probing on this issue, three roles for design to play in research emerged: (i) design researcher in service of a research community—working to help researchers ground and frame problems and communicate the impact; (ii) design researcher as critic of the HCI community—making artifacts that stimulate discussion of critical issues; and (iii) design researcher as pattern finder, finding patterns that lead to pattern languages.

One challenge a few interviewees noted for designers participating in research comes from the consultancy model that drives most design work. Since the majority of design research is paid for by the development industry, it is unlikely that this information, which provides a significant competitive advantage, would be openly shared.

HCI researchers on design and design research

While interaction designers could articulate their role within an HCI team, researchers were far less articulate about the role of interaction design. In our interviews with leading HCI researchers, we heard views of design as "the discovery of mental models", "a discipline focused on the whole instead of the parts", and "desire to understand users". However, the dominant view was that designers focused solely on the surface structure, or the visual aesthetics of software and hardware artifacts. This idea of design and designers as having a focus on decoration is a commonly held belief of design by most people [5].

When asked about what design research is and what design researchers do, the HCI researchers we interviewed had no concrete ideas. This is not surprising given the lack of clarity within the interaction design community on what design is. Instead, a common theme we heard was that it was up to the interaction designers working in research to invent what design research should be within the context of HCI.

A MODEL OF INTERACTION DESIGN RESEARCH WITHIN HCI RESEARCH

Our research model attempts to unite and advance the findings from the literature review and interviews described above in a format that complements current methods of HCI research. It follows from Christopher Frayling's concept of conducting *research through design* [14] where design researchers focus on making the *right* thing; artifacts intended to transform the world from the current state to a preferred state.

In our model (Figure 1), interaction design researchers engage wicked problems found in HCI. Examples of wicked problems include: (1) The design of smart home services for families where parents address the paradox of wanting to care and protect their children while also wanting to make them independent and children face the paradox of desiring the comfort and security their home and family provide while also wanting to step out and discover and invent who they are and might be. (2) The role of ubiquitous, assistive technology in aiding an elderly population to "age in place" in their own homes. It is wicked in that the stakeholders have conflicting goals including adult children who often want their parents out of the home in an environment that can better ensure their safety, and elder parents who have huge identity investments in their homes, and desire to remain, even when doing so creates tremendous social isolation.

Using our model, interaction design researchers integrate the *true* knowledge (the models and theories from the behavioral scientist) with the *how* knowledge (the technical opportunities demonstrated by engineers). Design researchers ground their explorations in *real* knowledge produced by anthropologists and by design researchers performing the upfront research for a design project. Through an active process of ideating, iterating, and critiquing potential solutions, design researchers continually reframe the problem as they attempt to make the *right* thing. The final output of this activity is a concrete problem framing and articulation of the preferred state, and a series of artifacts—models, prototypes, products, and documentation of the design process.

This research through design approach produces several beneficial contributions for the HCI community. First, design researchers identify opportunities for new technology or for advancements of current technology that will have significant impact on the world. This type of design research provides research engineers with inspiration and motivation for what they might build. Design researchers also undertake problem framing that helps identify important gaps in behavioral theory and models. In evaluating the performance and effect of the artifact situated in the world, design researchers can both discover unanticipated effects and provide a template for bridging the general aspects of the theory to a specific problem space, context of use, and set of target users.

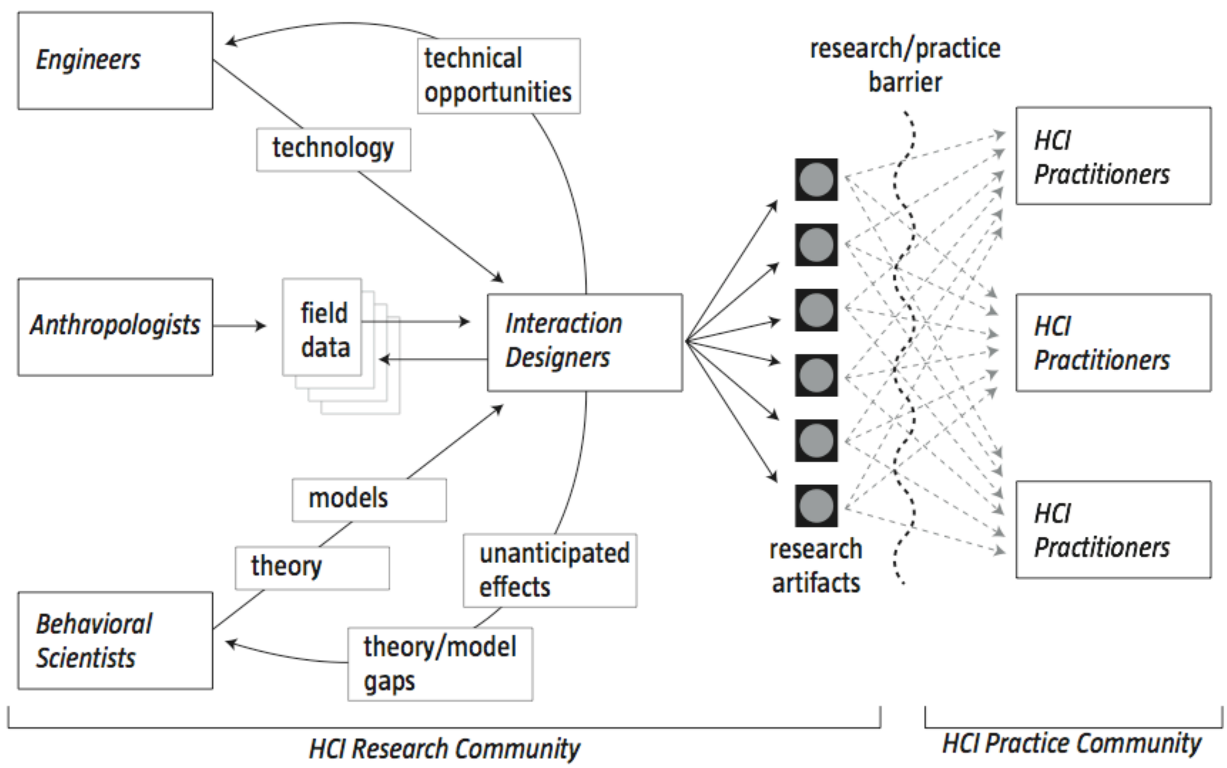


Figure 1. An illustration of the pathways and deliverables between and among Interaction Design Researchers and other HCI Researchers. The model emphasizes the production of artifacts as vehicles for embodying what "ought to be" and that influence both the research and practice communities.

Second, interaction design researchers create artifacts that provide concrete embodiments of theory and technical opportunities. These design exemplars then become an appropriate conduit for the transfer of HCI research to the practice community. Design artifacts are the currency of design communication. In education they are the content that teachers use to help design students understand what design is and how the activity can be done [7]. In research, they describe a vision of a preferred state, increasing the chance for knowledge transfer to the research, practice, and education communities. Through exposure to the ideas in the artifacts, the practice community can more easily observe the value of different theories, models, and technology, and this can motivate them to follow the threads back to the original research that might most impact their work.

Third, use of this model results in a holistic research contribution that reveals the framing of the problem and the balance the researchers have made between the intersecting and conflicting perspectives. The idea of contributing a whole closely resembles the work of systems engineers in HCI who focus on building whole systems. The Aware Home constructed at Georgia Tech provides a good example [11]. In this case the novelty was not in the construction of the individual elements, but in the integration of many technical research contributions from a variety of disciplines, into a single working system. The difference between this type of contribution and the design research contributions we propose involve both the intent and the process of the research. In making a technical contribution of a whole, engineers first develop a specification of what they need to make to meet a specific need. Next, they take a research focus asking questions such as can this be built? Is there a better way to build this?

In proposing a model of design research with a focus on the production of artifacts, we build on Nigel Cross's idea that design knowledge resides in the product [7]. The artifact reflects a specific framing of the problem, and situates itself in a constellation of other research artifacts that take on similar framings or use radically different framings to address the same problem. These research artifacts provide the catalyst and subject matter for discourse in the community, with each new artifact continuing the conversation. When several related research artifacts have been created, then researchers can use more traditional design research methods to analysis the artifacts and search for similar approaches designers have taken in addressing common problems. The artifacts made through design research have the potential to become pre-patterns [6] from which design patterns [1] can begin to emerge.

Our model departs from the roles of the design researcher discovered in our literature review and interviews: (i) design researcher as member of design practice team doing upfront project research; (ii) traditional knowledge producing design researcher studying design process and analyzing artifacts to discover patterns; (iii) design researcher as critic; and (iv) design research as framing and communication consultant in service of other researchers. Using our model, design researchers work in a way very similar to design practitioners, applying their strength at addressing under-constrained problems. This is in no way intended to discount the other design research roles that have already had significant impact on HCI. Instead, we are proposing an additional model of design research that seems particularly suited for interaction design researchers working in HCI research and allows design researchers to work more as a collaborative equal with other HCI researchers. An obvious criticism of this model is how in its use design researchers can distinguish their contributions as research and not as practice. This is a concern raised by Nigel Cross, who cannot consider normal works of practice to be regarded as research contributions [7].

We differentiate research artifacts from design practice artifacts in two important ways. First, the intent going into the research is to produce knowledge for the research and practice communities, not to make a commercially viable product. To this end, we expect research projects that take this research through design approach will ignore or de-emphasize perspectives in framing the problem, such as the detailed economics associated with manufacturability and distribution, the integration of the product into a product line, the effect of the product on a company's identity, etc. In this way design researchers focus on making the *right* things, while design practitioners focus on making *commercially successful* things.

Second, research contributions should be artifacts that demonstrate significant invention. The contributions should be novel integrations of theory, technology, user need, and context; not just refinements of products that already exist in the research literature or commercial markets. The contribution must demonstrate a significant advance through the integration. This aspect of a design research contribution makes particular sense in the interaction design space of HCI. Meteoric technological advances in hardware and software drive an aggressive invention of novel products in HCI and interaction design domains that are not as aggressively experienced by other design domains. While product designers might find themselves redesigning office furniture to meet the changing needs of work, interaction designers more often find themselves tasked with inventing whole new product categories.

Our model of design research allows interaction design researchers to do what designers do best: to study the world and then to make things intended to affect change. Our model provides a new channel for the power of design thinking, desired by many disciplines, to be unleashed as in a research context. Design researchers can contribute from a position of strength, instead of aping the methods of other disciplines as a means of justifying their research contribution.

CRITERIA FOR EVALUATING INTERACTION DESIGN RESEARCH WITHIN HCI

Many design researchers have made contributions using a research through design approach. While the idea is not new within the HCI and interaction design research community, there is no agreed upon standard of what research through design means nor what a high quality contribution should be. To help to formalize this research method, we propose a set of criteria, or four lenses for evaluating an interaction design research contribution: process, invention, relevance, and extensibility.

Process: One of the critical elements for judging the quality of an interaction design research contribution is the process. Like anthropologists making contributions in this science-dominated domain, there is no expectation that reproducing the process will produce the same results. Instead, part of the judgment of the work examines the rigor applied to the methods and the rationale for the selection of specific methods. In documenting their contributions, interaction design researchers must provide enough detail that the process they employed can be reproduced. In addition, they must provide a rationale for their selection of the specific methods they employed.

Invention: The interaction design research contribution must constitute a significant invention. Interaction design researchers must demonstrate that they have produced a novel integration of various subject matters to address a specific situation. In doing so, an extensive literature review must be performed that situates the work and details the aspects that demonstrate how their contribution advances the current state of the art in the research community. In addition, in articulating the integration as invention, interaction designers must detail how advances in technology could result in a significant advancement. It is in the articulation of the invention that the detail about the technical opportunities is communicated to the engineers in the HCI research community, providing them with guidance on what to build.

Relevance: Scientific research has a focus on validity. In engineering, this often means a demonstration of the performance increase or the function of their contribution. In behavioral science, validity means an experiment that disproves the null hypothesis. In both cases, the work must be documented in such a way that peers can reproduce the results. As mentioned above, this does not make sense to have as a requirement for a research through design approach. There can be no expectation that two designers given the same problem, or even the same problem framing, will produce identical or even similar artifacts. Instead of validity, the benchmark for interaction design research should be relevance. This constitutes a shift from what is true (the focus of behavioral scientists) to what is real (the focus of anthropologists). However, in addition to framing the work within the real world, interaction design researchers must also articulate the preferred state their design attempts to achieve and provide support for why the community should consider this state to be preferred.

Today, many design research contributions claiming to follow a research through design approach neglect to cast the work in terms of relevance. The design researchers follow a design process, but the motivation for their work, the detail on current situation, and on the preferred state are missing. Without this critical component, a research through design approach appears to be a self-indulgent, personal exploration that informs the researcher but makes no promise to impact the world.

Extensibility: The final criterion for judging successful design research is extensibility. Extensibility is defined as the ability to build on the resulting outcomes of the interaction design research: either employing the process in a future design problem, or understanding and leveraging the knowledge created by the resulting artifacts. Extensibility means that the design research has been described and documented in a way that the community can leverage the knowledge derived from the work.

EXAMPLES OF INTERACTION DESIGN RESEARCH WITHIN HCI

In order to demonstrate how the lenses might work for evaluating the quality of interaction design research in HCI, we provide examples of three interaction design cases that help illustrate different aspects of this model.

XEROX reprographics

FitchRichardsonSmith's work with Xerox in the early 1980s on the interaction design of reprographics machines provides an early example of research through design that produced design exemplars and a design language (an intentional pattern language) that can still be seen today in the interaction and behavior of copiers and printers. The design process was documented in an extended rationale, called *Principles for Constructing Communicative Objects and Object Systems for Interactive Dialogs*, and detailed the design and rationale for every element of a machine to support *positive* interaction [28].

Prior to this work, reprographics machines used in offices generally had a key operator: a trained technician who held the key to operate, maintain, and repair the machine. Design researchers working on this project reframed the problem from making a machine that was easier for a key-operator to maintain to making a machine that any office worker could walk up to, use, and fix if it had a paper jam. The prototypes produced (Figure 2) illustrated the idea that people could learn to operate the machines as they used them—rather than being trained, which was unheard of in the industry at that time. The design language included the use of green on the copy button and on the edge of the glass panel to indicate points of entry, and the color blue to indicate where users should interact with paper. Lighter shades indicated areas of frequent interaction and darker shades indicated areas with less frequent use. Texture indicated specific touch points. Finally, the prototypes provided concrete illustrations of how to provide instructions at the point of need. Evaluations of the prototypes revealed a shift in work practice that came about as a result of the new way of interacting with the machines [20].

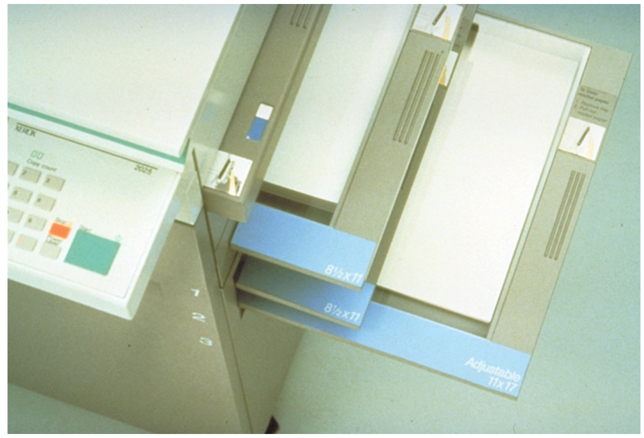


Figure 2. Xerox prototype machine.

In terms of invention, this work demonstrated an integration of the latest cognitive research on how people learn to interact with systems. In terms of relevance, it connected with the increasing need in the work place to empower workers to take more ownership and responsibility for the individual documents they were working on. One of the most valuable contributions was extensibility. The Xerox guidelines and rationale document communicated reusable information for extensions in design [27] but the machines themselves became objects that could be read by other designers outside of Xerox. This worked to transfer the knowledge to the practice community. Today, elements and resources from this interaction design research project can still be seen in almost every copier and printer.

Philips vision of the future

In 1995, Philips Design's Vision of the Future project explored possibilities for life and technology in the near future. Using a rigorous design process documented in the book Vision of the Future [19], this project examined how advances in technology would change family life along with other aspects of society. Multidisciplinary teams were brought together to propose directions for new products and services in four different domains of life: personal, domestic, public, and mobile. In terms of relevance, the work detailed how changes to traditional forms and behaviors of technical products technology could allow products to more easily integrate into the social life of people outside of work environments. For example, Figure 3 shows a mobile communication device housed in an aesthetic form not unlike a flower vase. The novel designs clearly had the intention of improving the quality of people's lives and provided a view of a preferred state. The work and the design process have also proven to be extensible, as numerous undergraduate and graduate programs in interaction design have imitated the project.

Figure 3. Concept from Philips' vision of the future project.

Apple Guides

This research through design project helps illustrate how interaction design research can feed back ideas to the HCI research community. In designing an interface to a multimedia database, the design team used theory from cognitive psychology to address the real world issue of people getting lost in hypertext interfaces [18]. The team chose to use black and white renderings of people dressed in historic costumes and set in historic contexts.

In evaluating this interface to see if the guide improved the navigation of the content, the design team discovered an unanticipated of effect. Participants no longer viewed the content as having the voice of an unbiased encyclopedia. Instead, they felt the content represented the opinion of the individual guide. Through dissemination of these evaluation findings, this design through research project helped to stimulate new technical research on the underlying technology to produce embodied agents and new behavioral research to understand the effect embodied agents had on users.

CONCLUSION

This paper has presented two years of iterative design efforts to explore and advance methods for interaction design researchers to make design research contributions that both integrate with and benefit the HCI research and practice communities. The work has resulted in a new model of interaction design research within HCI that allows design researchers to collaborate on an equal footing with HCI engineering and behavioral science researchers. In addition, it provides a set of critical lenses for evaluating what constitutes a good interaction design research contribution for researchers following this model.

The model provides five main benefits. First, it allows the HCI research community to engage with wicked problems that cannot be easily addressed through science and engineering methods. Second, it feeds back technology opportunities to the engineers and gaps in behavior theory and unexpected behaviors to the behavioral scientists, motivating new research. Third, it provides a new method for transferring knowledge produced in the HCI research to the HCI practice community, potentially increasing the likelihood this knowledge will move into products in the world. Fourth, it allows interaction designers to make research contributions that take advantage of the real skill designers possess—reframing problems through a process of making the *right* thing. Fifth, it motivates the HCI community to discuss preferred states and to reflect on the potential impacts research might have on the world.

We hope that in proposing this model, we can begin a serious discussion of the role of design and design thinking in HCI research. We will continue to evaluate and refine our model with practitioners and researchers. Additionally, we hope to formulate some changes to both HCI and interaction design education that will allow interaction design research to continue to grow in importance.

ACKNOWLEDGEMENTS

We wish to thank our colleagues at Carnegie Mellon's Human-Computer Interaction Institute (HCII) and at the School of Design for their participation, insights, and patience. In addition, we would like to thank Professor Dan Sieworiek, the director of the HCII, for his championing and financial support of this work.

REFERENCES

1. Alexander, C., Ishikawa, S., Silverstein, M., Jacobson, M., Fiksdahl-King, I., Angel, S. *A Pattern Language: Towns, Buildings, Construction*. Oxford University Press, 1977.
2. Apple Computer, Inc. *Macintosh Human Computer Interface Guidelines*, Addison-Wesley Professional, Reading, MA, 1992.
3. Bayazit, N. Investigating Design: A Review of Forty Years of Design Research. *Design Issues 20*, 1 (2004), 16-29.
4. Buchanan, R. Design Research and the New Learning. *Design Issues 17*, 4 (2001), 3-23.
5. Blevis, E., Lim, Y.K., & Stolterman, E. Regarding Software as a Material of Design. *Proc. of Wonderground*, Design Research Society, (2006).
6. Chung, Eric, Jason I. Hong, James Lin, Madhu K. Prabaker, James A. Landay, and Alan Liu. Development and Evaluation of Emerging Design Patterns for Ubiquitous Computing. *Proc. DIS 2004,* ACM Press (2004), 233-242.
7. Cross, N. Design Research: A Disciplined Conversation. *Design Issues 15*, 2 (1999), 5-10.
8. Cross, N. Designerly Ways of Knowing: Design Discipline Versus Design Science. *Design Issues 17*, 3 (2001), 49-55.
9. Dunne, A., Raby. F. *Design Noir: The Secret Life of Electronic Objects*. Birkhäuser, Basel, Switzerland, 2001.

10. Gaver, WW., Bowers, J., Boucher, A., Gellerson, H., Pennington, S., Schmidt, A., Steed, A., Villars, N., Walker, B. The drift table: designing for ludic engagement. *Ext. Abstracts CHI '04*, ACM Press (2004), 885-900.
11. Kidd, C.D., Orr, R., Abowd, G.D., Atkeson, C.G., Essa, I.A., MacIntyre, B., Mynatt, E.D., Starner, T., Newstetter, W. The Aware Home: A Living Laboratory for Ubiquitous Computing Research, *Proc. of the Second International Workshop on Cooperative Buildings, Integrating Information, Organization, and Architecture*, (1999) 191-198.
12. Fallman, D. Design-Oriented Human-Computer Interaction. *Proc. CHI 2003*, ACM Press (2003), 225-232.
13. Fallman, D. Why Research-Oriented Design Isn't Design-Oriented Research. *Proc. NordiCHI 2005*, Umea Institute of Design Press (2005).
14. Frayling, C. Research in Art and Design. *Royal College of Art Research Papers 1,* 1 (1993),1-5.
15. Laurel B. Design Research: Methods and Perspectives. MIT Press, Cambridge, MA, 2003.
16. Löwgren, J. Applying Design Methodology to Software Development. *Proc. of DIS 1995*, ACM Press (1995), 87-95.
17. Nelson, H.G. and Stoltermann, E. *The Design Way: intentional change in an unpredictable world*. Educational Technology Publications, Englewood Cliffs, NJ, 2003.
18. Oren, T., Salomon, G., Kreitman, K., Don, A.: Guides: Characterizing the Interface. In (Eds: Laurel, B.): *The Art of Human-Computer Interface Design*, Addison-Wesley (1990) 355-365.
19. Philips Design, Vision of the Future *V&K* (1995)
20. Rheinfrank, J., Hartman, W., Wasserman, A., *Design for usability: crafting a strategy for the design of a new generation of Xerox copiers, Usability: turning technologies into tools*, Oxford University Press, Inc., New York, NY, 1992.
21. Rittel, H.W.J., Webber, M.M. Dilemmas in a General Theory of Planning. *Policy Sciences 4*, 2 (1973), 155-66.
22. Schön, D.A. *The Reflective Practitioner: How professionals think in action*. Temple Smith, London, 1983.
23. Simon, H.A. *The Sciences of the Artificial*, MIT Press, Cambridge, MA, 1969.
24. van Duyne, D.K., Landay, J.A., Hong, J. I. *The Design of Sites: Principles, Processes, and Patterns for Crafting a Customer-Centered Web Experience*, Addison-Wesley, Reading, MA, 2003.
25. Wolf, T.V., Rode, J.A., Sussman, J., Kellogg, W.A. Dispelling Design as the 'Black Art' of Chi. *Proc. of CHI 2006*, ACM Press (2006), 521-530.
26. Wright, P., Blythe, M., McCarthy, J. User Experience and the Idea of Design in HCI. *Lecture Notes in Computer Science*, Stephen W. Gilroy and Michael D. Harrison eds. Springer, Berlin/Heidelberg (2006). 1-14.
27. Xerox Corporation. *Design Guidelines: Industrial design and Human Factors for Reprographics Products* (1985).
28. Xerox Corporation. *Principles for Constructing Communicative Objects and Object Systems as Interactive Dialogs* (1985).

Available online at www.sciencedirect.com
ScienceDirect
Int. J. Human-Computer Studies 65 (2007) 737–743
International Journal of Human-Computer Studies
www.elsevier.com/locate/ijhcs

HCI reality—an 'Unreal Tournament'?

Christoph Bartneck*, Matthias Rauterberg

Department of Industrial Design, Eindhoven University of Technology, Den Dolech 2, 5600MB Eindhoven, The Netherlands

Received 21 February 2006; received in revised form 6 March 2007; accepted 10 March 2007
Communicated by K.S. Severinson Eklundh
Available online 23 March 2007

Abstract

The cooperation between designers, engineers and scientists in the human–computer interaction (HCI) community is often difficult, and can only be explained by investigating the different paradigms by which they operate. This study proposes a paradigm model for designers, engineers and scientists, using three barriers to separate the professions. We then report on an empirical study that attempted to validate the understand/transform world barrier in the paradigm model using an online questionnaire. We conclude that the used 'Attitude About Reality' scale was unsuitable for measuring this barrier, whereas information about the educational background of the participants was a good predictor for the self-reported profession (designer, engineer or scientist). Interestingly, among the three professions, engineers appear to be the cohesive element, since they often have dual backgrounds, whereas very few participants had dual science/design backgrounds. Engineers could, therefore, build a bridge between designers and scientists, and through their integrative role, could guide the HCI community to realizing its full potential.

Keywords: Human–computer interaction; Community; Paradigm; Design; Engineering; Science; Attitude About Reality

1. Introduction

The human–computer interaction (HCI) community is diverse. Academics and practitioners from science, engineering and design contribute to its vivid development, but communication and cooperation between the different groups is often challenging. The Association for Computing Machinery (ACM) Computer Human Interaction (CHI) conference, 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 CHI2006 conference ignited a shouting match between academics and practitioners (Arnowitz and Dykstra-Erickson, 2005). The intensity of the situation could be compared to scenes from the multiplayer video game, 'Unreal Tournament'. Both groups defended their access to the conference through the different publication formats, such as paper sessions, panels, and case studies, similar to how, 'Unreal Tournament' players fight for markers in the 'domination' game mode. This outbreak of emotion illustrates the tension between the different groups and it can be explained by taking a closer look at the paradigms by which they operate, and at the barriers that separate them. Snow (1964) 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. (Snow, 1964, p. 16)

After addressing these theoretical aspects, we will present an empirical study that attempts to verify one of the barriers between the paradigms and discuss its consequences. A better understanding of the different paradigms within the HCI community could help to prevent wasting

*Corresponding author. Tel.: +31 40 2475175; fax: +31 40 2475376.
E-mail addresses: c.bartneck@tue.nl (C. Bartneck), g.w.m.rauterberg@tue.nl (M. Rauterberg).

1071-5819/$ - see front matter © 2007 Elsevier Ltd. All rights reserved.
doi:10.1016/j.ijhcs.2007.03.003

any additional time and energy on shouting matches, and could lead to a mutual beneficial cooperation.

'Paradigm' is defined in the Kuhnian sense as a disciplinary matrix that is composed of those (a) shared beliefs, (b) values, (c) models, and (d) exemplars that guide a community of theorists and practitioners (Kuhn, 1970). We propose three barriers that can be used to distinguish the paradigms of the three different disciplines (see Fig. 1): designers {D}, engineers {E} and scientists (in particular social scientists), {S}:

(1) knowledge representation (explicit {S, E} versus implicit {D});
(2) view on reality (understanding {S} versus transforming reality {D, E}); and
(3) main focus (technology {E} versus human {D, S}).

Barrier 1: Engineers {E} and scientists {S} make their results explicit by publishing in journals, books and conference proceedings, or by acquiring patents. Their body of knowledge is externalized and described outside of the individual engineer or scientist. These two communities revise their published results through discussion and control tests among peers. On the other hand, designers'{D} results are mainly represented by their concrete designs. The design knowledge necessary to create these designs lies within the individual designer, mainly as implicit knowledge, often referred to as *intuition* (see Dorfman et al., 1996). To make better designs, the designer has to become more experienced. After gaining considerable experience and intuition, designers tend to reflect (Schön, 1991) and publish their views on design (Dorst, 2003). Even so, the foundation of these reflections lies within the individual designer's experiences of reality.

Barrier 2: Engineers {E} and designers {D} transform the world into preferred situations (Simon, 1996; Vincenti, 1990), while scientists {S} mainly attempt to understand the world through the pursuit of knowledge covering general truths or the operation of general laws (definition taken from Encyclopaedia Britannica). This difference between {E, D} and {S} is of particular interest to our investigation since a preferred state could also be the state of knowing and since understanding also requires the use of synthesis. The following model illustrates the relationship between 'abstracting' from reality (for understanding) and 'concretization' (for transforming reality; see Fig. 2).

Barrier 3: Scientists {S} and designers {D} are predominantly interested in humans in their role as possible users. Designers are interested in human values, which they transform into requirements and eventually solutions. Scientists in the HCI community are typically associated with the social or cognitive sciences. They are interested in the users' abilities and behaviors such as perception, cognition and action. Engineers {E} are mainly interested in technology, which includes software for interactive systems. They investigate the structure and operational principles of these technical systems to solve certain problems.

Given a reality at time t_1, science in the positivistic paradigm observes and analyzes particular phenomena in this reality, makes proper abstractions, and tries to predict similar phenomena for reality at time t_2. To preserve a stable reality [reality (t_1) = reality (t_2)], science in the positivistic paradigm has to operate under the essential assumption that model and theory are not a part of reality [(model, theory) $\notin$ {reality}]. The theory ('*res cogitans*') itself is clearly separated from and does not influence the described phenomena ('*res extensa*'; see Descartes, 1644 and more recently Dreyfus, 1979, Chapter 7, and Blackmore, 1999, Chapter 17).

For example, the theory of gravity explains and predicts certain phenomena, such as falling apples, but it neither influences nor changes the phenomenon of 'gravity'. In this sense, models and theories of science in our modern positivistic paradigm are not part of the investigated and described reality; but they are apart from this reality. For this concept of reality we will use the lowercase style. We will use the term REALITY in the uppercase style for the

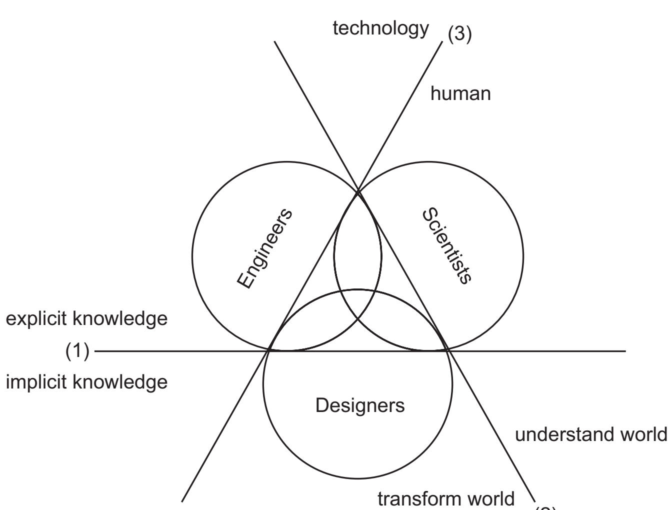


Fig. 1. Theoretically hypothesized paradigm model of designers, engineers and scientists with the three main barriers distinguishing them.

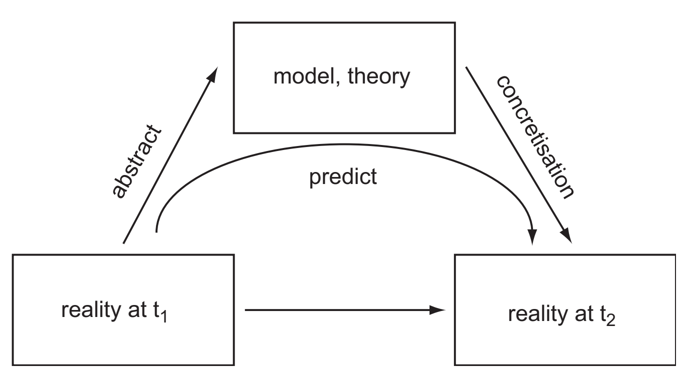


Fig. 2. Progress model (adapted from Rauterberg, 2006). Scientists create models and theories of reality through abstraction with the aim of predicting reality. Designers and engineers concretize the abstract models and theories into artifacts that improve reality.

broader meaning of the term as the union of model, theory and reality. The underlying mechanism to guarantee the fulfillment of the assumption [reality (t_1) = reality (t_2)] is reductionism via abstraction. Any differences in empirical measurements and observations between t_1 and t_2, such as noise, are interpreted as just accidental factors, which do not contradict the theory and/or the underlying principle. Armed only with knowledge based on theories developed under the positivistic paradigm, the design of a concrete artifact is impossible, because the knowledge in these theories is purified from the concrete and changing contextual factors between reality at t_1 and at t_2. This lack of specific knowledge for any concretization, such as craft skills provided via experiences and intuition, gives design and engineering disciplines their right to exist. Dreyfus and Dreyfus (1992) and Dreyfus et al. (1986), stimulated a very important discussion about the importance and boundaries of and necessity for *intuitive* expertise, complementary to artificial expert systems which just follow explicitly given rules.

Therefore, activities governed by a *constructivistic paradigm* claim to influence the reality and to change this reality via the developed artifacts [reality $(t_1) \neq$ reality (t_2)], and in fact they do! The design and engineering disciplines develop knowledge to make concretization possible. This knowledge realized in the form of models and artifacts can be interpreted as part of the REALITY, and not apart from it [(model, artifact) $\in$ {reality}]. But how can design and engineering disciplines guarantee a stable reality? If models and artifacts are seen as part of this REALITY, such as a subset of this REALITY under consideration, then any action, which changes this subset, changes the whole REALITY set as well. So, constructive disciplines such as design and engineering cannot guarantee a stable reality, and indeed they do *not* want to (Klemm, 1964).

Scientists, with their logical positivistic paradigm on the one side, and engineers and designers with their constructivistic paradigm on the other side, appear to have different attitudes toward REALITY. Our study attempted to find empirical proof of this difference. We hypothesized that the Attitude About Reality (AAR) scale (Unger et al., 1986) might be useful for measuring this difference. This bi-polar scale ranges from 'logical positivism' on the one side to 'social constructivism' on the other side. Unger et al. (1986) defines *logical positivism* as follows:

> Reality is relatively fixed and objectively accessible. Logical positivism also states that meaning is operationally defined and therefore replicable across social contexts, and that reality will be increasingly uncovered by the use of more and more refined measurement techniques.

Social constructivism is defined by Unger et al. (1986) as follows:

> Meaning is defined by our linguistic and conceptual categories; these categories are the product of social-consensus process that is neither predictable nor progressive in nature. This view 'invites one to challenge the objective basis of conventional knowledge' (Gergen, 1985) and to focus instead upon processes of negotiated understanding as the critical events through which to analyze reality, as we know it. The degree to which a particular form of understanding prevails or is sustained across time is not seen as fundamentally dependent upon its empirical validity.

If scientists operate under the logical positivistic paradigm, then they should score higher on the AAR scale compared to designers and engineers that work under the constructivistic paradigm. Our main research question is whether the different paradigms of designers, engineers and scientists do indeed lead to different views on reality as measured through the AAR scale.

2. Method

We conducted a study in which the participant's self-reported profession, their educational background and the AAR score was recorded. The data were gathered through an online questionnaire. Invitations to participate in the study were posted on several HCI mailing lists, including the ACM's CHI-Announcements list, British HCI Group's BCS-HCI news list and the German GI-Fachgruppe Software-Ergonomie SW-Ergo list. While online surveys have some methodological difficulties (Kaye and Johnson, 1999), such as the submission of duplicates, it still offers the broadest access to a given community. Our study focuses on the HCI community and not on the general public. Therefore sampling problems apply only to a lesser degree. It can be assumed that most HCI community members have access to a computer and the Internet. We checked the submission dates and times in combination with the IP addresses of the computer from which the survey was completed, to prevent duplicate submissions.

2.1. Measurements

First, the participants were asked to declare in what academic fields they had a bachelor's, master's or doctoral degree. The participants could select multiple answers, even within the different educational levels. They could also decide not to give an answer for a certain educational level. If, for example, the participant did not have a Ph.D. degree then they would not check any of the offered choices for the Ph.D. degree.

Given the diversity of educational programs it appeared to be unwise to include an open question for the academic fields. The option 'Other' was included to identify participants that diverged from the given set of academic fields.

The scores on the educational background were transformed in the following manner. We categorized each education field as either design, engineering or science according to the following schema:

Design: Architecture, Fashion Design, Industrial Design (Product, Packaging, etc.), Visual Design (Graphic, Interaction, Information, etc).

Engineering: Bioengineering/Biomedical Engineering, Civil Engineering, Electrical Engineering, Environmental Engineering, Material Science and Engineering, Mechanical Engineering, Software Engineering.

Science: Biology, Chemistry, Computer Sciences, Geology, History, Mathematics, Medicine, Philosophy, Physics, Psychology, and Sociology.

The schema might have a certain ambiguity. For example, it can be argued if Mathematics is really a science (Jaffe, 1997). Furthermore, its relationship to Computer Science and Software Engineering is not completely clear. However, it appears to have been generally categorized as a science by encyclopedias such as 'Encyclopedia Britannica'.

Afterwards, the questionnaire inquired how participants would categorize their main profession as it is now: designer, engineer or scientist. This self-reported categorization was a forced choice selection question and they were not allowed to make multiple choices. Finally, the participants had to fill in the AAR questionnaire (Unger et al., 1986) which consisted of 40 questions. Each question had to be answered using a sevenpoint Likert scale (see Fig. 3). The answers of all inverse items were transformed before undergoing the analysis. The AAR score is the sum of all items ranging from a minimum of 40 to a maximum of 280 points.

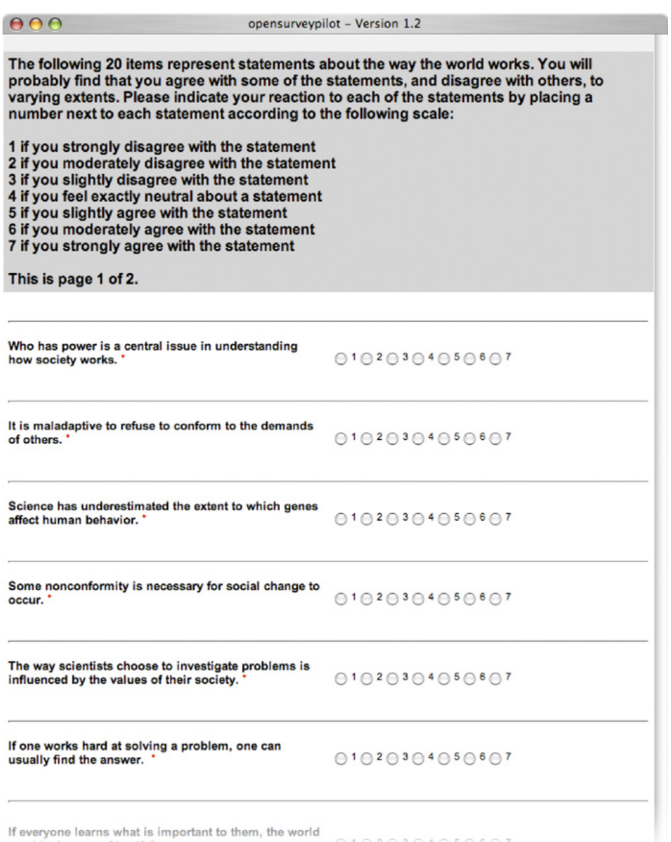


Fig. 3. Example screenshot of the questionnaire.

We then estimated how much time each participant spent in design (d), engineering (e), and science (s). Since the education systems vary among different countries we assumed that on average a bachelor's degree requires four years, a master's requires two years, and a Ph.D. requires four years of education. If a person had a bachelor's in mechanical engineering (e) and a master's in industrial design (d) then the person would receive a score of $d = 2$ and $e = 4$. In addition, we noted the highest educational level (*eduLevel*) of the participant (1 = bachelor' s, 2 = master' s and 3 = Ph.D.). In the above example, the highest level would have been the master's degree. We then calculated the specialization of the participant (*propEduInProfession*). The number of years of education the participant spent on his/her profession were divided by the total years of education ($d + e + s$). If the participant in the previous example considers themselves to be a designer then their score would be *propEduInProfession* = 0.33 (2/6) and if they considered themselves to be an engineer then their score would be *propEduInProfession* = 0.66 (4/6).

To summarize, we recorded the participants' age, gender, profession, highest education degree (*eduLevel*), and AAR score.Based on an estimation of the participants' education years in design (d), engineering (e), and science (s), we calculated the participants' self-reported specialization (*propEduInProfession*).

2.2. Participants

Given the international scope of the various mailing lists to which the invitations were sent, it can be assumed that the participants originated from several different countries. Of the 128 people that filled in the questionnaire, a total of 114 were used for the analysis. Participants who indicated 'Other' for education were excluded from the analysis ($N = 12$), since no further information on them was available. The remaining 114 participants identified themselves as designers ($N = 26$), engineers ($N = 33$) and scientists ($N = 55$). Table 1 shows the gender and profession frequencies. More men ($N = 81$) than women ($N = 33$) participated in the study and most women were scientists. A χ^2 test revealed no significant correlation between profession and gender ($\chi^2 = 1.850$; df = 2; $p = 0.397$). Table 2 summarizes the education levels per

Table 1
Frequencies of gender and profession

	Self-reported profession			Total
Gender	Designer	Engineer	Scientist	
Female	7	7	19	33
Male	19	26	36	81
Total	26	33	55	114

Table 2
Frequencies of the education levels per profession

Self-reported profession				Total
Education level	Designer	Engineer	Scientist	
Bachelor	5	9	6	20
Master	18	16	23	57
Ph.D.	3	8	26	37
Total	26	33	55	114

self-reported profession. Most participants had at least a master's degree ($N = 57$) and the scientists in particular tended to have a Ph.D. degree ($N = 26$). A Chi-Square test revealed that there was a significant correlation ($\chi^2 = 13.788$; $\mathrm{df} = 4$; $p = 0.008$) between profession and *eduLevel*.

3. Results

A reliability analysis for the 40 AAR items for all 114 participants resulted in a Cronbach's alpha of 0.614, which gives us sufficient confidence in the reliability of this questionnaire. An analysis of variance (ANOVA) was conducted with self-reported profession as the independent variable and the total AAR score as the dependent variable. The Levene's Test for equality of variance was not significant ($p = 0.278$) and therefore the variance can be assumed to be homogeneous. Profession does not have a significant influence on the total AAR ($F(2, 111) = 0.046$, $p = 0.955$). The scores for designer (168.15), engineer (168.33) and scientist (167.42) were only slightly above the middle value of the AAR (160). Next, we conducted a factor analysis of the 40 items on the AAR scale using a varimax rotation. It revealed 13 factors with eigenvalues greater than 1 after 25 iterations. The interpretation of these 13 factors was not the main aim of our study.

A discriminant analysis was performed to determine to what degree the variables measuring the participants' educational background (*d, e, s, eduLevel* and *propEduInProfession*) and gender, predicts the participants' self-reported profession. Table 3 shows the pooled within groups correlation between the variables and the two discriminant functions. The variables *s* and *eduLevel* largely correlate with the first function while the remaining variables correlate with the second function. The variable *propEduInProfession* correlates with both functions, which comes as no surprise, since the proportion spent within a certain profession is independent of the professions themselves. Function 1 may be labeled 'abstract orientation' and function 2 'concrete orientation'. This interpretation would be congruent with Fig. 2, which illustrates the role of abstracting and concretizing in the progress of science.

The sample data were randomly split into two groups. The first group was used for the creation of the discriminant model (original group) and the second group was used to validate the model (cross-validated group). In total 81.6% of the original cases and 77.2% of the cross-validated cases were correctly classified by the discriminant model (see Table 4). Interestingly, designers are sometimes incorrectly classified as engineers, and engineers are sometimes wrongly classified as scientists. Scientists, however, are rarely classified as anything but scientists (see Table 4).

Table 3
Pooled within groups correlation between the discriminating variables (*d, e, s, eduLevel, propEduInProfession*, gender) and the discriminant functions

	Function	
	1	2
s	0.397[a]	0.018
eduLevel	0.242[a]	0.180
d	−0.498	0.755[a]
propEduInProfession	0.513	0.565[a]
e	0.006	−0.542[a]
Gender	−0.075	−0.160[a]

[a]Largest absolute correlation between each variable and any discriminant function.

Table 4
Predicted membership accuracy of the self-reported profession category based on the variables *d, e, s, eduLevel, propEduInProfession*, and gender

		Profession	Predicted group membership			Total
			Designer	Engineer	Scientist	
Original	Count	Designer	16	10	0	26
		Engineer	1	28	4	33
		Scientist	1	5	49	55
Cross-validated	Count	Designer	16	10	0	26
		Engineer	1	23	9	33
		Scientist	1	5	49	55

A second discriminant analysis was performed to investigate the predictive power of age, gender, educational background (*d,e,s,eduLevel*, *propEduInProfession*) and AAR on the participants' profession (see Table 5).

The sample data were randomly split into two groups. The first group was used for the creation of the discriminant model (original group) and the second group was used to validate the model (cross-validated group). The 79.8% of original cases and 74.6% of the cross-validated cases were correctly classified by the discriminant model. The classification accuracy did not improve (compare Tables 4 and 6). Including AAR as a predictor did not improve the prediction accuracy.

4. Discussion and conclusion

Among our three chosen disciplines the differences in AAR are small and insignificant. The AAR scale appears to be ineffective in detecting the differences between designers, engineers and scientists. The results of the factor

Table 5
Pooled within groups correlation between the discriminating variables (age, gender, educational background [*d,e,s,eduLevel* and *propEduInProfession*] and AAR) and the discriminant functions

	Function	
	1	2
propEduInProfession	0.505[a]	0.472
s	0.381[a]	−0.029
eduLevel	0.237[a]	0.141
d	−0.459	0.767[a]
e	−0.007	−0.510[a]
Age	−0.178	−0.227[a]
Gender	−0.076	−0.141[a]
AAR	−0.020	−0.023[a]

[a]Largest absolute correlation between each variable and any discriminant function.

Table 6
Predicted membership accuracy of the self-reported profession category based on the variables *d, e, s, eduLevel, propEduInProfession*, gender and AAR

		Profession	Predicted group membership			Total
			Designer	Engineer	Scientist	
Original	Count	Designer	16	10	0	26
		Engineer	1	28	4	33
		Scientist	1	7	47	55
Cross-validated	Count	Designer	16	10	0	26
		Engineer	3	22	8	33
		Scientist	1	7	47	55

analysis strengthen this impression. Jackson and Jeffers (1989) identified only three factors of the AAR scale and labeled them, 'societal determinism', 'individual determinism', and, 'variable determinism'. In contrast, our factor analysis revealed 13 different factors. This difference might be explained by the diversity of the participants in our study. Both, Jackson and Jeffers (1989) and Unger et al. (1986), used a homogenous populations (e.g., undergraduate psychology students), whereas our study included participants from diverse backgrounds and ages. The AAR scale appears to be ineffective for heterogenous groups, which is unfortunate, since its potential value lies in explaining differences between conflicting groups. In contrast, the educational background of the participants has a much higher prediction accuracy of 77–82% for the participants' self-reported profession.

We therefore focus our further discussion on the educational backgrounds of the designers, engineers and scientists. The term 'scientist' is somewhat ambiguous. People may consider themselves to be scientists because they work for an academic institute or because they have a Ph.D. degree, or because they conduct scientific studies. An engineer working for a university might consider himself a scientist even if he only works on engineering tasks. Most medical doctors have a Ph.D. degree but work as general practitioners. However, to have a successful university/academic career, it is generally necessary to have a Ph.D. degree, which understandably requires time. Our data show that 70% of the participants that consider themselves to be scientists have a Ph.D. degree and as a result, have spent more time on their education. Furthermore, *propEduInProfession* is the best predictor for scientists when all variables are considered (see Table 5). They tend to stay within the same scientific fields while engineers and designers change their fields more often. In general, the education history of a person predicts in over 80% of the cases their self-reported profession correctly. This result indicates the importance of the education for the development of professions and thereby the differentiations of the disciplines within the HCI community. Perhaps an overhaul of our education systems, to include more diverse courses, is required in order to obtain increased cooperation between disciplines. Designers could, for example, avail of basic courses in experimental methodology and statistics, while scientists could be granted access to interface design courses.

It also becomes apparent that engineers are the binding element between designers and scientists. Designers and scientists sometimes have an engineering education and engineers sometimes have a design or science education. In contrast, scientists and designers rarely have a design or science education, respectively. The intersection between science and design as shown in Fig. 1 appears to be small. A more realistic model of the people within the HCI community is shown in Fig. 4 in which the three disciplines are aligned with engineering in the centre, flanked on either side by science and design.

Engineering shares much of its knowledge with science (Vincenti, 1990) and it can be argued that on the very grounds on which the claim of superiority is made for scientific knowledge, engineering knowledge is shown to be far more reliable, secure and trustworthy than scientific knowledge (Pitt, 2001). Engineers require such knowledge to build artifacts on which our lives depend, e.g., cars and houses must be safe and engineers have no margin for error when evaluating the strengths of materials.

Since engineers are more likely to have education in multiple fields than scientists or designers, they tend to acquire the knowledge and skills of the other professions. This enables them to speak and empathize with the other fields. They therefore could bridge the gap between

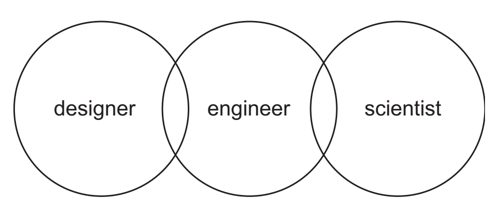


Fig. 4. Empirically improved paradigm model of designers, engineers and scientists.

designers and scientists and guide the HCI community to its full potential. To do so they need to be open-minded and collaborate with both the disciplines. In general, people with educational backgrounds in at least two disciplines might play the key role.

5. Future work

This study tried to find empirical proof of the AAR barrier between scientists on the one side and engineers and designers on the other side (see Fig. 1) by utilizing the AAR questionnaire. Unfortunately, we found that the AAR questionnaire was unsuitable for this task. To further empirically evaluate our model of the HCI community, it would be necessary to systematically test *all* barriers. In addition, one could consider a variation of the methodology in which the participants would be allowed to give multiple-choice answers for the self-reported profession.

Acknowledgments

The authors would like to thank the participants of this study who took the time to complete the questionnaire and who provided relevant and interesting feedback. In addition, we would like to thank Paul Locher for his support and in particular Rhoda Unger for making the AAR questionnaire available to us. We also would like to thank Aoife Currid for editing the manuscript.

References

Arnowitz, J., Dykstra-Erickson, E., 2005. CHI and the Practitioner Dilemma. Interactions 12 (4), 5–9.

Blackmore, S.J., 1999. The Meme Machine. Oxford University Press, Oxford, England, New York.

Descartes, R., 1644. Principia Philosophiae. Danielem Elzevirium, Amsterdam.

Dorfman, J., Shames, V.A., Kihlstrom, J.F., 1996. Intuition, incubation, and insight: implicit cognition in problem solving. In: Underwood, G. (Ed.), Implicit Cognition. Oxford University Press, Oxford, UK.

Dorst, C.H., 2003. Understanding Design. BIS Publisher, Amsterdam.

Dreyfus, H.L., 1979. What computers can't do: the Limits of Artificial Intelligence, revised ed. Harper & Row, New York.

Dreyfus, H.L., Dreyfus, S.E., 1992. What Computers Still Can't Do : a Critique of Artificial Reason. MIT Press, Cambridge, MA.

Dreyfus, H.L., Dreyfus, S.E., Athanasiou, T., 1986. Mind over Machine : the Power of Human Intuition and Expertise in the Era of the Computer. Free Press, New York.

Gergen, K.J., 1985. The social constructionistic movement in modern psychology. American Psychologist 40, 266–275.

Jackson, L.A., Jeffers, D.L., 1989. The Attitude About Reality Scale: a new measure of personal epistemology. Journal of Personality Assessment 53 (2), 353–365.

Jaffe, A., 1997. Proof and the evolution of mathematics. Synthese 111 (2), 133–146.

Kaye, B.K., Johnson, T.J., 1999. Research methodology: taming the cyber frontier-techniques for improving online surveys. Social Science Computer Review 17 (3), 323–337.

Klemm, F., 1964. A History of Western Technology. MIT Press, Cambridge.

Kuhn, T.S., 1970. The Structure of Scientific Revolutions, second ed. University of Chicago Press, Chicago.

Pitt, J.C., 2001. What engineers know. Techne 5 (3), 17–30.

Rauterberg, M., 2006. HCI as an engineering discipline: to be or not to be!? African Journal of Information and Communication Technology 2 (4), 163–184.

Schön, D.A., 1991. The Reflective Practitioner: How Professionals Think in Action. Arena, Aldershot England.

Simon, H.A., 1996. The Sciences of the Artificial, third ed. MIT Press, Cambridge, MA.

Snow, C.P., 1964. The Two Cultures: and a Second Look, second ed. University Press, Cambridge.

Unger, R.K., Draper, R.D., Pendergrass, M.L., 1986. Personal epistomology and personal experience. Journal of Social Issues 42 (2), 67–79.

Vincenti, W.G., 1990. What Engineers Know and How They Know it: Analytical Studies from Aeronautical History. Johns Hopkins University Press, Baltimore.

Subjects, Objects, Data and Values

Paper presented by Robert M Pirsig at the Einstein Meets Magritte Conference held in Brussels from May 29 to June 3 1995.

The title of what I have to say today is "Subjects, Objects, Data and Values." It concerns the central theme of this conference -- the meeting of art and science. Science is all about subjects and objects and particularly data, but it excludes values. Art is concerned primarily with values but doesn't really pay much attention to scientific data and sometimes excludes objects. My own work is concerned with a Metaphysics of Quality that can cross over this division with a single overall rational framework.

When I sent in the title of this paper in February I hadn't written this paper yet and so I kept the title very general to allow myself plenty of room. Now the paper is finished and I can add a subtitle that is more specific. The subtitle is, "Some Connections Between the Metaphysics of Quality and Niels Bohr's Philosophy of Complementarity." As I see it, Bohr's Complementarity and the Metaphysics of Quality stand midway between Einstein and Magritte. I have concentrated on Bohr's work as a way of making the larger connection.

I want to start with a famous conference that occurred here in Brussels in October 1927. It was the Fifth Physical Conference of the Solvay Institute. Here is a brief account of what happened, described by Bohr's biographer, Ruth Moore:

Bohr and Einstein were there, "as well as nearly all others who were contributing to theoretical physics. Lawrence Bragg and Arthur Compton came from the United States. DeBroglie, Born, Heisenberg, and Schrodinger all were to speak on the formulation of the quantum theory.

"The subject was 'Electrons and Photons.' To leave no doubt that it was directed to the main question, the theme embroiling all of physics, discussion was centered around the renunciation of certainty implied in the new methods [of physics] ... Bohr was invited to give the conference a report on the epistemological problems confronting quantum physics. By asking him to speak on the science of knowledge and the grounds for it, the conference gave him full opportunity to present Complementarity. There was no avoidance; the issue had to be directly faced.

"Excitement mounted as Einstein rose to speak. He did not keep them long in suspense. He did not like uncertainty. He did not like the abandonment of 'reality'. He did not think Complementarity was an acceptable solution, or a necessary one. 'The weakness of the theory lies in the fact that on the one hand, no closer connection with the wave concept is obtainable,' he said, 'and on the other hand that it leaves to chance the time and the direction of the elementary processes.'

"A dozen physicists were shouting in a dozen languages for the floor. Individual arguments were breaking out in all parts of the room. Lorentz, who was residing pounded to restore order. He fought to keep the discussion within the bounds of amity and order. But so great was the noise an the commotion that Ehrenfest slipped up to the blackboard, erased some of the figures that filled it, and wrote: 'The Lord did there confound the language of all the earth.'

"As the embattled physicists suddenly recognized the reference to the confusion of languages that beset the building of the tower of Babel, a roar of laughter went up. The first round had ended." (Moore 164)

The conference was carried on in events but also in private meetings and personal conversations with "thought experiments" carried out where physical conditions were imagined and results were predicted on the basis of known scientific facts. Behind the thought experiments was an all-important question of scientific certainty. Bohr was saying that the particles that constitute our material universe can only be described in terms of statistical probability and never in terms of absolute certainty. He regarded the development of the quantum revolution as in a certain sense "complete." Quantum theory need no longer await some enlightening revelation that would put everything right from a classical point of view.

Einstein wasn't having any of it. Quantum theory was not complete, he said. The universe is not ultimately a set of statistics. It was at one of those meetings that Einstein asked his famous question, "Do you really believe God resorts to dice playing?"

Thus began the controversy over Complementarity that continued for the rest of Bohr's live. It seems that I have heard about this famous schism all my life and wondered what it was about but never thought I would ever study it because I have no background in physics or mathematics. However, after my second book Lila came out in 1991 a friend in Norway wrote me that there was some attention being paid to Lila in Copenhagen by followers of Niels Bohr. It was suggested that the Metaphysics of Quality was similar to the Copenhagen Interpretation of the Quantum Theory 1. That sounded like good news to me and something I should look into. When similarities of this sort exist, they can either be an odd coincidence or they can be evidence that both systems of thought are describing something that is true independently of either thinker. Where the approaches are very different each can sometimes throw new light on the other. So when the invitation came to speak here I decided to make it the topic of today's paper. If the Copenhagen Interpretation, which is a dominant explanation of quantum theory today, agrees with the Metaphysics of Quality and if the Metaphysics of Quality is a correct theory of art, then there may be here a unified theory of art and science. Einstein will have met Magritte and the purpose of this conference will have been to some extent fulfilled.

Quantum theory

The first thing I discovered is that the volume of literature on quantum theory is enormous, and to a non-mathematician much of it is inscrutable. Physicists who do try to explain quantum theory in common language point out what a terrible burden it is to try to discuss it in non-mathematical terms. For me, a non-mathematician, it is also a burden to deal with secondary sources on the problem without knowing what the original mathematical language means. But there are two aspects to quantum theory: the mathematics of quantum theory and the philosophy of quantum theory. They are very deeply separated. The first seems to work very well. The second does not seem to work very well. Most physicists use the mathematics of the quantum theory with complete confidence and completely ignore the philosophy. I'm going to reverse that and concentrate on the philosophy and bypass the math. I have been working almost entirely from secondary sources and have placed heaviest reliance on a book by Henry J. Folse called The Philosophy of Niels Bohr. I have read that there are many variations of the Copenhagen Interpretation and Bohr's philosophy of Complementarity is only one of them. But it is the first and to simplify matters I have stayed with it alone.

For those of you who are as unfamiliar with quantum physics as I am I'll try to give a minimum summary here of what brought things to this state of conflict in 1927:

Before 1900 there existed in physics a problem known as "the ultra-violet catastrophe." Radiation from black bodies was not behaving according to predictions. In 1900 Max Planck solved this problem by theorizing that the radiant energy was coming in packets, rather than in a continuous flow. In 1905 Einstein saw that light was doing the same thing and named these packets "quanta." In 1913 Niels Bohr, who had developed the most widely accepted picture of the atom at that time, saw that a description of the way these quanta behaved also fitted the behavior of the electron in the atom.

With this new picture of the universe came a number of paradoxes: the disappearance of space-time locality, the abandonment of causality, and the contradictory appearance of atomic matter as both particles and waves.

The record of the period just before the conference of 1927 is best given by physicist Werner Heisenberg who worked with Bohr on this problem:

"I remember discussions with Bohr which went through many hours till very late at night and ended almost in despair, and when at the end of the discussion I went alone for a walk in the neighboring park I repeated to myself again and again the question: "Can nature possibly be as absurd as it seemed to us in these atomic experiments?" (Heisenberg 42)

At another point Heisenberg said, "When you speak about the model, you mean something which can only be described by means of classical physics. As soon as you go away from classical physics, then, in a strict sense you don' t even know what a model could possibly mean because then the words haven't got any meaning any more. Now this was a dilemma.... Bohr tried to keep the picture while at the same time omitting classical mechanics. He tried to keep the words and the pictures without keeping the meanings of the words of the pictures. Both things are possible in such a situation because your words don't really tackle the things any more. You can't get hold of the things by means of your words, so what shall you do?... Bohr's escape would be into the philosophy of things." (qtd in Folse 111)

Heisenberg remembers, "Those paradoxes were so in the center of his mind that he just couldn' t imagine that anybody could find an answer to the paradoxes, even having the nicest mathematical scheme in the world.... The very strange situation was that now by coming nearer and nearer to the solution the paradoxes became worse and worse. That was the main experience.... nobody could know an answer to the question, 'Is an electron now a wave or is it a particle, and how does it behave if I do this or that and so on.' Therefore the paradoxes became so much more pronounced in that time.... only by coming nearer and nearer to the real thing to see that the paradoxes by no means disappeared but on the contrary got worse and worse because they turn out more clearly.... like a chemist who tries to concentrate his poison more and more from some kind of solution, we tried to concentrate the poison of the paradox" (qtd in Folse 85)

Heisenberg said, "Bohr was more worried than anybody about the inconsistencies of quantum theory. So he tried really to understand what is behind these difficulties.... Bohr really suffered from it, and Bohr couldn't talk of anything else....He in some ways directly suffered from this impossibility to penetrate into this very unanschaulich, unreasonable behavior of nature.... But that was Bohr's whole philosophical attitude -- he was a man who really always wanted to get the last degree of clarity. He would never stop before the end.... Bohr would follow the thing to the very end, just to the point where he was just at the wall.... He did see that the whole theory was on the one hand extremely successful, and on the other hand was fundamentally wrong. And that was a contradiction which was very difficult to bear, especially for a man who had formulated the theory. So he was in a continuous inner discussion about the problem. He always worried 'what has happened?'" (qtd. in Folse 36-37)

Bohr, Heisenberg and concepts of reality

During this early development of quantum theory there appeared a disagreement between Bohr and Heisenberg that is important to notice. Heisenberg was satisfied that the mathematical solution, matrix mechanics, gave all the understanding of atomic systems that was needed. Verbal pictures of what was going on were not necessary. Classical theoretical notions as "objects" are no more than conceptual instruments for predicting successfully the outcome of various experiments.

Heisenberg said, "Well, we have a consistent mathematical scheme and this consistent mathematical scheme tells us everything which can be observed. Nothing is in nature which cannot be described by this scheme.... Since classical physics is not true there, why should we stick so much to these concepts? Why not say just that we cannot use these concepts with a high degree of precision.... and therefore we have to abandon the classical concepts to a certain extent. When we get beyond this range of the classical theory we must realize that our words don't fit. They don't really get a hold in the physical reality and therefore a new mathematical scheme is just as good as anything because the new mathematical scheme then tells what may be there and what may not be there." (qtd in Folse 94)

This early view of Heisenberg's is, I understand, the view of most physicists today. If the mathematics works who needs the philosophy? But Bohr did not agree at all with this view.

Bohr saw that the quantum theory's mathematical formulation had to have a connection to the cultural world of everyday life in which the experiments are performed. If that connection were not made there would be no way to run an experiment that would prove whether a quantum prediction was true or not. Quantum theory must be verified by classical concepts that refer to observable properties of nature.

Heisenberg remembers, ".... Sometimes Bohr and I would disagree because I would say, 'Well I'm convinced that this is the solution already.' Bohr would say, 'No there you come into a contradiction.' Then sometimes I had the impression that Bohr really tried to lead me onto Glatteis, onto slippery ground, in order to prove that I had not the solution. But, this was, of course, exactly what he had to do from his point of view. It was perfectly correct. He was also perfectly correct in saying, 'So long as it is possible that you get onto slippery ground, then it means that we have not understood the theory.'" (qtd, in Folse 86-87)

Heisenberg said the controversy was so intense, "I remember that it ended with my breaking out into tears because I just couldn't stand this pressure from Bohr." (qtd in Jammer 65) But Heisenberg concluded, ".... just by these discussions with Bohr I learned that the thing which I in some way attempted could not be done. That is one cannot go entirely away from the old words because one has to talk about something .... So I could realize that I could not avoid using these weak terms which we always have used for many years in order to describe what I see. So I saw that in order to describe phenomena one needs a language.... The terms don't get hold of the phenomena, but still, to some extent they do. I realized, in the process of these discussions with Bohr, how desperate the situation is. On the one hand we knew that our concepts don't work, and on the other hand we have nothing except the concepts with which we could talk about what we see.... I think this tension you just have to take; you can't avoid it. That was perhaps the strongest experience of these months." (qtd in Folse 96)

The language problem

As I read these statements it occurred to me that the tension that Heisenberg referred to still exists to day and may be in part the reason for this conference here in Brussels this week. Although scientists have great problems in their work with the use of the every day language of literature and the arts, they cannot do without it.

When Bohr formulated his philosophy of Complementarity that was what he was trying to do -- find a common ground between the new quantum theory and the language of everyday life. It was this effort that Einstein attacked here in Brussels in October 1927. Bohr was really caught in the middle between anti-realists like Heisenberg who said, forget the philosophy and the realists like Einstein who said, if you stay with statistics without specifying what it means in terms of real external objects, then you are leaving reality behind.

The debate was always in terms of thought experiments. Although Bohr had said, "Reality is a term we must learn to use," the debate was never raised to the level of a discussion of what this "physical reality" is whose description is either complete or incomplete. The reason may be that in those days a philosophic discussion of "reality" was greatly discouraged. Discussions of reality were metaphysics and metaphysics was something associated with medieval religious mysticism. Yet as I read through the material even I could see that this was not primarily a quarrel about physics, it was about metaphysics. And I saw that others had noted that too 2. There is no way one can possibly construct a scientific experiment to determine whether or not an external reality exists if there is a difference in metaphysical interpretation. Whatever results you come up with can still be explained differently in each metaphysical system.

Complementarity

So now it is time to get into a closer look at the metaphysical system of Complementarity itself. As almost everyone comments, it is not easy to understand. I have been over the materials dozens of times and still am not at all sure I have it completely right. I want to show some simple diagrams first to make it clearer.

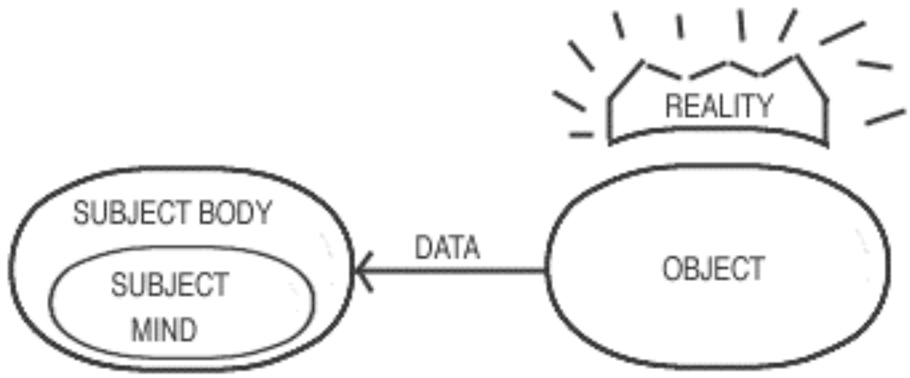


FIGURE 1

This first drawing represents the classical view of science. We are the subject. The external world is the object. We study the object with measuring instruments to collect data about the object, work with logic and math on this data and develop a theory to explain what this object really is. This view is so well known to us today we think of it as common sense. If there were time it would be valuable to get into the history of how this view came into being. In 400 of the last 500 years it has worked with enormous success. It is only in the last hundred years or so that our measurements are showing that the objects we are studying are apparently impossible. Since the phenomena from the measurements are not about to change, Bohr concluded that the logic of science must change to accommodate them.

Here is the second diagram

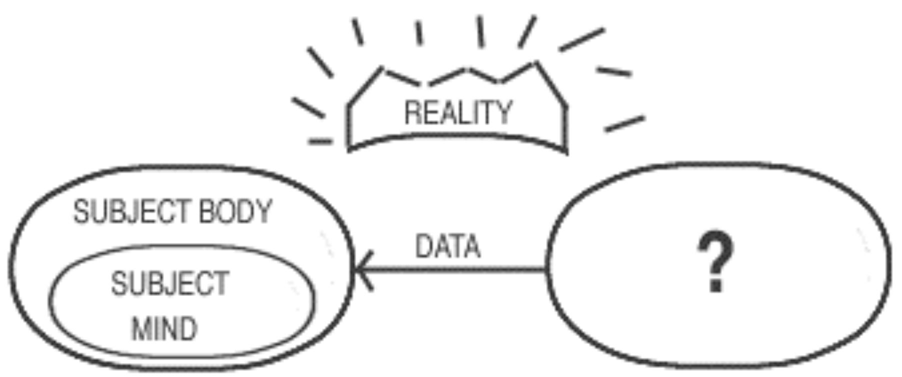


FIGURE 2

I found Complementarity easier to understand when I describe it in two steps, of which this is the first. There is a shift in reality shown here from the object to the data. This view known as phenomenalism, says that what we really observe is not the object. What we really observe is only data. This philosophy of science is associated with Ernst Mach and the positivists. Einstein did not like it and assumed Bohr shared it, but Bohr did not reject objectivity completely. He did not care so much which philosophical camp he was in, he was mainly concerned with whether Complementarity provided an adequate description to go with the quantum theory. In this third diagram we get down to the details of Complementarity.

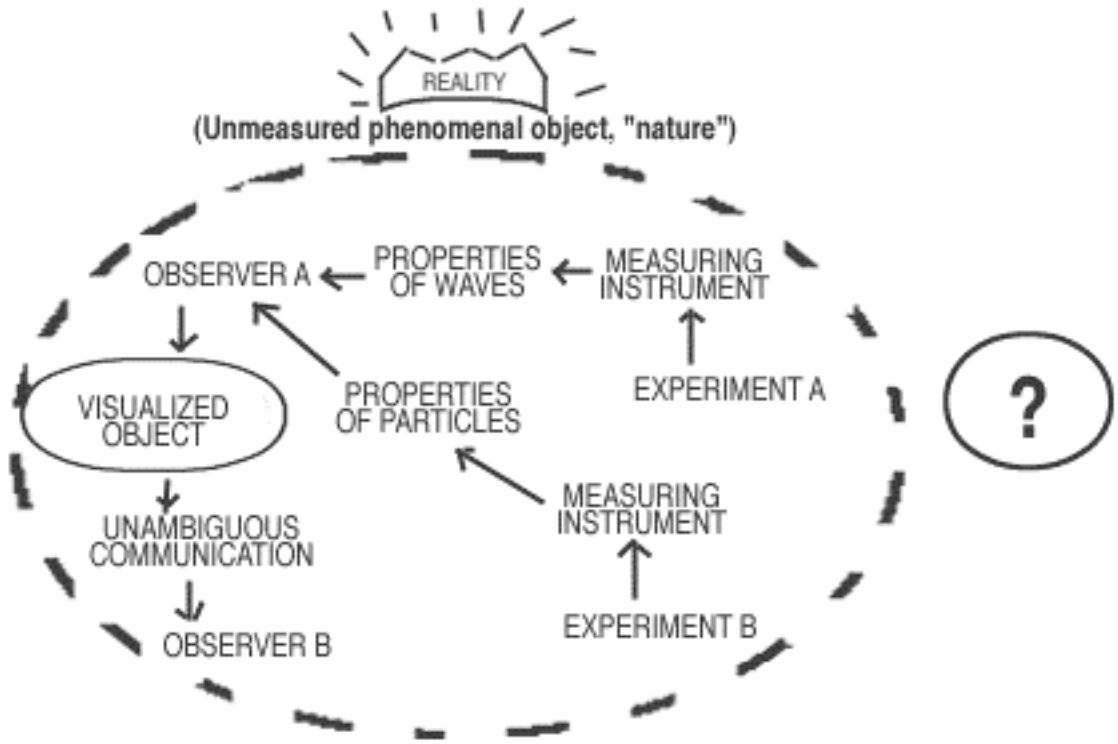


FIGURE 3

This diagram is not anything Bohr generated. It is something I have assembled myself and although I have revised it many times I would still expect Bohr to find things wrong with it, and others too who are more familiar with this subject than I am. But it is the best I can do. Bohr saw the Complementarity that is diagramed here a a way of solving many paradoxes but the wave-particle paradox he seems to have given the most attention to and I will use this paradox only.

First, notice that within this phenomenal object all things are together except the visualized object that is surrounded by an inner oval. There is no sharp exclusion of the observer from the observation. There is no sharp distinction of the measuring instrument from the experiment. The whole phenomenon is treated as one big observational interaction in which the distinction between observing system and observed phenomenal object is clear but is arbitrary.

Second, notice that on the right hand side of this larger oval there are two experiments: Experiment A and Experiment B. From Experiment A the observer observes waves. From Experiment B the observer observes particles. The experiments never put these two together. It is wrong to say that the experiments are on the same object or on any object at all. It is wrong to say that waves or particles are there before the experiment takes place. We can never say what goes into the experiment. We can only comment on what comes out.

Third, notice that when observer A observes experiment A and then, at another time, he observes experiment B, he may afterward in his mind combine the results of experiment A and experiment B to produce a "visualized" or "idealized" object. This visual object is a sort of mental collage created by the observer. Experiment A and Experiment B have been combined in a Complementarity for his philosophy.

Fourth, notice that this "visualized' object, that now may be called "light," is both waves and particles. Its description is what we must mean when we speak of objectivity. When Bohr says " It is wrong to think that the task of physics is to find out how nature is. Physics concerns what we can say about nature." (Herbert p45) He means that this visualized object is all we can talk about. It is an abstraction, but there is no other object. There is no "deep reality."

Fifth, notice that observer A then communicates this visualized object in an unambiguous way to observer B. By "unambiguous" is meant that A communicates it through a mathematical formalism combined with a word picture. All measuring equipment must be included in an unambiguous description. Later observer B can run his own experiment using the same measuring instruments and testing conditions to confirm the unambiguous communication from observer A. The proved unambiguity of this communication verifies the true objectivity of A's visualized object.

It can now be said that, because of this way of understanding things, a truly objective description has been given of light as both waves and particles without involving nature in a contradiction.

Finally, notice that this largest oval, the unmeasured phenomenal object shown with the dashed line, contains everything that Bohr talks about. He never discusses the old physical reality shown with the question mark off to the right that is external to this unmeasured phenomenal object. But, more importantly, he never mentions this larger oval, this unmeasured phenomenal object itself, presumably because to do so would be meaningless. It has no properties. The properties result only from the experiment that occurs within this oval. I have made this oval with a dashed line because I have a feeling Bohr wouldn't approve of it. But I think this larger unmeasured phenomenal object with the dashed line has to be there because if it were not there the only thing the experiments would be measuring is the measuring instruments themselves. Though Bohr doesn't describe it, something as to go into the front end of each experiment. I may be missing something but I don't see how you can have an experiment where nothing goes in but phenomena come out. Bohr may say that what goes in the front end of the experiment is "meaningless" and by the use of that term invite us to never think of it at all. But there has to be something going in whether it is meaningless or not. I make this point now because I will be coming to it later.

It has been said that neither Einstein nor Bohr seemed explicitly aware that although they conducted their dispute in terms of thought experiments, the dispute is nevertheless about metaphysics. The metaphysical issue at the root of it all is the old mind-versus matter issue, the subject-versus-object issue that has dogged philosophy since the days of Isaac Newton and David Hume and Immanuel Kant.

Subjectivity

Bohr's Complementarity was accused of being subjectivistic. If the world is composed of subjects and objects, and if Bohr says the properties of the atom are not in the objects, then Bohr is saying that the properties of the atom are in the subject. But if there is one thing science cannot be it is subjective. You cannot seriously say that science is all in your head. However in his early writing on Complementarity that is what Bohr seemed to be saying. (Folse 24) Bohr was trying to work out a problem in quantum physics, not just juggle a lot of philosophic categories, and Henry Folse says it didn't seem to occur to him what the implications of this might be. In his first paper on Complementarity Bohr made no mention of objectivity and actually made the gross mistake of calling his Complementarity subjective. He also spoke of scientific observation as "disturbing the phenomenon" which suggested that either he was talking about thoughts disturbing objects or else talking about phenomena being subjective.

Given this attack on his subjectivity it can be seen why Bohr developed the concepts of "phenomenal object" and "visual object" as independent of the subject in the diagram I have just shown you. He was constantly under pressure to prove that what he was talking about was not subjective.

His repeated argument is that Complementarity is not subjective because it provides unambiguous communication. When the results of the experiment exist unambiguously in the mind of several scientists Bohr says it is no longer subjective.

However, in my own opinion, that still doesn't get him out of the charge of subjectivity. When Bohr says the test of objective, scientific truth is "unambiguous communication" he is saying that it is not nature but society that ultimately decides what is true. But a society is not an objective entity. As anthropologists well know, societies are subjective too. The only truly objective aspects of "unambiguous communication" are the brain circuits that produce it; the larynx; the sound waves or other media that carry it; the ear drum, and the brain circuits that receive it. These can process falsehood just as easily as truth.

Folse says that Bohr never overcame the criticism that his philosophy was subjectivistic. "Bohr had envisioned Complementarity spreading out into wider and wider fields, just as the mechanical approach of Galileo had started in astronomy and simple phenomena of motion and gradually spread to all of the physical science." (Folse 168) But that never happened. Quantum physics dominates the scientific scene today but not because of Bohr's philosophy of Complementarity. It dominates because the mathematical formalisms of quantum theory correctly predict atomic phenomena.

Bohr was disappointed all his life by what he regarded as the failure of philosophers to understand Complementarity. Except for William James he "felt that philosophers were very odd people who really were lost." (Folse 44) Late in his life he remarked, "I think it would be reasonable to say that no man who is called a philosopher really understands what is meant by the Complementary descriptions." And as Folse concludes, " that somewhat wistful comment by this great pioneer of modern atomic theory is sadly true today as it was over fifty years ago." (Folse 44) Although Bohr had intended to write a book that contained and developed his philosophical ideas he never wrote it. This leads me to think that he realized his philosophy wasn't working the way he hoped it would but didn't know what to do about it. He talked as though he was sure it was right but was frustrated and disappointed that it never seemed to have caught on with others.

Henry Folse said that, "in what was to be his last interview, the day before his death, Bohr was questioned by Thomas Kuhn about the nature of his interest in fundamental philosophical problems. His answer was direct: 'It was in some ways my life you see.'" (Folse 31) That reply has an understatement and sadness to it that left me quiet for a long time.

The Metaphysics of Quality

I want to make a sharp shift now from Copenhagen to the town of Bozeman, Montana and the English department of Montana State College in 1959 when I was a teacher there. Sometimes people come at me when I talk about quality as though I had made the whole problem up by myself. But I was under legal contract with the state government of Montana to teach quality even though I had no clear idea what is was, and nobody else did either. Anthropologists know that every culture has strange and bizarre practices that make no sense from a practical view, but it is much easer to spot those practices in other cultures than in our own. I will point out to you that for centuries rhetoric instructors in our culture have been paid to pass and fail students on the quality of their writing without ever having any viable definition of what that quality is or even if there is such a thing at all. This is a bizarre practice that I tried to end.

In Zen and the Art of Motorcycle Maintenance I described how the question, "What is quality?" had been arrived at, and I described the first attempt to solve it where Phaedrus thinks to himself: "Quality ... you know what it is, yet you don't know what it is. But that's self-contradictory. But some things are better than others, that is, they have more quality. But when you try to say what the quality is, apart from the things that have it, it all goes pouf! There's nothing more to talk about. But if you can't say what Quality is, how do you know what it is, or how do you know that it even exists? If no one knows what it is, then for all practical purposes it doesn't exist at all. But for all practical purposes it really does exist. What else are the grades based on? Why else would people pay fortunes for some things and throw others in the trash pile? Obviously some things are better than others ... but what's the 'betterness'? ... So round and round you go spinning mental wheels and nowhere finding anyplace to get traction."

It was a common mischievous practice for students to send the same rhetoric paper to different teachers and observe that it got different grades. From this the students would argue that the whole idea of quality was meaningless. But one instructor turned the tables on them and handed a group of papers to several different students and asked each student to grade them for quality. As he expected, the student's relative rankings correlated with each other and with those of the instructor. This meant that although the students were saying there is no such thing as quality, they already knew what is was, and could not deny it.

So what I did is transfer that exercise into the classroom, having the students judge four papers day after day until they saw that they knew what quality is. They never had to say in any conceptual way what kind of object quality is but they understood that when you see it you know it. Quality is real even though it cannot be defined.

Eventually my unusual teaching methods came to the attention of the other professors in the department and in a friendly way they asked the question that connects all this with the struggles of Niels Bohr: "is quality in the subject or in the object?" The answer that was finally given was, "neither, Quality is a separate category of experience that is neither subject or object." This was the beginning of the system of thought called the Metaphysics of Quality. It has lasted for more than 35 years now. The question today is, if Niels Bohr had given that answer would his system of Complementarity have been improved?

In the Metaphysics of Quality the world is composed of three things: mind, matter and Quality. Because something is not located in the object does not mean that it has to be located in your mind. Quality cannot be independently derived form either mind or matter. But it can be derived from the relationship of mind and matter with each other. Quality occurs at the point at which subject and object meet. Quality is not a thing. It is an event. It is the event at which the subject becomes aware of the object. And because without objects there can be no subject, quality is the event at which awareness of both subjects and objects is made possible. Quality is not just the result of a collision between subject and object. The very existence of subject and object themselves is deduced from the Quality event. The Quality event is the cause of the subjects and objects, which are then mistakenly presumed to be the cause of the Quality!

The most striking similarity between the Metaphysics of Quality and Complementarity is that this Quality event corresponds to what Bohr means by "observation." When the Copenhagen Interpretation "holds that the unmeasured atom is not real, that its attributes are created or realized in the act of measurement," (Herbert xiii) it is saying something very close to the Metaphysics of Quality. The observation creates the reality.

Zen and the Art of Motorcycle Maintenance left one enormous metaphysical problem unanswered that became the central driving reason for the expansion of the Metaphysics of Quality into a second book called Lila. This problem was: if Quality is a constant, why does it seem so variable? Why do people have different opinions about it? The answer became: The quality that was referred to in Zen and the Art of Motorcycle Maintenance can be subdivided into Dynamic Quality and static quality. Dynamic Quality is a stream of quality events going on and on forever, always at the cutting edge of the present. But in the wake of this cutting edge are static patterns of value. These are memories, customs and patterns of nature. The reason there is a difference between individual evaluations of quality is that although Dynamic Quality is a constant, these static patterns are different for everyone because each person has a different static pattern of life history. Both the Dynamic Quality and the static patterns influence his final judgment. That is why there is some uniformity among individual value judgments but not complete uniformity.

Here is a drawing of the basic framework of the Metaphysics of Quality:

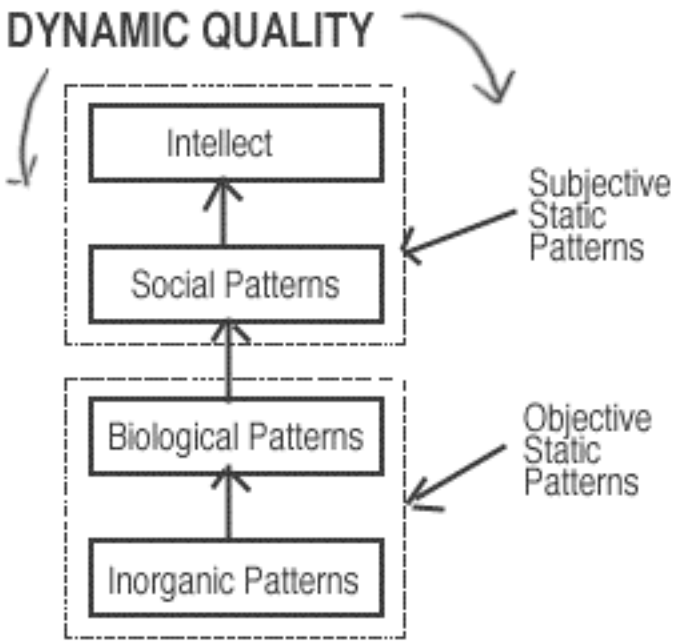


FIGURE 4

In this diagram you will notice that Dynamic Quality is not shown in any block. It is in the background. This seems the best way to represent it. It is not only outside the blocks, it pervades them but it goes on where the blocks leave off.

The blocks are organized in the order of evolution, with each higher block more recent and more Dynamic than the lower ones. The block at the top contains such static intellectual patterns as theology, science, philosophy, mathematics. The placement of the intellect in this position makes it superior to society, biology and inorganic patterns but still inferior to Dynamic Quality. The Metaphysics of Quality says there can be many competing truths and it is value that decides among them. This is the very essence of William James' philosophy of Pragmatism which Bohr greatly admired. The name "Complementarity" itself means there can be multiple truths.

The social patterns in the next box down include such institutions as family, church and government. They are the patterns of culture that the anthropologist and sociologist study.

In the third box are the biological patterns: senses of touch, sight, hearing, smell and taste. The Metaphysics of Quality follows the empirical tradition here in saying that the senses are the starting point of reality, but -- all importantly -- it includes a sense of value. Values are phenomena. To ignore them is to misread the world. It says this sense of value, of liking or disliking, is a primary sense that is a kind of gatekeeper for everything else an infant learns. At birth this sense of value is extremely Dynamic but as the infant grows up this sense of value becomes more and more influenced by accumulated static patterns. In the past this biological sense of value has been called the "subjective" because there values cannot be located in an external physical object. But quantum theory has destroyed the idea that only properties located in external physical objects have reality.

The bottom box shows inorganic patterns. The Metaphysics of Quality says objects are composed of "substance" but it says that this substance can be defined more precisely as "stable inorganic patterns of value." This added definition makes substance sound more ephemeral than previously but it is not. The objects look and smell and feel the same either way. The Metaphysics of Quality agrees with scientific realism that these inorganic patterns are completely real, and there is no reason that box shouldn't be there, but it says that this reality is ulimately a deduction made in the first months of an infant's life and supported by the culture in which the infant grows up. I have noticed that Einstein in his 1936 essay Physics and Reality also held this view. (Jammer 230) Bohr is sometimes mistakenly thought to say that this inorganic level does not exist. However both Folse and Max Jammer argue at length that this is not true. He does not deny this inorganic reality. He simply says that the properties the physicist describes cannot be said to reside at this level.

I can now say some general things about this diagram:

Values

First, each higher pattern grows out of the lower one so we tend to think of the higher pattern as the property of the lower one. However if you study the world you will observe that the higher patterns often oppose the lower ones. Biological values of life oppose physical values of gravitation and entropy. Social values of family and law and order oppose biological values of lust and greed. Intellectual values of truth and freedom of opinion often oppose social patterns of government. This opposition of levels of static patterns offers a good explanation of why science in the past has rejected what it has called "values." The "values" it has rejected are static social prejudices and static biological emotions. When social patterns such as religion are mixed in with the scientific method, and when biological emotions are mixed in with the scientific method these "values" are properly considered a source of corruption of the scientific method. Science, it is said should be "value free", and if these were the only kind of values the statement would be true.

However, the Metaphysics of Quality observe that these two kinds of values are lower on the evolutionary ladder than the intellectual pattern of science. Science rejects them to set free its own higher intellectual pattern. The Metaphysics of Quality calls this a correct moral judgment by science. However science never rejects the value of truth. It never rejects the value of experiment. It never rejects the value of mathematical precision. Most important, it never rejects Dynamic Quality. The greatest strength of the scientific method is that it always allows new experiences, new ideas and a new evaluation of what it learns.

Next, notice that the Metaphysics of Quality provides a larger framework in which to integrate subjectivity and objectivity. Subjectivity and objectivity are not separate universes that have no connection to each other. They are instead separate stages of a single evolutionary process called value. I can find no place where the words subjective and objective are used where they cannot be replaced by one of these four categories. When we get rid of the words "subjective" and "objective" completely often there is a great increase in the clarity of what is said. One person who I'm sure would agree with me on this would be Niels Bohr.

A third piece of evidence that reveals the similarity between the Metaphysics of Quality and Complementarity occurs when Bohr says, "We are suspended in language," the Metaphysics of Quality completely agrees. In the block diagram of the Metaphysics of Quality we see that each higher level of evolution rests on and is supported by the next lower level of evolution and cannot do without it. There is no intellect that can independently reach and make contact with inorganic patterns. It must go through both society and biology to reach them. In the past science has insisted on the necessity of biological proofs, that is, proofs in terms of sense data, and it has tried to discard social patterns as a source of scientific knowledge. When Bohr says we are suspended in language I think he means you cannot get rid of the social contexts either. That was his argument to Heisenberg. The Metaphysics of Quality supports it.

The fourth evidence of similarity is that the Metaphysics of Quality substitutes the word "value" for cause. It says that to say "A causes B" can be better said as "B values precondition A." This has seemed to me to be a better terminology for describing quantum phenomena. The term "cause" implies an absolute certainty that quantum theory says does not exist.

The fifth evidence of similarity is that probability itself may be expressed as value, so that "a static pattern of inorganic values," which is a definition the Metaphysics of Quality gives to "substance," is the same as "a pattern of probabilities," which is a definition quantum theory gives to substance. If the atomic world is composed of probability waves and if probability is equal to value then it follows logically that the atomic world is composed of value. The literature on probability is very large and I haven't read it but I have noted that Heisenberg has said that "the possibility or 'tendency' for an even to take place has a kind of reality -- a certain intermediate layer of reality, halfway between the massive reality of matter and the intellectual reality of the idea or the image ... it is formulated quantitatively as probability and subject to mathematically expressible laws of nature." (Qtd in Jammer 44) This intermediate reality Heisenberg talked about may correspond to value, but I'm not sure of that. Although probability may equal inorganic value it certainly doesn't equal any of the other value patterns. All of these patterns -- all of life -- seem to be in a war against it. In biology, conformity to inorganic probability is another name for death.

The Conceptually Unknown

The sixth piece of evidence is that the Metaphysics of Quality answers a problem that Bohr refused to answer. His refusal has weighed against him. Bohr "refused to comment on the relationship between Complementarity and the nature of physical reality." (Folse 223) "Bohr never makes clear in what sense we can have knowledge of the reality which causes our experiences." (Folse 241) He leaves it just hanging in limbo.

The question is why would Bohr do that? It is absurd to think that he forgot about it, that it just slipped his mind. He must have had a reason. The explanation, I think, is that Bohr is prohibited from speaking about any external physical reality ahead of the experiment. Before the experiment he must say there is nothing to know. In the old classic physics an external object was put into the front end of the experiment. It was subject to one or another forces and the results studied. Now that external object is gone. Whatever Bohr says about anything that goes into the front end of the experiment will be taken as a property of an independent physical reality. It is vital to Complementarity that there are no properties until after the observation.

So Bohr never mentions the unmeasured phenomenal object shown as the larger dashed oval in the diagram of Complementarity. But as was said before, something has to be there. If it were not there the measuring instruments would just be measuring their own internal characteristics. It is clear from what Bohr does say that the unmeasured phenomenal object is unpatterned. The patterns only emerge after an experiment. This unmeasured phenomenal object is not the subject of classical physics. So what is left to conclude? It seems to me that it is not a very large jump of the imagination to see that this unmeasured phenomenal object is in fact a third category, which is not subject and not object because it is independent of the two. When this assertion is made Complementarity is out from under its lifelong accusation of subjectivity. We no longer need to claim that we ourselves alter scientific reality when we look at it and know about it -- a claim that Einstein regarded as part of a "shaky game."

The similarity between Dynamic Quality and Bohr's unmeasured phenomenal object does not at first seem very great. It is only when one sees that the unmeasured phenomenal object is not really phenomenal and not really an object that the two draw closer together. The unmeasured phenomenal object is not really phenomenal because it has no characteristics before an observation takes place. It is not really an object because objects are over in that right oval with the question mark in it. Those objects are what are being rejected in the first place. So what is this unmeasured phenomenal object?

It seems to me that a keystone in a bridge between the Metaphysics of Quality and Complementarity may be established if what has been called the "unmeasured phenomenal object" is now called the "The Conceptually Unknown" and what is called "Dynamic Quality" is also called "The Conceptually Unknown." Then the two come together. I would guess that the Conceptually Unknown is an unacceptable category in physics because it is intellectually meaningless and physics is only concerned with what is intellectually meaningful. That also might be why Bohr never mentioned it. However I think that this avoidance of The Conceptually Unknown should be revised. It is like saying that the number zero is unacceptable to mathematics because there's nothing there. Mathematics has done very well with the number "zero" despite that fact. The Conceptually Unknown, it seems to me is a workable intellectual category for the description of nature and it ought to be worked more. As a starting axiom I would say, "Things which are intellectually meaningless can nevertheless have value." I don't know of an artist who would disagree with that. Certainly not Rene Magritte.

For those who would like more information about this "Conceptually Unknown" than I can give them today there is a valuable book called Zen in the Art of Archery by Eugen Herrigel from which I derived the title for my own first book. When the Zen Archer refers to an "it" that shoots the arrow he is referring to what I mean by Dynamic Quality. For those who prefer to stay more within the confines of Western analytical though there is a book by Prof. F.S.C Northrop of Yale University called The Meeting of East and West. It is the book that really started me on this philosophic quest that has now lasted 47 years.

Northrop's name for Dynamic Quality is "the undifferentiated aesthetic continuum." By "continuum" he means that it goes on and on forever. By "undifferentiated" he means that it is without conceptual distinctions. And by "aesthetic" he means that it has quality.

I think that science generally agrees that there is something that has to enter into experiments other than the measuring instruments, and I think science would agree that "Conceptually Unknown" is an acceptable name for it. What science might not agree on is that this Conceptually unknown is aesthetic. But if the Conceptually Unknown were not aesthetic why should the scientific community be so attracted to it? If you think about it you will see that science would lose all meaning without this attraction to the unknown. A good word for the attraction is "curiosity." Without this curiosity there would never have been any science. Try to imagine a scientist who has no curiosity whatsoever and estimate what his output will be.

This aesthetic nature of the Conceptually Unknown is a point of connection between the sciences and the arts. What relates science to the arts is that science explore the Conceptually Unknown in order to develop a theory that will cover measurable patterns emerging from the unknown. The arts explore the Conceptually Unknown in other ways to create patterns such as music, literature, painting, that reveal the Dynamic Quality that produced them. This description, I think, is the rational connection between science and the arts.

In Zen and the Art of Motorcycle Maintenance art was defined as high quality endeavor. I have never found a need to add anything to that definition. But one of the reasons I have spent so much time in this paper describing the personal relationship of Werner Heisenberg and Niels Bohr in the development of quantum theory is that although the world views science as a sort of plodding, logical methodical advancement of knowledge, what I saw here were two artists in the throes of creative discovery. They were at the cutting edge of knowledge plunging into the unknown trying to bring something out of that unknown into a static form that would be of value to everyone. As Bohr might have loved to observe, science and art are just two different complementary ways of looking at the same thing. In the largest sense it is really unnecessary to create a meeting of the arts and sciences because in actual practice, at the most immediate level they have never really been separated. They have always been different aspects of the same human purpose.

welding gives a tremendous feeling of power and control over the metal. You can do anything. He brings out some photographs of things he has welded and these show beautiful birds and animals with flowing metal surface textures that are not like anything else.

Later I move over and talk with Jack and Wylla. Jack is leaving to head an English department down in Boise, Idaho. His attitudes toward the department here seem guarded, but negative. They would be negative, of course, or he wouldn't be leaving. I seem to remember now he was a fiction writer mainly, who taught English, rather than a systematic scholar who taught English. There was a continuing split in the department along these lines which in part gave rise to, or at least accelerated the growth of, Phædrus' wild set of ideas which no one else had ever heard of, and Jack was supportive of Phædrus because, although he wasn't sure he knew what Phædrus was talking about, he saw it was something a fiction writer could work with better than linguistic analysis. It's an old split. Like the one between art and art history. One does it and the other talks about how it's done and the talk about how it's done never seems to match how one does it.

DeWeese brings over some instructions for assembly of an outdoor barbecue rotisserie which he wants me to evaluate as a professional technical writer. He's spent a whole afternoon trying to get the thing together and he wants to see these instructions totally damned.

But as I read them they look like ordinary instructions to me and I'm at a loss to find anything wrong with them. I don't want to say this, of course, so I hunt hard for something to pick on. You can't really tell whether a set of instructions is all right until you check it against the device or procedure it describes, but I see a page separation that prevents reading without flipping back and forth between the text and illustration...always a poor practice. I jump on this very hard and DeWeese encourages every jump. Chris takes the instructions to see what I mean.

But while I'm jumping on this and describing some of the agonies of misinterpretation that bad cross- referencing can produce, I've a feeling that this isn't why DeWeese found them so hard to understand. It's just the lack of smoothness and continuity which threw him off. He's unable to comprehend things when they appear in the ugly, chopped-up, grotesque sentence style common to engineering and technical writing. Science works with chunks and bits and pieces of things with the continuity presumed, and DeWeese works only with the continuities of things with the chunks and bits and pieces presumed. What he really wants me to damn is the lack of artistic continuity, something an engineer couldn't care less about. It hangs up, really, on the classic-romantic split, like everything else about technology.

But Chris, meanwhile, takes the instructions and folds them around in a way I hadn't thought of so that the illustration sits there right next to the text. I double-take this, then triple-take it and feel like a movie cartoon character who has just walked beyond the edge of a cliff but hasn't fallen yet because he hasn't realized his predicament. I nod, and there's silence, and then I realize my predicament, then a long laughter as I pound Chris on the top of the head all the way down to the bottom of the canyon. When the laughter subsides, I say, ``Well, anyway -- " but the laughter starts all over again.

``What I wanted to say," I finally get in, ``is that I've a set of instructions at home which open up great realms for the improvement of technical writing. They begin, `Assembly of Japanese bicycle require great peace of mind.' "

This produces more laughter, but Sylvia and Gennie and the sculptor give sharp looks of recognition.

``That's a good instruction," the sculptor says. Gennie nods too.

``That's kind of why I saved it," I say. ``At first I laughed because of memories of bicycles I'd put together and, of course, the unintended slur on Japanese manufacture. But there's a lot of wisdom in that statement."

John looks at me apprehensively. I look at him with equal apprehension. We both laugh. He says, ``The professor will now expound."

``Peace of mind isn't at all superficial, really," I expound. ``It's the whole thing. That which produces it is good maintenance; that which disturbs it is poor maintenance. What we call workability of the machine is just an objectification of this peace of mind. The ultimate test's always your own serenity. If you don't have this when you start and maintain it while you're working you're likely to build your personal problems right into the machine itself."

They just look at me, thinking about this.

``It's an unconventional concept," I say, ``but conventional reason bears it out. The material object of observation, the bicycle or rotisserie, can't be right or wrong. Molecules are molecules. They don't have any ethical codes to follow except those people give them. The test of the machine is the satisfaction it gives you. There isn't any other test. If the machine produces tranquillity it's right. If it disturbs you it's wrong until either the machine or your mind is changed. The test of the machine's always your own mind. There isn't any other test."

DeWeese asks, ``What if the machine is wrong and I feel peaceful about it?"

Laughter.

I reply, ``That's self-contradictory. If you really don't care you aren't going to know it's wrong. The thought'll never occur to you. The act of pronouncing it wrong's a form of caring."

I add, ``What's more common is that you feel unpeaceful even if it's right, and I think that's the actual case here. In this case, if you're worried, it isn't right. That means it isn't checked out thoroughly enough. In any industrial situation a machine that isn't checked out is a `down' machine and can't be used even though it may work perfectly. Your worry about the rotisserie is the same thing. You haven't completed the ultimate requirement of achieving peace of mind, because you feel these instructions were too complicated and you may not have understood them correctly."

DeWeese asks, ``Well, how would you change them so I would get this peace of mind?"

``That would require a lot more study than I've just given them now. The whole thing goes very deep. These rotisserie instructions begin and end exclusively with the machine. But the kind of approach I'm thinking about doesn't cut it off so narrowly. What's really angering about instructions of this sort is that they imply there's only one way to put this rotisserie together...their way. And that presumption wipes out all the creativity. Actually there are hundreds of ways to put the rotisserie together and when they make you follow just one way without showing you the overall problem the instructions become hard to follow in such a way as not to make mistakes. You lose feeling for the work. And not only that, it's very unlikely that they've told you the best way."

``But they're from the factory," John says.

``I'm from the factory too," I say ``and I know how instructions like this are put together. You go out on the assembly line with a tape recorder and the foreman sends you to talk to the guy he needs least, the biggest goof-off he's got, and whatever he tells you...that's the instructions. The next guy might have told you something completely different and probably better, but he's too busy." They all look surprised. ``I might have known," DeWeese says.

``It's the format," I say. ``No writer can buck it. Technology presumes there's just one right way to do things and there never is. And when you presume there's just one right way to do things, of course the instructions begin and end exclusively with the rotisserie. But if you have to choose among an infinite number of ways to put it together then the relation of the machine to you, and the relation of the machine and you to the rest of the world, has to be considered, because the selection from many choices, the art of the work is just as dependent upon your own mind and spirit as it is upon the material of the machine. That's why you need the peace of mind."

``Actually this idea isn't so strange," I continue. ``Sometime look at a novice workman or a bad workman and compare his expression with that of a craftsman whose work you know is excellent and you'll see the difference. The craftsman isn't ever following a single line of instruction. He's making decisions as he goes along. For that reason he'll be absorbed and attentive to what he's doing even though he doesn't deliberately contrive this. His motions and the machine are in a kind of harmony. He isn't following any set of written instructions because the nature of the material at hand determines his thoughts and motions, which simultaneously change the nature of the material at hand. The material and his thoughts are changing together in a progression of changes until his mind's at rest at the same time the material's right."

``Sounds like art," the instructor says.

``Well, it is art," I say. ``This divorce of art from technology is completely unnatural. It's just that it's gone on so long you have to be an archeologist to find out where the two separated. Rotisserie assembly is actually a long-lost branch of sculpture, so divorced from its roots by centuries of intellectual wrong turns that just to associate the two sounds ludicrous."

They're not sure whether I'm kidding or not.

``You mean," DeWeese asks, ``that when I was putting this rotisserie together I was actually sculpting it?"

``Sure."

He goes over this in his mind, smiling more and more. ``I wish I'd known that," he says. Laughter follows.

Chris says he doesn't understand what I'm saying. ``That's all right, Chris," Jack Barsness says. ``We don't either." More laughter.

``I think I'll just stay with ordinary sculpture," the sculptor says.

``I think I'll just stick to painting," DeWeese says.

``I think I'll just stick to drumming," John says.

Chris asks, ``What are you going to stick to?"

``Mah guns, boy, mah guns," I tell him. ``That's the Code of the West."