Philosophy of science underpinnings of prototype validation: Popper vs. Quine

Running headline: Philosophy of science -- prototype validation

Esther E. Klein, Ph. D.*
Department of Business Computer Information Systems/Quantitative Methods
Zarb School of Business
Hofstra University
211 Weller Hall
Hempstead, NY 11549, USA
(516) 463 – 4529
eklein9@aol.com OR acseek@hofstra.edu

Paul J. Herskovitz, M.B.A., J.D.
Department of Business (Law)
College of Staten Island, City University of New York
2800 Victory Boulevard, Room #3N – 206
Staten Island, NY 10314, USA
(718) 982 – 2959
pjkersko@aol.com OR herskovitz@mail.csi.cuny.edu

17,701 words

* Corresponding author.
Abstract. In this paper, we aim to provide prototype validation with a theoretical framework located within the philosophy of science. Toward that end, we consider Popperian and Quinean accounts of scientific knowledge and argue that the theoretical underpinning of prototype validation is Quine’s holistic philosophy of science, whose cornerstone principle is that all beliefs are revisable. Specifically, our thesis is that the systems developer and prototype end user join forces — not as Popperian falsifiers, who make a decision rule to reject the prototype on account of any divergence, major or minor, from the user’s mental model, but — as Quinean revisers, with the objective of fine-tuning the prototype (or the user’s mental model) so that the prototype and the user’s mental model are congruent with each other. This paper suggests that prototype revisions are belief revisions, and, as such, should be guided — and are guided — by pragmatic norms, such as conservatism, simplicity, and generality, and are influenced by social, or sociological, factors as well.

Keywords: Belief revision, coherence, epistemology, falsificationism, holism, pragmatic norms, philosophy of science, prototype validation

INTRODUCTION

The past decade, bracketed by the waning years of the second millennium and the opening years of the third, has been witness to a growing, robust, and richly-textured literature on the philosophical foundations and dimensions of information systems (IS) in general and information systems development (ISD) in particular (e.g., see Hirschheim, Klein & Lyytinen, 1995; Iivari & Hirschheim, 1996; Winder, Probert & Beeson, 1997; Lee, 1999; Bharadwaj, 2000; Milton, Kazmierczak & Thomas, 2000; Dobson, 2001; Gregg, Kulkarni & Vinzé, 2001; Probert, 2001; Rose, 2002; Probert, 2003; King & Kimble, 2004; Klein, 2004; Monod, 2004). The aim of this paper is to contribute to these efforts by providing prototype validation with a theoretical framework located within the philosophy of science.
IS is one of three computer, or information technology (IT), disciplines, the others being computer science (CS) and software engineering (SE). Glass, Ramesh, and Vessey have explained the difference between IS and its two sister disciplines thus:

Researchers in CS, and to some extent, SE, primarily expect to produce new things — processes, methods, algorithms, products. IS researchers, on the other hand, expect to explore things — theories, concepts, techniques, projects. … The explorations of IS are usually performed in an organizational and therefore behavioral context. … IS researchers emphasize their work is theory-based, and, as well as using theories based in IS, researchers in this realm also explore the relevance of theories extracted from other disciplines.

(Glass, Ramesh & Vessey, 2004, p. 93)

In contrast to CS and SE, which are predominantly technical disciplines, IS “in essence is an applied social science pertaining to the use and impact of technology” (Elliot & Avison, 2004, p. 5). IS, then, can be viewed as “social systems that are technically implemented” (Hirschheim, Klein & Lyytinen, 1995, p.1). Accordingly, Hevner and March (2003) have argued that the IS discipline is governed by both behavioral (social) science and design (technical) science paradigms (see also Hevner, March, Park & Ram, 2004).

Prototyping is a flexible “[ISD] methodology based on building and using a model of a system for designing, implementing, testing, and installing the system” (Lantz, 1987, p. 1). Under such an approach, “[information] systems are developed through an iterative rather than a systematic process,” whereby “[systems] developers and users are constantly interacting, revising, and testing the prototype system until it evolves into an acceptable application” (Oz, 2004, p. 599). (The terms “user” and “users,” as employed in this paper, refer to the end users of the information systems and encompass situations where the end user is an individual or a team.) Thus, prototyping fosters “constructive dialogue between users and designers” (Molin, 2004, p. 425). Hirschheim, Klein, and Lyytinen (1995, p. 35) have defined a prototype as “an experimental version of a system which is used to improve communication and user feedback (and sometimes to demonstrate technical feasibility or efficiency of new software designs)” (see also Michener, Mohan, Astrachan & Hale, 2003, p. 227). Put another way, a prototype “is a scaled down variant of the final [information] system which exhibits some of its salient features and thereby allows the users hands-on experimentation to understand the interfaces or computational power,” permitting the revision of the prototype in a
subsequent iteration (Hirschheim et al., 1995, p. 35; see also Bahn & Naumann, 1997, p. 240). Avison and Fitzgerald (1999, p. 260) have captured the essence of a prototype by defining it as “an approximation of a type that exhibits the essential features of the final version of that type” (see also Avison & Fitzgerald, 1995, p. 76).

Following Floyd (1984, pp. 6-12), Bischofberger and Pomberger (1992) have divided prototyping approaches into three categories on the basis of their objectives: exploratory (“to obtain a requirements definition,” p. 16), experimental (to attain “a concise specification of the components which form the system,” p. 17), and evolutionary (for incremental system development, p. 17). The commonality that underlies these three prototyping approaches is the revisability of the prototype and the assumption that there will indeed be revisions to the prototype. For Naumann and Jenkins (1982, p. 37), the activity of prototyping typically involves prototype revision: “The prototype builder constructs successive versions of the system, compromising and resolving conflicts between the context (i.e., user needs and desires) and the form as constrained by technology and economics.” Beynon-Davies, Tudhope, and Mackay (1999), citing Dearnley and Mayhew (1983), have similarly depicted prototyping in terms of prototype revision:

[T]he information system developer, after some initial investigation, constructs a working model which s/he demonstrates to the user. The developer and the user then discuss the prototype, agreeing on enhancements and amendments. This cycle of inspection-discussion-amendment is repeated several times until the user is satisfied with the system.

(Beynon-Davies, Tudhope & Mackay, 1999, p. 108)

According to Bischofberger and Pomberger (1992), the prototypes themselves can be categorized on the basis of the degree of incorporation in the target information system: complete, incomplete, throwaway, and reusable. These four kinds of prototypes are built with the expectation that they will be tested against the user’s mental model, and if necessary — which is often the case — revised.

As is true with many disciplines that have a pragmatic component, theory in IS has not kept pace with practice. An aim of this paper is to redress this imbalance with respect to prototyping. Specifically, in our quest to provide a theory for prototype validation, we will consider two seminal and well-developed philosophies of science as
theoretical frameworks — Popper’s falsificationism and Quine’s holism — and we will argue that it is Quinean holism that is the appropriate theory for prototype validation.

Philosophy of science, broadly construed, is concerned with “understand[ing] the meaning, method, and logical structure of science” (Klemke, Hollinger & Kline, 1980, p. 2), including the social, or human, sciences, as well as the natural sciences, and “intersects with other areas of philosophy, such as epistemology, metaphysics, and philosophy of language” (Curd & Cover, 1998, p. xvii). According to Okasha (2002, p. 12), “[t]he principal task of philosophy of science is to analyse the methods of enquiry used in the various sciences.” The focus of this paper is the epistemology of science, the branch of philosophy of science that deals with issues pertaining to how claims to scientific knowledge are justified and what constitutes evidential support for scientific statements, hypotheses, and theories. (In this paper, we follow the traditional definition of knowledge as “‘true belief’ that is ‘rationally justified’ or ‘well-grounded,’” Campbell, 1987, p. 53.) Evidence, its assessment, and “[t]he untidy process of groping for truth” (Haack, 1999, p. 12), then, are the concerns of epistemology of science, concerns that, not surprisingly, are shared by the discipline of law. In fact, there is a burgeoning cross-fertilization, with the attendant cross-referencing and trespassing over each others’ disciplinary boundaries, that is characteristic of both the epistemology of science and law literatures (e.g., see Uebel, 1993; Patterson, 1996; Goldman, 1997; Burney, 2002; Allen & Pardo, 2003; Beecher-Monas, 2003; Caudill & LaRue, 2003; Edmond, 2003; Haack, 2003; Imwinkelried, 2003; Mueller, 2003; Benedict, 2004; Ulen, 2004).

Contemporary philosophy of science has two prevailing — and competing — schools of thought: scientific realism and instrumentalism (Dogan, 2002). Scientific realism posits that scientific theories with long records of predictive success are at least approximately true and thus science offers, for the most part, a true account of reality (see Putnam, 1978, p. 18; McMullin, 1984). Instrumentalism holds that scientific theories are merely useful instruments or tools in making predictions and thus do not necessarily represent an underlying reality (see Rosenberg, 1994). Although not dominant, the philosophies of science of Popper and Quine have been exceptionally influential — Popper among scientists (Boyd, 1991, p. 11; Short, 1997, p. 85) and social scientists (Skinner, 1985, p. 5) and Quine among philosophers (Hookway, 1988, p. 1; Creath,
1990b, p. 1) — and, we suggest, may be useful in understanding discrete scientific activities within a larger discipline. According to Webb (1995, p. 87), in the social sciences, “we use theories to highlight aspects of reality that are deemed important with respect to a particular phenomenon” (see also Mitchell, 2002, p. 130). In that spirit, recently, Klein and Herskovitz (in press) have analyzed Popperian falsificationism, Quinean holism, and early-Putnamean scientific realism with respect to computer simulation validation, and have concluded that a Popperian perspective is the appropriate stance there. Similarly, here, we will consider Popperian falsificationism and Quinean holism not as overarching or grand theories for IS but merely as frameworks limited to exploring one aspect of IS: prototype validation.

Because an entire information system is implemented on the basis of a prototype, it is essential that the user make a determination whether the prototype is an accurate representation of the user’s mental model. Such a determination is referred to as prototype validation, which consists of a comparison of the prototype with the user’s mental model (e.g., see Dussart, Aubert & Patry, 2004; see also Baskerville, 1999). By building a prototype, the systems developer (prototype builder) is proposing a mini-theory of the user’s mental model. If the user finds that the prototype does not correspond in some way to the user’s mental model and thus cannot be validated, the prototype need not be rejected outright. Instead, either the prototype or the user’s mental model will need to be revised, or adjusted. Accordingly, we argue that the activity of prototype validation does not adhere to Popper’s falsificationist philosophy of science but rather follows a Quinean approach to belief revision, which is a central plank of Quine’s philosophy of science.

**POPPER’S PHILOSOPHY OF SCIENCE**

The two core ideas in Popper’s philosophy of science are falsifiability, or refutability, as the touchstone of science and the rejection of induction as the method of science (Popper, 1959; 1965; 1979). For Popper (1987, p. 116), “all human knowledge is fallible and conjectural … a product of the method of trial and error.” In Popper’s account of science, scientists put forth bold conjectures for trial and then systematically attempt to falsify these conjectures. Those conjectures that are not falsified are retained
“for the time being” (Popper, 1959, p. 104). According to Popper’s falsifiability criterion, for a proposition — statement, hypothesis, or theory — to qualify as scientific, it must put itself at risk by specifying certain predictions and forbidding certain observations:

Statements, or systems of statements, convey information about the empirical world only if they are capable of clashing with experience; or, more precisely, only if they can be \textit{systematically tested} [italics in original], that is to say, if they can be subjected … to tests which \textit{might} [italics in original] result in their refutation.

(Popper, 1959, pp. 313-314)

Accordingly, each scientific proposition must state under what conditions it will be deemed as having been disconfirmed for there cannot be a “genuine test [italics in original] of a theory [or other scientific proposition]” without “an attempt to falsify it” (Popper, 1957, p. 160). Under Popper’s normative, or prescriptive, epistemology of science, the task of the Popperian scientist is to earnestly endeavor to bring about the disconfirmation — not the verification — of scientific assertions. Popper has not advanced the claim that scientists in the real world necessarily proceed by attempting to falsify their assertions. What Popper has suggested is that making attempts at falsification \textit{should} be the way science is practiced.

According to Popper (1959), a scientific proposition can expose itself to the risk of being refuted by undergoing “severe tests.” If an experimental finding or observation is inconsistent with a proposition’s predictions, the proposition is conclusively falsified. If, however, the proposition withstands attempts at falsification, the proposition is not deemed verified (proven true) or probable, but rather merely “corroborated,” which in the Popperian lexicon means not yet falsified, only tentatively true (Popper, 1959, p. 266). Under a Popperian philosophy of science, the greater “the severity of the various tests” that the proposition has successfully endured, the higher the degree of corroboration (p. 267).

On a Popperian view, although a scientific proposition can be refuted upon the occurrence of a single, isolated counterinstance, such proposition can never be established as true or even probable, irrespective of the number of confirming instances. Under Popper’s falsificationist perspective, even well-tested theories are merely provisionally true, always subject to being rejected upon an experimental result or observation contrary to its predictions. “This view imply[s] that \textit{scientific theories are}
either falsified or for ever remain hypotheses or conjectures [italics in original]” (Popper, 1974, p. 62). Hence, there is a logical asymmetry between falsification and verification, whereby falsifications are regarded as unambiguous and decisive but verifications are viewed as conjectural and tentative.

**Falsificationism vs verificationism**

Popper’s approach is contrary to the traditional induction-based verificationist view of science, known as the “Received View” (Suppe, 1977), espoused by the logical, or empirical, positivists, that a theory can be proven true, or at least probable, by cumulative observations consistent with the theory’s predictions (see Carnap, 1966, p. 20; see also Ayer, 1955, p. 9; Agassi, 1988, p. 77). Thus, for verificationists, the method of science is induction, a mode of reasoning that proceeds from specific observations to general theories. According to the verificationist position, for a statement to be meaningful, it must be empirically verifiable in principle, that is, it must be possible to describe the type of test or “the sort of observation … which would confirm or disconfirm [the statement]” (Edwards & Pap, 1965, p. 677). For the verificationists, then, what distinguishes scientific assertions from nonscientific (metaphysical or pseudoscientific) ones is this verifiability criterion of meaningfulness.

The verificationist conception of science is a foundationalist position, holding that the use of empirical verifiability will ensure that our beliefs and knowledge are justified in that they rest on secure foundations. Popper has asserted that this verificationist quest for justification should be abandoned (Popper, 1979, p. 29). In contrast to the verificationist view, Popper’s philosophy of science is nonfoundationalist, advancing the notion that certainty in science is an elusive ideal and hence all knowledge is necessarily tentative:

We cannot reasonably aim at certainty. Once we realize that human knowledge is fallible, we realize also that we can never be completely certain [italics in original] that we have not made a mistake …. Since we can never know anything for sure, it is simply not worth searching for certainty; but it is well worth searching for truth; and we do this chiefly by searching for mistakes, so that we can correct them.

(Popper, 1992, p. 4)

**The problem of induction**
Popper has attacked verificationism’s use of induction, whereby knowledge develops by generalizations derived from observations. According to Popper, employing induction to establish the truth of a theory is not rationally justified because of the logical difficulty eponymously known as “Hume’s problem of induction” after its discoverer the philosopher David Hume, “who argued that from the strict logical point of view we have no justification in generalizing from instances we have experience of to those of which we have no experience” (O’Hear, 1989, p. 27). Born (1949) has articulated the problem thus:

No observation or experiment, however extended can give more than a finite number of repetitions; [therefore,] the statement of law — B depends on A — always transcends experience. Yet this kind of statement is made everywhere and all the time, and sometimes from scanty material.

(Born, 1949, p. 6)

Gillott and Kumar (1997) have offered this succinct illustration of the problem of induction:

How can we say all swans are white just because we have not seen any black ones? It could be the case that the next swan seen is black. No amount of observing white swans allows any inferences to be made about the probability of the next swan being white.

(Gillott & Kumar, 1997, p. 16)

Popper’s (1992) solution to the problem of induction is that induction does not exist because it is impossible to have observations that have not been influenced, or tainted, by theory. Observations — the “bedrock” of the traditional empiricist approach of verificationism (Laudan, 1990, p.35) — are never theory-free, but rather are embedded within “a frame of expectations” or “a frame of theories” (Popper, 1965, p. 47). Thus, theories, or conceptual pigeonholes, necessarily precede and taint observations so that all observations and observation reports are “interpretations in light of theories” (p. 38, note 3; see also Hanson, 1972, pp. 4-30; Kuhn, 1970c). Knowledge develops not by inductive inferences but by a process of trial and error, of learning from our mistakes.

The problem of demarcation

Popper has put forward the falsifiability criterion as not only the solution to the problem of induction but also the solution to the problem of demarcation: “[H]ow can you distinguish the theories of the empirical sciences from pseudo-scientific or non-
scientific or metaphysical speculations?” (Popper, 1983, p. 159). Simply put, “When should a theory be ranked as scientific?” (Popper, 1957, p. 155). According to the verificationist view a scientific theory is demarcated, or distinguished, from a nonscientific speculation by the former’s use of induction. For Popper, it is the falsifiability of a theory or other assertion that confers scientific status: “A theory which is not refutable by any conceivable event is non-scientific. Irrefutability is not a virtue of a theory (as people often think) but a vice” (Popper, 1957, p. 159).

According to Popper’s falsifiability criterion, Einstein’s theory of relativity is ranked as scientific because it puts itself at risk by making specific predictions concerning the distance of stars from the sun, thereby prohibiting any observation other than the one predicted:

If observation shows that the predicted effect is definitely absent, then the theory is simply refuted. The theory is *incompatible with certain possible results of observation* [italics in original] — in fact, with results which, before Einstein, everybody would have expected.

(Popper, 1957, p. 159)

Popper has contrasted Einstein’s theory of relativity with the psychological theories of Freud and Adler, which do not identify any type of human behavior that would refute their theories and thus are not regarded by Popper as scientific:

I could not think of any conceivable instance of human behaviour which could not be interpreted in terms of either [Freudian or Adlerian] theory. It was precisely this fact — that they always fitted, that they were always confirmed — which, in the eyes of their admirers, constituted the strongest arguments in favour of these theories. It began to dawn on me that this apparent strength was in fact their greatest weakness.

(Popper, 1957, p. 158)

Popper (1957, p. 160) has recognized that, consistent with Duhem’s and Quine’s holistic insights (see discussion below), “[s]ome genuinely testable [falsifiable] theories, when found to be false, are still upheld by their admirers” by the introduction of auxiliary hypotheses, or background assumptions, that save the theories from refutation. For example, scientists may explain away a contrary experimental result by positing that the experimental apparatus was faulty. According to Popper, “[s]uch a procedure is always possible, but it rescues the theory from refutation only at the price of destroying or at least lowering its scientific status” (Popper, 1957, p. 160) and thus should be generally avoided (Popper, 1959, pp. 82-83). For Popper, the scientists must adopt a decision rule
to conclusively reject theories that have been falsified without resort to post-experiment or post-observation introduction of auxiliary hypotheses (Popper, 1959, p. 37).

In *Daubert v. Merrell Dow Pharmaceuticals, Inc.* (1993), the United States Supreme Court has adopted Popper’s notion of falsifiability as a demarcation criterion with respect to the admissibility of scientific evidence. Subsequent court cases have reaffirmed this view (*General Electric Co. v. Joiner*, 1997; *Kumho Tire Co. v. Carmichael*, 1999; see also *US v. Starzecpyzel*, 1995; *City of Tuscaloosa v. Harcros Chems., Inc.*, 1998; *US v. Havward*, 2000; *Moore v. State*, 2001), and fresh scholarship on Popperian philosophy of science has penetrated the realm of law reviews (e.g., see Edmond & Mercer, 2002; Adelman, 2004; Case, 2004).

Boyd (1991, p. 11) has characterized Popperian falsificationism as a “variant version of verificationism which, although it has had rather little impact on recent philosophy of science, has had a deep influence on the thinking of many philosophically inclined scientists and other thinkers.” Writing from the perspective of a practicing scientist, Medawar was of the view that Popper’s falsificationist philosophy of science accurately reflected the work of scientists in real-life laboratories:

> In real laboratories there is no constant clamour of affirmation or denial. We are all very conscious of being engaged in an exploratory process as we cautiously grope our way forwards by the method which has come to be summed up by the now familiar cliché of *conjecture and refutation* [italics in original].
> (Medawar, 1959, p. 100)

**Criticisms of Popper’s philosophy of science**

The strand that unites most of the criticism of Popper’s philosophy of science is that it does not conform to reality, and more importantly, “that it is virtually impossible to put into practice” (Sayer, 1984, p. 205). Popper’s falsificationist philosophy has been attacked by Quine (1951; 1975), Kuhn (1970a; 1970b; 1970c), and Putnam (1974) as being contrary to the way science is actually practiced. Kuhn (1970c) has advanced the view, adopted by Putnam (1974), that:

> the sciences do not and cannot emulate a Popperian account of their practice [because] our access to the facts in the light of which we test our beliefs is always filtered by what Kuhn has called our existing “paradigms” or frameworks of understanding.
> (Skinner, 1985, p. 10)
Thus, in the real world, a theory is usually not abandoned because of a sole experimental result or observation contrary to the theory that is being tested. The typical scenario is that the scientist saves the theory from refutation by attributing the recalcitrant finding to a faulty implicit auxiliary hypothesis, or background assumption (e.g., the instrumentation is in good working order), a belief that now has to be revised. Such rescue of a theory under investigation is the central insight of Quine’s holist, or Duhem-Quine, thesis (see discussion below), which renders conclusive falsifications impossible and thereby poses a significant challenge to Popper’s philosophy of science. (For a recent illustration of how a single contrary experimental finding does not overthrow an established theory, see Myers et al., 2004; see also Things fall apart, 2004.)

Lewthwaite (2003) has pointed out that scientists do not go about their labors seeking the falsification of their theories:

It is unlikely that a scientist or theorist would accept the denouncement of an entire scientific theory simply on the grounds that an error has been identified in one of the many premises in an experiment test of that theory. Indeed, no legitimate ‘knowledge hunter’ would be so hasty in dismissing a life’s amount of work.

(Lewthwaite, 2003, ¶ 17)

Moreover, Papineau (1995) has suggested that were scientists to be solely concerned with refuting their theories that would undermine the goals of the scientific enterprise:

[T]here is an obvious flaw [in Popper’s anti-inductivist philosophy of science]. Popper’s falsificationist strategy of conjectures and refutations can only deliver negative knowledge. It can show that certain scientific theories are false, but it never shows that any theory is true. …There would be no point to science unless its conjectures sometimes acquired enough inductive evidence to graduate to the status of established truths. This is the real reason for testing hypotheses against predictions. The aim is not to falsify them, but to identify those that can be turned into the kind of positive knowledge that enables us to build bridges and treat diseases.

(Papineau, 1995, pp. 4-5)

Straub, Gefen, and Boudreau (2004) have observed that many contemporary methodologists have recognized that inquiry in the human sciences does not and should not strictly adhere to a Popperian philosophy of science, and thus social scientists should not be quick to reject a theory because of a sole contrary result:

It is also critical to note that more recent methodologists like Campbell and Stanley (1963) and Cook and Campbell (1979) are not as enamored as Popper with respect to the need for a theory to be falsified. Cook and Campbell (1979), in particular, go to great lengths to argue that the social sciences will almost
never be able to prove deterministically that a cause leads to an effect, as can sometimes be shown in the natural sciences. Statistical relationships can be instructive, however, and this means that one must be more cautious about asserting that a theory has been disconfirmed by a single study. In short, later methodologists are more willing to stress the extent to which a theory has been confirmed, given a cumulative tradition of work and the fact that these researchers found statistically significant relationships between certain causes and certain effects.

(Straub et al., 2004, p. 4)

A literal interpretation of Popper’s falsificationist criterion, then, will mandate the abandonment of a theory upon a single counterinstance, making it virtually impossible for the human sciences to retain any theories and ultimately resulting in the disappearance of the human sciences as disciplines. Webb (1995) has observed:

It is … the case that if social scientists followed the advice of Karl Popper and considered falsification as the decisive criterion of demarcation, there would be remarkably few, if any, theories in social science…. [I]t is the case in social science that there is greater difficulty in specifying the limiting conditions within which any particular theory will apply. Thus in natural science, a theory, law, or hypothesis is stated in such a way that the conditions under which it will operate are specified…. In social science the identity of cases is far more difficult to establish and hence the conditions under which a theory may be said to be applicable or not applicable are much more difficult to establish.

(Webb, 1995, pp. 87-88)

A key concept in Popper’s falsificationist philosophy of science is the asymmetry between verifications, which are deemed ambiguous, and falsifications, which are regarded as conclusive. McGinn (2002) has attacked this distinction as specious:

[T]here is something contrived and artificial about setting up an opposition here: for falsifying a statement is equivalent to verifying its negation. If I make an observation that falsifies the statement that Jones is in the next room (I go there and have a look), I thereby verify the statement that Jones is not [italics in original] in the next room. A reason to reject a statement is a reason to accept its negation; so it cannot be that rational inquiry consists solely [italics in original] in the rejection of statements.

(McGinn, 2002, p. 49)

Anderson, Hughes, and Sharrock (1986) have made the point that many theories in the human sciences are not falsifiable, and thus Popper’s falsifiability criterion “rules out famous theories in the social [human] sciences” (pp. 241-242). Accordingly:

As Popper himself saw, the implication of falsificationism for the social sciences is enormous. Many known and trusted theories — Marxism, psychoanalysis, behaviourism — simply do not match up to the theoretical requirements that falsification entails. They do not express testable [italics in original] hypotheses; they do not make precise [italics in original] predictions; they do not state the
grounds on which they would count themselves as refuted. In all cases, at least in Popper’s eyes, what social scientific theories are offering are sets of organising categories, ways of looking at social life, not scientific theories.

(Anderson et al., 1986, p. 240)

Arguing that Popper’s falsifiability standard, as embraced by the United States Supreme Court in the *Daubert* case, is not the appropriate criterion for the introduction of scientific evidence, Mason (2001) has commented thus:

The goal of science is knowledge. Whether knowledge will ever involve acquaintance with truth is a mystery. Even so, Popperian uncertainty is not necessarily the appropriate view for courts to take in determining the admissibility of scientific evidence. The Daubert standard is flawed in that it is overly restrictive concerning evidence that might be valuable (but not falsifiable), and it is overly inclusive of evidence that might be worthless (but meets the requirements of so-called “good science”). Thus the standard, and the difficulty with which it has been implemented in trial courts, tends to thwart the goals of the law-justice and the pursuit of truth. An alternative should be sought that recognizes differently the value of scientific knowledge and that can be incorporated more easily by a trial judge or jury.

(Mason, 2001, pp. 902-903)

Reduced to its essence, the core arguments of Popper’s critics are that scientists do not — and should not — go about their work with the aim of discrediting and rejecting the very theories that they have created. One such critic, Quine, has presented an alternative account of science wherein Popper’s skeptical falsifiers are replaced by even more skeptical revisers, who are “hoping to save [their theories] rather than refute [them]” (Quine, 2000a, p. 6).

**QUINE’S PHILOSOPHY OF SCIENCE**

For Quine, falsifications are on the same epistemological footing as verifications, and hence both are viewed as equally suspect and inconclusive. Quine has starkly distinguished his approach from that of Popper thus: “Karl Popper argued that experiment can only refute hypotheses, not prove them. I hold that experiment is fallible both ways” (Quine, 2000c, p. 412). Accordingly, Quine has viewed *both* verifications and falsifications as tentative and defeasible, true only for the time being and subject to revision. This skeptical stance toward all evidence, confirming and disconfirming, derives from Quine’s holism, a nonfoundationalist philosophy of science that views all knowledge as a seamless, interconnected “web of belief” (Quine & Ullian, 1970) that, unlike Popper’s falsificationism, admits of no demarcation between science and
nonscience (Quine, 1951, p. 20). Thus, philosophy, logic, mathematics, and the human, or social, sciences are all continuous with natural science (see Smart, 1969, p. 3; Ben-Menahem, 2003, p. 4). At the core of holism is the notion that all beliefs — even the laws of logic and mathematics — are in principle revisable (universal revisability). “No statement [in a larger theoretical network] is immune to revision” (Quine, 1951, p. 40). Murphy (1990) has described Quine’s philosophy of science as “postmodern”:

Quine not only replaced the foundationalist theory of knowledge with a holist account, but also provided a new picture or metaphor — that of a web or network of beliefs — to replace the ‘layer-cake’ model [the logical positivist bottom-up conception of science as consisting of three levels, with observation reports on the bottom, empirical generalizations in the middle, and theories on top]. (Murphy, 1990, p. 294)

For Quine, like for Popper, all observation is theory-laden, refracted through the lens of preconceived theoretical assumptions (see Quine, 2000a, p. 5; see also Hanson, 1972; Brewer & Lambert, 2001). However, both Quine and Popper have agreed to provisionally accept observation statements, or reports, as correct as a matter of convention (Laudan, 1990, p. 44) because, despite the limitations of evidence perceived by our senses, “whatever evidence there is [italics in original] for science is [italics in original] sensory evidence” (Quine, 1969, p. 75), with “observation, however inconclusive, [being] the locus of evidence” (Quine, 2000c, p. 412).

According to Quine, a theory is actually a complex whole, or “theory-bundle,” consisting of “a substantial bundle of interlocking [components]” (Quine, 2000c, p. 412), including the theory or hypothesis under investigation and various auxiliary hypotheses, or background assumptions. Contrary to a Popperian perspective, under which even isolated, individual hypotheses of a larger theory are falsifiable piecemeal (see Vuillemin, 1986, p. 595; Simkin, 1993, p. 167), on a Quinean view, only an entire theory — the “theory-bundle” — is subject to being tested, with the constituent individual hypotheses of that theory not being separately falsifiable (Quine, 1951; see also Harding, 1976).

**The holism, or Duhem-Quine, thesis**

Quine (1951) has brought renewed scholarly attention to the holism thesis that had been initially advanced by Duhem (1906), which suggested that the researcher confronting a contrary experimental result is not compelled to reject or alter the theory or...
hypothesis under investigation, but instead can change any of the auxiliary hypotheses and thereby rescue the theory or hypothesis under investigation.

Boghossian has illustrated the Duhemian argument thus:

Suppose that an experimental observation is inconsistent with a theory that you believe: the theory predicts that the needle will read “10” and the needle does not budge from zero. What Duhem pointed out is that this does not necessarily refute the theory. For the observational prediction is generated not merely on the basis of the theory, but, in addition, through the use of auxiliary hypotheses about the functioning of the experimental apparatus. In light of the recalcitrant observational result, \textit{something} [italics in original] has to be revised, but so far we do not yet know exactly what: perhaps it is the theory; perhaps it is the auxiliary hypotheses. Perhaps, indeed, it is the very claim that we recorded a genuinely recalcitrant result, as opposed to merely suffering some visual illusion.  

(Boghossian, 2001, p. 8)

Duhem’s argument was informed by his conventionalism, a position within philosophy of science that holds that a scientist’s choice of theory or hypothesis is not governed solely by empirical findings, but rather is determined by conventions that assist in the organization of observation and knowledge. Thus, “theories evolve by convention, on the basis of considerations like simplicity, not merely on the basis of their ability to withstand falsification” (Rosenthal & Rosnow, 1991, p. 34).

Quine has extended Duhem’s thesis, which was solely concerned with physical theory, to all knowledge (Quine, 1986). Moreover, “[i]n taking logic and mathematics to be continuous with science, and therefore revisable when experience so mandates, Quine’s holism goes beyond Duhem’s” (Ben-Menahem, 2003, p. 4). Quine has articulated his generalization of Duhem’s notion — variously referred to as the holism thesis (see Gibson, 2000, p. 82), the Duhem-Quine thesis (see Quine, 1975, p. 313), the Quine-Duhem thesis (see Bechtel, 1988, p. 42), and “the problem of auxiliary hypotheses” (see Fitelson & Waterman, forthcoming, ¶ 1) — thus:

[Si]cientific statements are not separately vulnerable to adverse observations, because it is only jointly as a theory that they imply their observable consequences. Any one of the statements can be adhered to in the face of adverse observations, by revising others.

(Quine, 1975, p. 313)

According to Quine (1975, pp. 314-315), “[i]n the face of recalcitrant observations, we are free to choose what statements [in a larger theoretical network or “theory-bundle”] to revise and what ones to hold fast.” Thus, “[a] recalcitrant experience can … be accommodated by any of various alternative re-evaluations in various
alternative quarters of the total [theoretical] system” (Quine, 1951, p. 40). For Quine, these revisions — even revisions of the laws of logic — can be made on pragmatic grounds (e.g., see Sher, 2002, p. 574; see also Sher, 1991). Accordingly, Quine has rejected the notion that there exists a special category of statements, referred to as analytic statements, that are not revisable because the meaning of the words and the laws of logic — not experience — make them necessarily true (e.g., “All and only bachelors are unmarried men,” Quine, 1951, p. 28). Discarding the idea of analyticity, Quine has held that there cannot be any knowledge independent of experience. “[A]ll statements are, in principle, answerable to experience, and, conversely, all statements can be maintained in the face of recalcitrant experience as long as we adjust other parts of our picture of the world” (Leiter, 2003, p. 44). Anderson et al. (1986) have put the matter thus:

There is no distinction between statements which can be revised and statements which it is impossible to revise [because all statements are revisable], though there is one between those we are quite willing to revise and those that we are more than unwilling to alter. There is a difference in the intensity of our commitment to different statements. The fact is we are loath to alter some because we are so committed to them, not that we are committed to them because they are in principle unalterable.

(Anderson et al., 1986, p. 155)

The underdetermination thesis

This holism thesis advanced by Quine has given rise to his notion of the underdetermination of theories by evidence, referred to commonly as the underdetermination thesis, which holds that empirical evidence cannot support, or determine, the choice of one theory over another (Quine, 1975; see also Duhem, 1954; Kuhn, 1970c). Put another way, in principle, “[a]ny set of data can be fit by many different [mutually inconsistent] theories” (Weinberg, 1998, p. 51), “[a]nd so, the argument concludes, we are never in a position to know that any of these theories is the truth” (Papineau, 1996, p. 302). According to Hacking (1999, p. 73), Quine, in his underdetermination thesis, was making “a logical point” that “[e]ven if all possible data were in, there would still ‘in principle’ be infinitely many theories that were formally consistent with such data.” Thus, for Quine, choice of theory, or revision of belief, is not made on purely logical or rational grounds.
Quine has explained that the underdetermination thesis follows from the holism thesis because “[i]f in the face of adverse observations we are free always to choose among various adequate modifications of our theory [holism thesis], then presumably all possible observations are insufficient to determine theory uniquely [underdetermination thesis]” (Quine, 1975, p. 313). Both the holism and underdetermination theses represent a significant challenge to Popperian falsificationism because a theory’s “[predictive] failure falsifies only a block of theory as a whole, a conjunction of many statements. The failure shows that one or more of those statements are false, but it does not show which” (Quine, 1969, p. 79). Hence, on a Quinean view, it is impossible to conclusively falsify a theory.

In positing “that there are in principle an indefinite number of theories that fit the observed facts more or less adequately” (Ariew, 1984, p. 313), the underdetermination thesis allows scientists and researchers, when encountering a contrary empirical result, to choose one of three alternative strategies, or “theory-bundle” configurations, in order to restore consistency: (a) abandonment of the theory or hypothesis under investigation and retention of the auxiliary hypotheses; (b) retention of the theory or hypothesis under investigation and revision of any one or more of the auxiliary hypotheses (referred to as “auxiliary fudging” [Lipton, 1991, p. 142]), thereby rescuing the theory or hypothesis under investigation (“auxiliary fudging,” Lipton, 1991, p. 142); or (c) revision (adjustment or tinkering) of the theory or hypothesis under investigation and retention of the auxiliary hypotheses (“theory fudging,” Lipton, 1991, p. 142).

How should a scientist or researcher choose among the three alternative strategies or “theory-bundle” configurations? According to Quine, pragmatic norms, such as conservatism, simplicity, and generality (see discussion below), should offer — and do offer — guidance but are not determinative. However, it is but a short distance from Quine’s relativist view (see Quine, 1984; see also Anderson et al., 1986, p. 148), derived from the underdetermination thesis, that belief revision, or theory choice, is not strictly governed by logic or reason to the social constructivist position that social factors account for the “theory-bundle” that is selected (see Gellner, 1986, p. 122; Laudan, 1990, pp. 146-170). Briefly put, social constructivism logically follows from a Quinean — or, for that matter, any relativist — philosophy of science (see Hesse, 1980, pp. 32-33).
Kukla (2000) has remarked on the relationship between social constructivism and philosophy of science thus:

The epistemological claim associated with [social] constructivism is the thesis of *epistemic relativism* [italics in original]. This is the view that there is no absolute warrant for any belief — that rational warrant makes sense only relative to a culture, or an individual, or a paradigm.

(Kukla, 2000, p. 4)

Social constructivism

Social constructivism, holding that “scientific knowledge itself [has] to be understood as a social product” (Pickering, 1992, p. 1), advances “the epistemic claim that the correct explanation for why we have some particular belief has to do with the role that that belief plays in our social lives, and not exclusively with the evidence adduced in its favour” (Boghossian, 2001, p. 6). Thus, “the way we think about things … are not just consequences of the way the world is, but are conditioned by our immersion in a particular society” (Weinberg, 2000, p. 8). As embodied in the work of the Science Studies Unit at the University of Edinburgh in the “Strong Programme in the Sociology of Scientific Knowledge,” social constructivism “recognise[s] the socially situated character of all knowledge, and hence the need to interpret scientific theories always with regard to their sociocultural contexts and conditions of emergence” (Norris, 2000, pp. 21-22). The “Strong Programme” has put forward the thesis that “the commitments of scientists to specific beliefs cannot be explained without reference to social interests, even if those beliefs are judged to be rational or adequate” (Kemp, 2003, p. 311).

The social constructivist view, then, is that “there are other [i.e., social] forces working on a scientist besides evidence and the rules of scientific method,” and that “[i]t is these other [social] causes which take up the slack left by the evidence in shaping scientists’ beliefs” (Laudan, 1990, p. 157). In addition to being influenced by the work of Quine, social constructivism was inspired by Kuhn, “who famously argued that the course of scientific activity is shaped by the scientific community’s choice of a paradigm” (Kukla, 2000, p. 8), defined by Kuhn (1970c, p. 182) as “a common disciplinary matrix,” and described by Weinberg (1998, p. 48) as “a consensus view.” In fact, an illustration of a social constructivist explanation is the Kuhnian notion, originally suggested by Planck (1949), “that older scientists are more resistant to challenges to the
prevailing paradigm than younger ones because more of their prestige is bound up in maintaining the status quo” (Laudan, 1990, p. 157; see also Shapin, 1982). Although appearing irrational according to the traditional view of scientific practice, “such behavior is rationally permissible” from a social constructivist perspective, especially in light of the “blunting [of] the impact of apparent refutations” by the underdetermination thesis (Laudan, 1990, p. 156).

Naturalized epistemology

Quine’s holistic philosophy of science is informed by his naturalized epistemology. Traditionally, the central task of epistemology, the branch of philosophy that “deals with the nature of knowledge and belief” (Hitchcock, 2004, p. 2), has been the justification of the truth claims of science (Fuller, 1988, p. 18; Norris, 2000, p. 1). In fact, “[e]pistemology is often regarded as the heart of philosophy,” because “[p]hilosophers want to understand what constitutes knowledge and how it is justified” (Bynum & Moor, 1998, p. 7). Traditional epistemology, then, is a normative undertaking seeking to answer “[h]ow ought we to arrive at our beliefs?” (Kornblith, 1994, p. 1) and “[h]ow can [we] assess the reliability of [our] knowledge?” (Sokal, 2001, p. 21). Asserting that the quest of the logical positivists for a “first philosophy” that would provide a justification for science outside of science has been a failure, Quine has advocated a naturalized epistemology, a descriptive undertaking whose task is to answer “[h]ow do we arrive at our beliefs?” (Kornblith, 1994, p. 1), thus “abandon[ing] … the goal of a first philosophy” (Quine, 2004a, p. 305) and the Cartesian search for placing science on secure foundations. Such naturalized epistemology is simply “an account of the way we adapt our systems of beliefs to changing experience” (Katz, 1998, p. 72).

Quine has attacked traditional epistemology’s attempts at justification of knowledge by resort to arguments a priori (that is, arguments that are independent of experience, or sensory evidence, but rather that appeal to reason and involve analysis of concepts or terms), counseling epistemologists to “[s]top dreaming of deducing science from observations” (Quine, 1969, p. 76), “surrender … the epistemological burden to psychology” (p. 75), and thereby allow epistemology to “[f]all into place as a chapter of
psychology and hence of natural science” (p. 82). In his canonical formulation of naturalized epistemology, Quine has asserted:

> The stimulation of his sensory receptors is all the evidence anybody has had to go on, ultimately, in arriving at his picture of the world. Why not just see how this construction really proceeds? Why not settle for psychology? … If we are out simply to understand the link between observation and science, we are well advised to use any available information, including that provided by the very science whose link with observation we are seeking to understand.

((Quine, 1969, p. 75)

Clarifying and elaborating on the above statement, Quine has remarked:

> In asking … “Why not settle for psychology?” I did not mean … that psychology would advance the justification process. I meant “Let us just get clear on the psychology of what we are actually doing, and look elsewhere if at all [italics added] for justification.”

((Quine, 2000c, p. 412)

Although a prodigious body of philosophical commentary on Quine’s naturalized epistemology has sprung up since the publication of his seminal paper “Epistemology Naturalized” in 1969 (for recent literature reviews and bibliographies, see Orenstein & Kotatko, 2000; Nelson & Nelson, 2003; Gibson, 2004), with various scholars advancing varying interpretations, derivations, and extensions, a close reading of Quine’s work reveals a clear conception of this new epistemological enterprise (e.g., see Orenstein & Kotatko, 2000; see also Quine, 1991; 1995a). Quine has advocated that naturalized epistemology should be continuous with empirical behavioral (perceptual) psychology, with the aim of exploring “how evidence relates to theory, and in what ways one’s theory of nature transcends any available evidence,” especially in light of the asymmetrical relationship “between the meager input [of sensory evidence] and the torrential output [of theory]” (1969, p. 83). Under Quine’s naturalized epistemology, “[t]he philosopher’s epistemological interests are not abandoned, they are pursued within empirical psychology” (Hookway, 1988, p. 54). However, in calling for “[t]he abdication of epistemology to psychology” (Quine, 2000b, p. 410), Quine has insisted that “[t]he pertinent motivations and aptitudes remain those of the analytic philosopher rather than the experimental psychologist” (p. 410). Accordingly:

[Quine] does not think that it is possible for the philosopher to make an independent inquiry into the kinds of things that there are in order to then use philosophical findings to judge whether science rightly identifies these things. Rather, philosophy depends on the findings of science for these are the best
information we have of what kinds of things there are. It is from science that philosophy must take its lead.

(Anderson et al., 1986, p. 146)

For Quine, naturalized epistemology and behavioral psychology would be “joint participants in the venture of finding out how the world is made up, distinguished only by the relative generality of their respective questions, those of philosophy being the more general ones” (Anderson et al., 1986, p. 147). Moreover, naturalized epistemology would not involve philosophy in the “direct investigation of the world” (p. 148), its method of inquiry being “semantic ascent,” whereby “[i]nquiries are moved ‘up’ a level and instead of examining things directly we examine, instead, how we talk about things” (p. 148). Thus, a Quinean naturalized epistemology would have “the task of making explicit what had been tacit, and precise what had been vague; of exposing and resolving paradoxes, smoothing kinks, lopping off vestigial growths …” (Quine, 1960, p. 275). In this way, there will be “philosophical progress here for which we would not look to psychology” (Quine, 2000b, p. 410).

On a Quinean view, justification for our beliefs and knowledge — the principal activity of traditional epistemology — is not a necessary task for naturalized philosophy, but if justification should be desired, it should be sought not in foundational, or indubitable, beliefs but rather in the coherence of beliefs with each other (see Quine, 1984, p. 293). Under such a coherence approach to justification, then, “[a] belief is justified not because it is indubitable or is derived from some other indubitable beliefs, but because it coheres with other beliefs that jointly support each other” (Thagard, 2000, p. 5).

Although Quine’s conception of epistemology carries a new naturalistic mandate for behavioral psychology, it still retains traditional epistemology’s normative, or prescriptive, function, not by focusing on justification of knowledge claims by seeking refuge in a “first philosophy, firmer than science” (Quine, 1970, p. 2), but in using pragmatic norms in the revision of beliefs, including scientific hypotheses. The notion that all beliefs are revisable is the central, overarching theme of Quine’s naturalized epistemology and of his philosophy of science in general. For Quine, all knowledge is tentative and subject to adjustment in light of “sensory evidence,” or experience, which is itself corrigible but which is “[all the] evidence there is [italics in original] for science”
Philosophy of science -- prototype validation

(Quine, 1969, p. 75). Quine has adumbrated his conception of the genesis of scientific reasoning thus:

The naturalist philosopher begins his reasoning within the inherited world theory as a going concern. He tentatively believes all of it, but believes also that some unidentified portions are wrong. He tries to improve, clarify, and understand the system from within.

(Quine, 1981, p. 72)

Under Quine’s naturalized epistemology, then, the starting point of theory formulation and thus of all scientific knowledge is our “scientific heritage” (Quine, 1951, p. 43) or “the lore of our fathers” (Quine, 1963, p. 406), by which Quine has meant currently-held basic beliefs that we have inherited by virtue of “growing up as we do in a going culture” (Quine & Ullian, 1970, p. 53). Thus, “we begin … in the middle of things with the beliefs we do have” (Creath, 1990a, p. 60), some of which will be “revised as experience prompts while taking care to save as much as we can and yet to achieve as simple a belief system as we can” (p. 60). For Quine, then, all theory formulation and hypothesis generation are instances of belief revision. As we are confronted by experiences contrary to our previous beliefs, we will make revisions somewhere in our “web of belief” so as to make our beliefs and experiences consistent with each other. Rott (2000, p. 503) has described Quinean epistemology as “paint[ing] a picture of how we should, and for the most part do, accommodate our scientific heritage if we meet with recalcitrant experiences.” For Quine (1951, pp. 39-40), “there is much latitude of choice as to what statements [of beliefs] to re-evaluate in the light of any single contrary experience,” however, these reevaluations, or revisions, should be guided by pragmatic norms.

These pragmatic norms are heuristics that assist the scientist in belief revision and hypothesis generation (Quine, 1951, p. 43; 1991, p. 269; 1995a, p. 49; 1995b, p. 258; 2000b, p. 411). Conservatism and simplicity are the chief pragmatic norms. Conservatism, also called by Quine as the “maxim of minimum mutilation” (Quine, 1991, p. 268), refers to the principle of “retain[ing] those hypotheses that clash least with the rest of our body of beliefs” (Orenstein, 1977, p. 83). According to Quine and Ullian (1970):

The less rejection of prior beliefs required, the more plausible the hypothesis — other things being equal. The plausibility of a hypothesis varies inversely with the plausibility of the prior beliefs that it disallows.
The norm of simplicity, also known as Ockham’s razor, admonishes us not to multiply theoretical entities unnecessarily and to prefer simpler theories over complex ones. “When there are hypotheses to choose between, and their claims are equal except in respect of simplicity, we choose the one that looks simpler” (Quine & Ullian, 1970, p. 45). Another norm to guide belief revision is generality, which advises us that our beliefs or hypotheses should be articulated with sufficient generality so that our initial experimental results will hold in subsequent test situations even though the latter do not correspond exactly to the first experimental run (p. 44).

**Criticisms of Quine’s philosophy of science**

Quine’s work has been almost universally admired in philosophical circles, the consensus view being that “[t]he last half-century in philosophy certainly belonged to Quine” (Blackburn, 2001, p. 37; see also Anderson et al., 1986, pp. 144-145; West, 1989, p. 184; Kuklick, 2001, p. 252). In addition, Quine’s philosophy of science has increasingly received attention within the realm of legal scholarship and United States case law (e.g., see Bethel v. Jefferson, 1978; Mercado v. Ahmed, 1991; Blue Cross and Blue Shield of New Jersey, Inc. v. Philip Morris, Inc., 2001; Schroeder, 2001; Blakey & Murray, 2002; Jenkins, 2002; Biancalana, 2003; Coleman, 2003; Redmayne, 2003; Gruber, 2004; Hansen, 2004; Mitchell, 2004; Solan, 2004; Ohlin, 2005; see also Fuller, 1967, p. xii). Most criticisms of Quine’s philosophy of science have acknowledged the main contours of his holism and underdetermination theses while disagreeing with their importance, taking issue with their implications, or disputing some elements therein (e.g., see Popper, 1957; 1959; Glymour, 1975; 1980; Grünbaum, 1984; Stove, 1999; see also Laudan, 1965).

Quine’s holism thesis — the notion that scientific statements are not tested in isolation but only as parts of a larger theoretical network, so that a contrary finding can be accommodated by making revisions somewhere in the network — has proven to be a hardy insight that could be attacked only at its margins. For example, Stove (1999) has conceded the correctness of Quine’s holism thesis but has played down its significance:

“This [holism] thesis is undoubtedly true. But what kind of truth is it? … In fact the thesis is simply the most trivial of contingent truths about human beings: that
given any proposition whatever, a scientist (or anyone) can take it into his head to affirm it, and can then stick to it through thick and thin.  

(Stove, 1999, p. 57)

Similarly, Glymour (1980) has acknowledged that there is a “kernel of truth” (p. 151) in Quine’s holism thesis in that “we cannot assess hypotheses in a complex theory in isolation from their fellows,” but has rejected the Quinean “insist[ence] that evidence must bear on all of a theory … or none of it or that we must accept or reject our theories as a single piece” (p. 152). Glymour’s criticism of Quine is rooted in the practice of science in the real world:

Scientists often claim that an experiment or observation tests certain hypotheses within a complex theory but not others. Relativity theorists, for example, are unanimous in the judgment that measurements of the gravitational red shift do not test the field equations of general relativity …. Observations are regarded as relevant to some hypotheses in a theory but not relevant to others in that same theory.

(Glymour, 1975, p. 403)

For Glymour (1980), like for Quine, the individual component hypotheses of a larger theoretical network, or “theory-bundle,” are initially tested together. However, according to Glymour, unlike for Quine, the scientist now chooses which of the individual hypotheses are to be retained for subsequent individual testing and which are to be rejected. Grünbaum (1984, p. 98) has referred to this procedure as “piecemeal testing within an overall theory [italics in original].” According to Glymour (1980):

When something goes wrong with a theory we may, after investigation, retain a best tested part of the theory and reject the rest. The interweaving of hypotheses means just that the pieces of our theory must be assessed together. What to believe and what to discard must depend on what else we believe and what else we discard.

(Glymour, 1980, p. 152)

As explored earlier, Popper (1957; 1959) has assailed Quine’s holism thesis for allowing of post-experiment or post-observation revisions of auxiliary hypotheses (e.g., instrumentation in good working order) to explain away a contrary finding and thereby save from refutation the theory under investigation. Admitting the possibility of such a Quinean rescue strategy, Popper has questioned the strategy’s legitimacy and has condemned its use in general, asserting that such preservation of the theory comes at the cost of diminished scientific status.
Quine’s underdetermination thesis — the argument that in principle there are many equally plausible and mutually inconsistent theories to explain a given set of experimental or observational data — has been challenged in some philosophical quarters as technically correct for the most part but of limited practical consequence (e.g., see Lipton, 1991). Okasha (2002) has summarized the two central criticisms of the underdetermination thesis:

In principle, there will always be more than one possible explanation of a given set of observations. But … it does not follow that all of these possible explanations are as good as one another. … [Moreover,] there are relatively few real cases of underdetermination in the history of science. … Indeed, when we inspect the historical record, the situation is almost exactly the reverse of the underdetermination argument would lead us to expect. Far from scientists being faced with a large number of alternative explanations of their observational data, they often have difficulty finding even one [italics in original] theory that fits the data adequately.

(Okasha, 2002, p. 73)

Similarly, Lipton (1991) has observed:

[Quine’s underdetermination thesis] is an important point, but it may blind the philosopher to the actual situation of the working scientist, which is almost the opposite of what underdetermination suggests. Often, the scientist’s problem is not to choose between many equally attractive theoretical systems, but to find even one. Where sensible accounts are scarce, there may be a great temptation to fudge what may be the only otherwise attractive account the scientist has been able to invent. The freedom to choose a completely different [theoretical] system is cold comfort if you can’t think of one.

(Lipton, 1991, p. 146)

The strengths and weaknesses of a Quinean stance make it the appropriate philosophy of science for some endeavors but not for others. Not all scientific activity conforms to a Quinean philosophy of science. For instance, Klein and Herskovitz (in press) have argued that a Quinean perspective is ill suited for computer simulation validation because, in accordance with Quine’s view that all falsifications as well as all verifications are ambiguous, a falsified model cannot be conclusively rejected. Hence, a Quinean philosophy of science discourages the improvement of computer simulation models as model developers will not know when to reject a model and make attempts to build a potentially better alternative. The adoption of a Quinean approach under such circumstances will lead to “epistemological nihilism” (Quine, 1969, p. 88). Accordingly, with respect to computer simulation validation, Klein and Herskovitz have concluded that a Popperian philosophy of science, which provides model developers with a firm decision
rule of conclusively rejecting a falsified model, is the appropriate stance as it would encourage the building of improved models.

PROTOTYPE VALIDATION AS A QUINEAN UNDERTAKING

In line with the work of Herskovitz (1991) and Klein and Herskovitz (in press), which suggests that different philosophies of science may be appropriate for various discrete scientific endeavors, the thesis of this paper is that the activity of prototype validation should follow — and does follow — a Quinean script. Specifically, we argue that the systems developer and prototype user are collaborators in a Quinean enterprise in which the systems developer, upon construction of the prototype, presents it to the prototype user for validation, that is, for an appraisal of how the prototype conforms to the user’s initial mental model or, if circumstances have changed for the user, to the user’s revised mental model. The user has three options: (a) accepting the prototype as completely conforming to the user’s mental model; (b) requesting revisions in the prototype or revising the mental model so that they are in conformity with each other; and (c) rejecting the prototype as not conforming to the user’s mental model.

Rejection of the prototype is reserved for only the most egregious situations, for example, when there is a total mismatch between the prototype and the user’s mental model, or when there has been a significant change in the user’s circumstances (see Carr, 2005; Lichtblau, 2005; Schmitt, 2005). In most circumstances, the user will not reject the prototype outright and scuttle the entire system if the prototype does not conform to the user’s mental model. Rather, the user will request modifications or adjustments to fine-tune the prototype. There may be, and usually are, several iterations, or rounds, to the prototype validation process.

Under a Popperian perspective, if the prototype in any way does not conform to the user’s mental model, the revision option is not available and the prototype is rejected. Such rejection belies the underlying assumption upon which prototyping is grounded: a prototype is constructed with the intention that it will be revised as the systems developer and prototype end user interact with each other and obtain a better or more precise understanding of the user’s requirements. The expectation is that fine-tuning will be
needed to bring the prototype and the user’s initial or revised mental model into conformity with each other. For example, Özcan (1998) has noted:

> Our experience suggests that the existence of a software tool to automate a task often alters users’ perception of what the task involves. As a result, it is possible that even those requirements that are perceived to be well-understood may still need to be modified.

(Özcan, 1998, p. 1373)

As revisions are at the very heart of prototyping (e.g., see Bødker & Grønbæk, 1991, p. 212), a Popperian falsificationist philosophy of science, which prohibits prototype revisions and mandates a rejection if the prototype and the user’s mental model diverge in any way, is ill suited as a theoretical framework for prototype validation.

By contrast, the application of a Quinean philosophy of science to prototype validation allows the prototype to be revised, which, in fact, is the main advantage of the prototyping approach to ISD (see Dearnley & Mayhew, 1983). As the prototype user is expected to modify, tweak, and tinker with the prototype until it conforms to the user’s initial or revised mental model, it is a Quinean philosophy of science that provides a suitable theory for prototype validation.

Anderson et al. (1986, p. 236) have captured the significance of the revision option under a Quinean approach thus: “In the face of disconfirmation we are not given only one alternative, namely rejecting the theory [as embodied in the prototype].” Rosenthal and Rosnow (1991, p. 35) have recognized that the essential distinction between the Popperian and Quinean philosophies of science is that the former holds that scientific theories that are empirically falsified must be unambiguously rejected and that only those that survive attempts at falsification should be retained, while the latter asserts that “there is no such thing as a completely decisive falsifying test because when a refutation occurs, it merely tells us that the general formulation needs to be adjusted [italics added], not that it needs to be discarded.” Quine’s holistic philosophy of science, then, is a philosophy of revision, “an epistemology of reevaluation” (Katz, 1998, p. 72). As such, it is well suited to serve as a theoretical framework for prototype validation.

**Pragmatic norms**

According to a Quinean philosophy of science, as embodied in Quine’s naturalized epistemology, revisions to prototypes should be — and are — guided, or
restrained, by pragmatic norms, such as conservatism, simplicity, and generality (see discussion above). Concerning the norm of conservatism, if in the user’s judgment the prototype does not conform to the user’s mental model or if circumstances have changed between the user’s construction of the mental model and the building of the prototype, the systems developer should make the least extreme changes to accommodate the user’s mental model. The systems developer needs to modify the prototype in such a way so as to preserve as much of the original features as possible while satisfying the user’s requirements. According to Dearnley and Mayhew (1983, p. 41), “[a] good prototype must be built on the basis of information already gathered, keeping in view the information required.” Drastic revisions and major overhauls may result in the loss of user confidence in the system and increased costs (see Dearnley & Mayhew, 1983).

Another aspect of the norm of conservatism is consistency with generally accepted information systems design practices and standards. Revisions to the prototype should cohere with current custom and usage in the profession. Morcor (2005) has made this point thus:

Like the stars in the sky, certain features and buttons must be in a certain place on the screen so users feel at home. … We spent a lot of time talking to users about the positioning of buttons. Closer to the top means a certain thing, over towards the right means another. There is a certain personality to the user interface that the users get to know like a reliable friend.

(Morcor, 2005, ¶ 9)

With respect to the norm of simplicity, both the systems developer and the prototype user should understand that the revisions should be as simple as possible to accomplish the task needed or the result requested. In fact, the user should bear in mind the simplicity norm when constructing the mental model. Simplicity is a key concept in ISD in general and in prototyping in particular. The benefits of simplicity are not insignificant:

Web sites and software often compete with each other based on the features they provide. The popular assumption is that the more features a product has, the better it will be. The truth is that features improve a product only if they are actually used by the customer. In most cases the proliferation of features in products creates more complexity than value. Each feature gets an icon or a link on a Web site or toolbar, and is yet another item that the user needs to wade through before they can find the one that they need.

(Berkun, 1999, ¶ 1)
For Morcor (2005, ¶ 10), simplicity has no shortage of virtues: “Minimum mouse clicks, minimum keystrokes, minimum screens. The fewer things that have to be done, the more powerful the user feels. Powerful users are happy and effective users.” Writing in the context of Web design, Nielsen (2000) has observed:

A general principle for all user interface design is to go through all of your design elements and remove them one at a time. If the design works as well without a certain design element, kill it. Simplicity always wins over complexity, especially on the Web where every five bytes saves is a millisecond less download time.

Nielsen (2000, p. 22)

Dearnley and Mayhew (1983) have made the case for simplicity in prototyping thus:

A prototype must be simple and [thereby] relatively quick to create, amend and rebuild. … The more simple a prototype is to build and modify, the faster the analyst can respond to the users’ criticisms and ideas. This speed is reassuring to the user, as he can see his comments being put into action, rather than have to wait weeks or months for the next version, by which time he may have lost enthusiasm.

(Dearnley & Mayhew, 1983, p. 41)

The norm of generality, in the context of ISD, prescribes that, in order to ensure maximum flexibility, prototypes, revisions to prototypes, and systems should not be hard-coded, but rather should contain “code that doesn’t need to change with every change in the details of its input data” (Aster, 1998, p. 5). For example, customer billing software that includes tax calculations should be written in generalizable code so that if tax conditions change (e.g., business expands to other tax jurisdictions or tax rate increases), the software can be easily adapted. After the systems developer presents the prototype to the user, the latter should test the prototype with a variety of different inputs and make the former aware of features that are subject to change. If it turns out that the prototype does not contain generalizable code for these features, the systems developer can make the appropriate revisions at that point.

**Social factors**

We argue that, in addition to being guided by pragmatic norms, prototype revision is determined by social, or sociological, factors, such as professional training and professional self-interest (see Taylor, 1989). This insight is derived from the social constructivist approach in the sociology of science (see discussion above), which holds
“that social causes are always present” (Brown, 1984, p. 9), along with other causes, as determinants of belief revision. As we have noted earlier, social constructivism has been held to be the logical implication of Quine’s underdetermination thesis that there are many — and in principle, an infinite number — alternative theories that fit one set of data. Accordingly, the argument goes, theory choice, or belief revision, is determined by extra-scientific beliefs, that is, beliefs which are not grounded in reason or experience, and so “it follows that social [sociological] factors must be invoked to explain why a scientist adopts a particular theory [or revises a particular belief]” (Ariew, 1984, p. 313).

To assert, as social constructivists do, that scientific beliefs are socially determined is to suggest that scientists form particular social beliefs as a consequence of belonging to a community of scholars, which sets professional standards, offers rewards of peer recognition, endorses certain formulae, provides exemplars of research, values some behaviors, and establishes model curricula for education and training of future scientists.

In accordance with a social constructivist approach, it is suggested that, for example, in the case of a prototype of a customer relationship manager (CRM) being built for a user law firm, there will be many revision requests from the lawyers on what will appear to the systems developer as trivial or debatable discrepancies between the prototype and the user’s model. We argue that this is so because finding flaws, attentiveness to minute details, and probing, interrogative questioning are behaviors that are encouraged by the legal community in law school, continuing legal education, and law reviews (e.g., see Byse, 1986; Silecchia, 1996; Kerr, 1999; Baker, 2000; Bennett, 2001; DeJarnatt, 2002; Butleritchie, 2002-2003; Ramsfield, 2003; Proctor, 2004). Moreover, these behaviors are in the lawyers’ professional self-interest as lawyers are rewarded for these activities in their professional practices (see Raasch, 2004a; 2004b).

Similarly, in developing for a hospital an electronic health record (EHR) system prototype with the aim of having data entry forms in fixed, discrete formats, the systems developer should anticipate that physician users — many of whom are reluctant to use computers in the first instance (Martin, 1999; Watkins et al., 1999; Quintero et al., 2001; Cross, 2002; News Target Network, 2005) — will request prototype revisions that will enable them to enter their clinical notes and observations in any format (see Ranganathan, Watson-Manheim & Keeler, 2004). For physicians, especially the older ones, writing
with pen on paper is a well-established medical tradition that was observed in medical school and that has continued in medical practice (see Landro, 2005).

IS researchers (e.g., see Naumann & Jenkins, 1982; Narayanan, Bailey, Tendulkar, Daley, Pliske & Wilson, 2002) have long recognized what Flynn and Jazi (1998, p. 53) have referred to as “the user-developer culture gap” in ISD, whereby systems developer and user have differing perspectives — and even different vocabularies (see Özcan, 1998, p. 1360) — so that they perceive the same problem from different vantage points. Systems developers, for example, tend to view user requirements as technical concerns, which are known from the outset and do not change, while paying “inadequate attention … to the social context within which the computer system will function, with the result that many systems eventually fail” (Flynn & Jazi, 1998, p. 54). By contrast, users view the requirements through the prism of their domain of interest (e.g., profession, specialization, occupation). Thus, according to the social constructivist insight, both the systems developer and user conceive user requirements differently by virtue of belonging to different professional or occupational communities with varying social beliefs.

Flynn and Jazi (1998) have argued that, in interacting with each other, the systems developer and user revise their socially determined beliefs concerning user requirements and arrive at a set of requirements that they have jointly socially constructed. According to Flynn and Jazi (1998):

[T]he requirements process is a social process and that it is based on the principles of iteration, which may occur within and between rounds, and user involvement. We take the view that requirements are not objective artefacts, available at the start of the requirements process …. Rather, requirements are emergent: they are socially constructed by the interactions involving users and developers in the requirements process [italics in original].

(Flynn & Jazi, 1998, p. 56)

The research of Flynn and Jazi, which dealt with ISD in general and did not address itself specifically to prototyping, suggests that the user’s requests for prototype revisions, and the systems developer’s responses, are influenced by beliefs widely held by other members of the professional or occupational group they belong to and that user-developer interaction will subsequently produce a new set of socially constructed prototype revisions.
The recent scholarly study by Lloyd and Sivin (2002) considerably bears on our treatment of how sociocultural circumstances shape prototype revision. Working within the history of science tradition, Lloyd and Sivin have argued that early Chinese and Greek science and medicine have developed differently because of the existence of different sociological factors, as encapsulated in a “cultural manifold,” a global term that captures, or summarizes, the interaction among social, institutional, and political dimensions:

[W]e found that we were investigating what we have come to call … a cultural manifold. Rather than comparing concepts or factors one at a time, we begin with the commonplace observation that scientific ideas or medical insights do not occur in a vacuum. They grow in the minds of people with a certain kind of education and a certain kind of livelihood and are inseparable from the rest of their experience.

(Lloyd & Sivin, 2002, p. xi)

Keyser (2004, p. 62) has characterized the cultural manifold as “the continuum of thinkers’ concepts, social goals, professional milieu, mode of discourse, and political associations.”

According to Lloyd and Sivin (2002, p. 247), “Ancient Greek culture encouraged disagreement and disputation in natural philosophy and science as in every other field; the Chinese emphasized consensus.” Starkly put, “the Chinese were collaborative, the Greeks competitive; in China agreement was sought out or else assumed to exist, in Greece rivalry flourished and was promoted” (Barnes, 2003, p. 24). These divergent stances, born of different social beliefs in different cultures, influenced, for instance, ancient Greek and Chinese modes of scientific inquiry. Thus:

The dominant … Greek way was through the search for foundations, the demand for demonstration, for incontrovertibility. … The principal … Chinese approach was to find and explore correspondences [and accordingly,] favor[ing] the formation of syntheses unifying widely divergent fields of inquiry.

(Lloyd & Sivin, 2002, p. 250)

Importing and extending the reach of Lloyd and Sivin’s historical and sociological insights concerning ancient Greek and Chinese scientific development to prototype validation, we suggest that different professions and occupations have different “cultural manifolds,” which will influence the content, form, and frequency of prototype revisions (see Huysman, 2002; Mark & Poltrock, 2004; see also Nakakoji, 1996; Lehto & Marttiin, 2000; Choe, 2004). Consider again our previous example of the construction of
a CRM prototype for a law firm. We conjecture that there is a cultural divide between systems developers, who hail from a professional community with a collaborative work culture and an emphasis on seeking consensus (Borsook, 2000, p. 233), and lawyers, who come from a competition-oriented work culture where a high premium is placed on adversariality (Ramsfield, 2003; Proctor, 2004). Thus, as suggested above, such lawyer users will generally seek prototype revisions for slight and arguable discrepancies between the prototype and the user’s mental model, although the collegial orientation of the systems developers should facilitate the joint social construction of revisions to the prototype. Future researchers from both the IS and sociology of knowledge disciplines may wish to further investigate how the differing “cultural manifolds” of systems developers and users influence prototype revisions.

CONCLUSION

In this paper, we have set for ourselves the task of situating a theoretical framework for prototype validation within the philosophy of science. Toward that end, we have compared the Popperian and Quinean accounts of scientific knowledge and have argued that a Quinean philosophy of science, whose cornerstone principle is that all beliefs are revisable, actually corresponds to the activity of prototype validation, and that such correspondence is well warranted. Specifically, we have suggested that prototype revisions are belief revisions, and, as such, are determined by pragmatic norms and social, or sociological, factors.

Our thesis, then, is that the systems developer and prototype end user have joined forces — not as Popperian falsifiers, who have a decision rule to reject the prototype on account of any divergence, major or minor, from the user’s mental model, but — as Quinean revisers, with the objective of fine-tuning the prototype (or the user’s mental model) so that the prototype and the user’s mental model are congruent with each other. Rather than focus on searching for inconsistencies between the prototype and the user’s mental model — which would invalidate the prototype — in the manner of Popperian falsifiers, the systems developer and prototype user adopt the stance of Quinean revisers to save the prototype from rejection by removing inconsistencies via adjustments.
REFERENCES
Agassi, J. (1988) *The gentle art of philosophical polemics: Selected reviews and comments*. Open Court, La Salle, IL.
November 2004,


City of Tuscaloosa v. Harcros Chems., Inc., 158 F.3d 548 (11th Cir. 1998).


Fitelson, B. & Waterman, A. (Forthcoming) Bayesian confirmation and auxiliary hypotheses revisited: a reply to Strevens. *British Journal for the Philosophy of Science*. 
Philosophy of science -- prototype validation


Quine, W. V. (1995b) Naturalism; or, living within one’s means. Dialectica, 49, 251-261.


