Proceedings of DETC '00
2000 ASME Design Engineering Technical Conferences
September 10-14, 2000, Baltimore, Maryland

DETC2000/DTM-14579

VALIDATING DESIGN METHODS & RESEARCH: THE VALIDATION SQUARE

Kjartan Pedersen
Kvaerner Oil and Gas AS
Postboks 169
1324 Lysaker
Norway

Jan Emblemsvåg
Considium Consulting Group AS
Postboks 1115
1361 Billingstad
Norway

Reid Bailey, Janet K. Allen[*] and Farrokh Mistree
Systems Realization Laboratory
The George W. Woodruff School of Mechanical Engineering
Georgia Institute of Technology
Atlanta, GA 30332-0405
USA

ABSTRACT

Validation of engineering research is typically anchored in the scientific inquiry tradition that is based primarily on logical induction and / or deduction. Since much engineering research is based on mathematical modeling, this kind of validation has worked – and still works – very well. There are, however, other areas of engineering research that rely on subjective statements as well as mathematical modeling, which makes this type of validation problematic. One such area is that of design methods within the field of engineering design. In this paper, we explore the question of *how one validates design research in general, and design methods in particular.*

Being anchored in the scientific inquiry tradition, research validation is strongly tied to a fundamental problem addressed in epistemology, namely, *what is scientific knowledge and how is new knowledge confirmed?* Thus, we first look to epistemology for answers to why an approach solely based on 'formal, rigorous and quantifiable' validation constitutes a problem, and for an augmented approach to research validation. We then propose the 'Validation Square' which we validate by testing its internal consistency based on logic in addition to testing its external relevance based on its usefulness with respect to a purpose.

We recognize that no one has *the* complete answer to the question we pose. To help us converge on an answer to these questions we "think aloud" and invite you to join us in doing the same. It is our hope that in so doing we, the members of this design research community, will all be the richer for it.

KEYWORDS: Philosophy of science, epistemology, engineering design, research validation.

NOMENCLATURE:

DSP	Decision Support Problem
EPV	Empirical Performance Validity
ESV	Empirical Structural Validity
TPV	Theoretical Performance Validity
TSV	Theoretical Structural Validity

1. WHAT IS SCIENTIFIC KNOWLEDGE? Searching for a New Approach to Design Method Validation

Validation refers to internal consistency (i.e., a logical problem), whereas *verification* deals with justification of knowledge claims. In modeling literature, these terms are swapped, and in this paper we use the terms as used in the modeling literature; i.e., *verification* refers to internal consistency, whereas *validation* refers to justification of knowledge claims (Barlas and Carpenter 1990). Validation of engineering research is anchored in the tradition of the scientific method. This tradition demands "formal, rigorous and quantitative validation" (Barlas and Carpenter 1990), which is based primarily on logical induction and / or deduction. Since much engineering research is based on mathematical modeling, this kind of validation has worked – and still works – very well. There are, however, other areas of engineering research that rely on subjective statements as well as mathematical modeling in which validation that is rooted in 'formal, rigorous and

---

[*] Corresponding Author, to whom correspondence should be addressed: janet.allen@me.gatech.edu, Phone/FAX: 404-894-8168/9342.

quantitative' measures becomes problematic. One such area is that of design methods within the field of engineering design. In this paper, we explore the question *how does one validate design research in general, and design methods in particular?*

In this section we address the question: *what are the problems encountered in implementing a formal, rigorous and quantifiable scheme to validate a method in engineering design?* We investigate this question by going to the roots of epistemology to evaluate the applicability of the fundamental assumptions upon which 'formal, rigorous and quantifiable' validation rest. Then, based on the literature, we propose new assumptions where old assumptions fail, before proposing a new approach to research validation based on a new set of assumptions.

1.1 The Historical Roots of Modern Epistemology

Epistemology (the theory of knowledge) started in ancient Greece with Phyrro and his skeptics. They tried to produce 'a criterion for truth', a search that was strongly influenced by Plato and Aristotle. Plato, who defined knowledge as "that over which there cannot be error", confined knowledge to a particular realm of perfect and unchangeable entities referred to as "Forms" (Honderich 1995). Plato later acknowledged that "correct belief can be turned into knowledge by means of a reason or cause", something Aristotle acknowledged. This constitutes the basis for foundationalism, where knowledge of the world rests on a "foundation of indubitable beliefs from which further propositions can be inferred to produce a superstructure of known truths" (Honderich 1995). Moreover, it is from this foundationalist basis that modern epistemology emerged in the seventeenth century, starting with Rene Descartes and rationalism.

1.2 The Foundationalist/Formalist/ Reductionist School of Epistemology

The rationalists asserted that "the truth is innate and prior to all experience" and that "human knowledge about the truth is based on reasoning" (Descartes [1641] 1931). The empiricists asserted that "all human knowledge about the truth is based on experience rather than reasoning" (Locke [1690] 1894). Both views, however, are based on the fundamental assumption that truths are absolute and innate, which links them to the views of Plato and Aristotle, and hence, to foundationalism.

The foundationalist view was brought forward by Bertrand Russell who introduced logical atomism, and by his student Ludwig Wittgenstein. With his "Tractatus Logico-philosophicus" Wittgenstein brought the atomist and foundationalist tradition to full fruition by asserting that the "function of philosophy is to monitor the bounds of sense, and to show that attempts to traverse the bounds of sense are futile" (Honderich 1995; Wittgenstein [1921] 1961). This became the basis for logical positivism, a movement which was in favor in the scientific community until the 1960's. Their doctrine was centered around the 'verification principle' asserting that "knowledge can only be claimed if judged true by meaning [analytically true] or true by virtue of experience [synthetically true]" (Honderich 1995). Hence, non-quantifiable, synthetic propositions are neither true nor false, that is, they are meaningless. From this it follows that unless statements can be formalized for analytical and/or empirical investigation, they are meaningless. Hence, most positivists consider metaphysical, religious, aesthetic, and ethical claims as inferior to those produced by science, resulting in an extreme focus on science in general and mathematical proofs in particular. This 'urge' to formalize statements into mathematics (to allow analytical judgements) links logical positivism to formalism, a view that is integral to many different philosophical schools which share the fundamental assumption that rational knowledge is the only valid knowledge.

Logical positivism became 'obsolete' in the late 1960's, however, many of the basic ideas of atomism and foundationalism are embodied in what later came to be known as reductionism. Reductionism is a wide term and is normally split into ontological, methodological and theory reduction. *Ontological reductionists* postulate that the whole of realities consists of a minimal number of materialistic substances. Hence, they deny the existence of immaterial phenomena and advocate "biological organisms to be no more than complex functioning machines". *Methodological reductionists* postulate that the properties of the whole are the sum of the properties of the parts. Hence, analysis of the parts is sufficient to gain knowledge about the whole. *Theory reductionists* assert that new theories absorb old theories rather than replace them. From this it follows that biology, for instance, will in the end be totally explained by chemistry and/or physics. In modern science, methodological reductionism has been the most influential reductive approach with the discovery of DNA as perhaps the most important triumph. Although successful, building on the assumptions that knowledge is innate and absolute and can only be verified by reason, reductionists are totally dependent on objective quantification. Hence, reductionism is based on the fundamental assumption that objectivity exists

In the preceding, we have documented that the tradition of scientific inquiry demanding 'formal, rigorous and quantitative' validation is anchored in the foundationalist/formalist/ reductionist school of epistemology. From this it follows that 'formal, rigorous and quantitative' validation is based on the fundamental assumptions that:

1) truths (knowledge) are innate and absolute,
2) that only rational knowledge is valid, and
3) that objectivity exists.

Having identified 'formal, rigorous and quantitative' validation as problematic when validating research that is based on subjective statements, we assert that the fundamental assumptions (1 through 3 above) are at the core of the problems. To substantiate this assertion we turn to the literature.

1.3 The Relativistic/Holistic/Social School of Epistemology – Challenging the 'Ruling' Fundamental Assumptions of Knowledge

The notion of innate and absolute truths was first challenged by Immanuel Kant who synthesized rationalism and empiricism in a search for knowledge on neutral ground by asserting that "all knowledge starts with experience" however, "not all knowledge arises out of experience" (Kant [1781] 1933). Hence, he suggested that not all truths are innate and absolute – there are some that might be added by the mind. This raises the question: *who is to determine what is given (i.e., innate and fundamental) and what is derived/added?* This question is referred to as "the myth of the given" in (Sellars 1963). Hegel, on the other hand, rejected the whole idea of innate truths and introduced a new logic in which conflict and contradictions are regarded as necessary elements of truth (thesis, antithesis, and synthesis). As a consequence Hegel regarded truth as a process rather than a fixed state of things. In his view knowledge is socially, culturally, and historically dependent, hence, there are no neutral foundations of knowledge, and entirely objective verification of knowledge claims is not possible (Hegel [1817] 1959). This view was supported by Thomas Kuhn who presented a historical analysis of how science progresses, and he argued that in any given epoch scientists work within and against the background of an unquestioned theory or set of beliefs (a paradigm). According to Capra: "scientific facts emerge out of an entire constellation of human perceptions, values, and actions – in one word, out of a paradigm – from which they cannot be separated" (Capra 1996). Based on this we assert that scientific knowledge is not innate nor absolute, strictly, but are inseparable from the social scientific context within which they are developed.

Science progresses, according to Kuhn, when the ruling paradigm cannot provide adequate explanations to scientific problems under investigation, and this inadequacy makes way for new paradigms. Central to Kuhn's view is that the change to a new paradigm cannot be based on strictly logical reason (Kuhn [1962] 1970). This is supported by Quine who argues that, "we choose a particular way of doing it [i.e., accommodate a new theory to an experiment] not because some absolute scientific principle [i.e., based on rationality] but because it is convenient, causing minimal disturbance in the existing theory". (Quine 1953). This links validation to preferences based on usefulness, similarly to Toulmin's proposal that a 'better' method is equivalent to a more useful method. This is important since it challenges the notion that only rational knowledge is valid knowledge.

Rational knowledge, or rational beliefs, is arrived at by accumulating and evaluating an adequate body of relevant evidence (Honderich 1995). The accumulation and evaluation of scientific evidence is addressed in the Scientific Method, wherein Sir Francis Bacon suggested that scientific knowledge is gained and claimed by a process of induction. This again requires rigorous rules, where formal logic and/or mathematics are preferred. The underlying assumption is of course that following the rules is a rational act in itself. This assumption fails, since determining which rules to follow also requires rules. Hence, total rational assessments are based on an infinite regress and therefore impossible. In reality the choice of rules is contextual as pointed out in the previous paragraph (i.e., dependent on the ruling paradigm). Hence, our ability to be rational depends on a basic ability to exercise intelligent judgement that cannot be completely captured in systems of rules, i.e., they are not accessible to investigation through the senses or calculation. This is the definition of intuitive knowledge, and it is ironic that hypotheses (the cornerstone of the scientific method) often are proposed as a result of intuitive processes (Honderich 1995). According to Albert Einstein expressed:

> *"The justification (truth content) of the system [physics] rests in the proof of usefulness of the resulting theorems on the basis of sense experiences, where the relations to the latter to the former can only be comprehended intuitively" (Einstein 1950)*

Based on this we assert that scientific knowledge is anchored in the rationality for facts, and on intuition for values. As a consequence, intuitive knowledge is linked to the application of rational knowledge through the determination of purpose.

The impossibility of total rational assessments also challenges the very existence of objectivity, the last of the fundamental assumptions upon which formal, rigorous and quantifiable validation rest. This assumption is also challenged by Werner Heisenberg who claims that a procedure for acquiring knowledge will affect the acquired knowledge itself (Capra 1991). Albert Einstein was also aware of this problem and he stated that "one may compare these rules [related to the scientific method] with the rules of a game in which, while the rules are arbitrary, it is their rigidity alone which makes the game possible. However, the fixation will never be final. It will have validity only for a special field of application". *What about the objectivity of mathematics?* Wittgenstein addressed the issue of objectivity in mathematics, and claimed that "logic [mathematics] is merely a tool consistent only within itself and hence content free" (Wittgenstein [1921] 1961). This view was supported by Kurt Gödel who claimed in his "Incompleteness Theorem" that "every formal number theory contains an indecisive formula, i.e., neither the formula nor its negation is provable in the theory" (Gödel 1931). From this it follows that attempting to prove something formally and/or objectively is an illogical and hence, an invalid act since the underlying axiom of such an approach, i.e., that objectivity exists, is already logically refuted by Wittgenstein and Gödel. Ultimately this leads to the proposition that a conversational, contextual and subjective validation approach is more logical, and therefore more formal, since it does not refute its own axioms, i.e., that subjectivity is unavoidable. Based on this we assert that total objectivity does not exist, and hence, that knowledge validation must be linked to contextual usefulness.

A new school of epistemology is based on the refutation of the fundamental assumptions upon which the foundationalist/ reductionist/formalist school of epistemology rests, namely, the relativist/holistic/social school of epistemology (Barlas and Carpenter 1990). Needless to say, we adhere to the relativist school of epistemology, and hence, we adopt a relativist view on scientific knowledge. What remains, however, is to evaluate what impact the different views on scientific knowledge have on research validation, which we do next.

1.4 Different Views of Knowledge: The Impact on Research Validation

Logical empiricist validation is a strictly formal, algorithmic, reductionist, and 'confrontational' process, where new knowledge is either true of false. The validation becomes a matter of formal accuracy rather than practical use. This approach is appropriate for closed problems that have right or wrong answers associated with them, like mathematical expressions or algorithms. *Relativist validation*, on the other hand, is a semiformal and communicative process, where validation is seen as a gradual process of building confidence in the usefulness of the new knowledge (with respect to a purpose). This approach is appropriate for open problems, where new knowledge is associated with heuristics and non-precise representations.

Through addressing how to validate engineering design research, we are addressing the fundamental nature of engineering design. We assert that engineering design is primarily concerned with open problems that involve objective and subjective elements and no single right answer. This separates design from most of the traditional engineering disciplines, in which a given problem has only one right answer. Engineering design requires both science *and* art to achieve a goal, separating it as fundamentally different from the analytical aspects of engineering. A formal, rigorous and quantifiable validation procedure only acknowledges the closed part of engineering design, while ignoring the significance of subjectivity. Hence, a relativist validation procedure is asserted in this paper.

As stated, the principal objective in this paper is to synthesize a framework for validating design methods. There are no right or wrong answers to this problem, there are many heuristics involved, and non-precise representations are common. Based on this assumption the relativist validation is adopted, and the validation strategy is based on the following statement.

We define scientific knowledge within the field of engineering design as socially justifiable belief according to the Relativistic School of Epistemology. We do so due to the open nature of design method synthesis, where new knowledge is associated with heuristics and non-precise representations, thus knowledge validation becomes a process of building confidence in its usefulness with respect to a purpose.

Accordingly, we assert that formal, rigorous and quantifiable validation (i.e., based on logic) can be applied to a design method's internal consistency but fails to validate its external relevance (i.e., its usefulness). Hence, formal, rigorous and quantifiable validation is necessary but not sufficient, and we therefore suggest including the validation of a method's usefulness with respect to a purpose as well. We further suggest that the validation of a method's usefulness be done using a set of carefully chosen example problems that will support a claim of generality.

The example problems that we are speaking of are synonymous with case studies from (Yin, 1994). Yin distinguishes cases studies from sampling units of an experiment based on the role of theoretical propositions. In an experiment, the sampling units are chosen randomly and their power is in their large numbers. In a case study, however, each case is connected to a specific theory. Accepting the usefulness of a design method for some case studies, then, is a matter of assessing whether the cases support or refute the theory. In this way, "individual case studies are to be selected as a laboratory investigator selects the topic of a new experiment" (Yin, 1994, pg. 31) and not as a laboratory investigator selects the samples used within an experiment.

In the next section we introduce the 'Validation Square' that provides a framework for validating internal consistency as well as external relevance for some particular instances in order to build confidence in its general usefulness with respect to a purpose.

2. THE VALIDATION SQUARE – A Process of Building Confidence in Usefulness –

In the previous section we asserted that research validation is a process of building confidence in its usefulness with respect to a purpose. We associate usefulness of a design method with whether the method provides design solutions 'correctly' (effectiveness), and whether it provides 'correct' design solutions (efficiency). Correct in this context are design solutions with acceptable operational performance, that are designed and realized with less cost and/or in less time. Hence, the process we present aims at evaluating the effectiveness and the efficiency of the method, based on qualitative and quantitative measures respectively. This is illustrated in Figure 1, where the Validation Square at the bottom is the synthesis of this process, and this process is detailed next.

2.1 Structural Validation – A Qualitative Process

As can be seen from Figure 1, being *effective* embodies three things:

(1) accepting the individual constructs constituting the method;
(2) accepting the internal consistency of the way the constructs are put together in the method; and
(3) accepting the appropriateness of the example problems that will be used to verify the performance of the method.

A description of each follows.

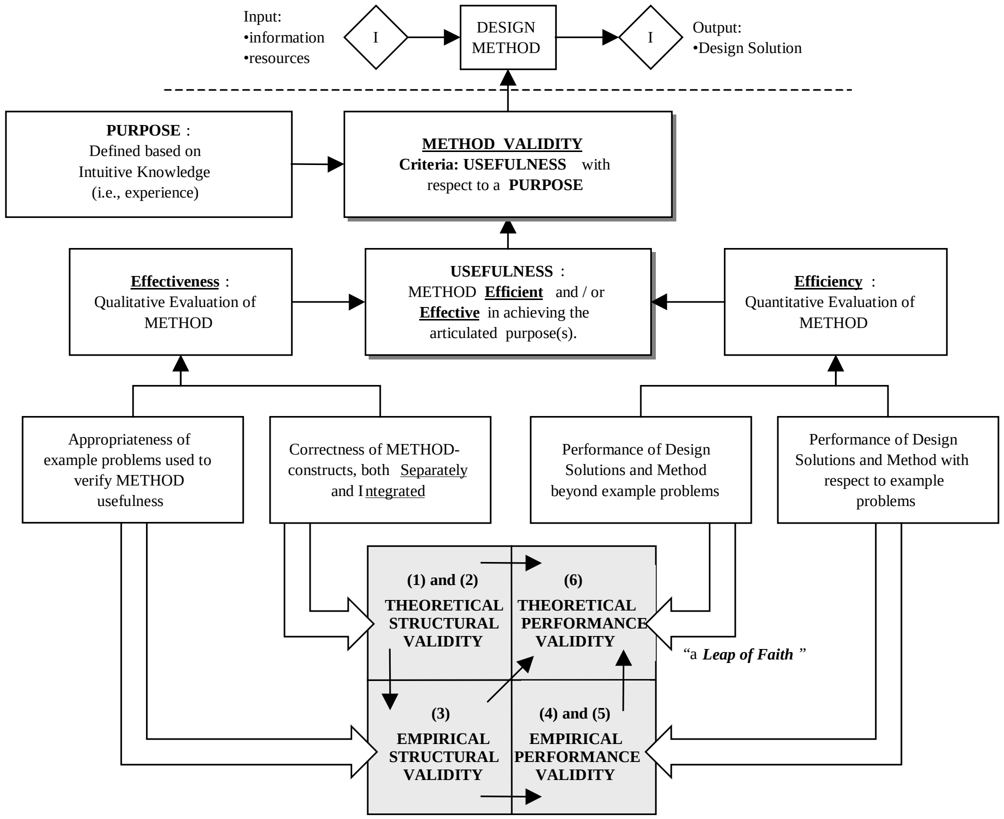


Figure 1 --

Design Method Validation - Building Confidence in Usefulness.
The Validation Square is in the gray box.

(1) Accepting the construct's validity: In order to build confidence in the validity of the individual constructs constituting the method, we suggest using the literature. Based on the name of the author and publisher, the number of references associated with the construct, how long the construct has been referenced, and so, an inference towards acceptance can be built. In addition, if the constructs are being used as benchmarking for new constructs, they should be demonstrated to be highly accepted and valued.

(2) Accepting method consistency: In order to build confidence in the way the constructs are put together in the method (i.e., in the method's internal consistency) we suggest using flowchart representations focusing on information flow. In this way it can easily be demonstrated that for each step (construct) there is adequate input available, that the anticipated output from the step (construct) is likely to occur based on the input, and that the anticipated output is an adequate input to another step (construct). Further, identifying the information flow unveils what information is assumed to be readily available, hence, facilitates evaluation against reality. Method inconsistency refers to generating information that is inadequate or not necessary, or invalid assumptions upon which the method rests.

(3) Accepting the example problems: In order to build confidence in the appropriateness of the example problems chosen for verifying the method performance, we suggest documentation in stages. First, document that the example problems are similar to the problems for which the method constructs are generally accepted. Then, document that the example problems represent the actual problem for which the method is intended. Finally, document that the data associated with the example problems can support a conclusion.

As can be seen, the validity of the method constructs – individually (1) and integrated (2) – deals with the structural soundness of the method in a more general sense, and are therefore denoted Theoretical Structural Validity. The validity of the example problems for which the method is to be tested (3) deals with the structural soundness for some particular instances, and are therefore denoted Empirical Structural Validity, However, both 'validities' are evaluated qualitatively.

2.2 Performance Validation – A Quantitative Process

As can be seen from Figure 1 being *efficient* embodies three things:

(4) accepting that the outcome of the method is useful with respect to the initial purpose for some chosen example problem(s);

(5) accepting that the achieved usefulness is linked to applying the method; and

(6) accepting that the usefulness of the method is beyond the case studies.

A description of each follows.

(4) Accepting usefulness of method for some example problems: To build confidence in the usefulness of the method, we suggest using representative example problems. In this way, the outcome of the method can be evaluated in terms of its usefulness. As indicated, metrics for usefulness are linked to the degree an articulated purpose has been achieved. However, the purpose for proposing a design method may vary. From an industrial perspective the purpose is typically linked to reducing cost and/or time and/or improving quality. From a scholarly perspective, the purpose is augmented to include addition of scientific knowledge that can help produce more scientific knowledge.

(5) Accepting that usefulness is linked to applying the method: To build confidence that the usefulness of the resulting example problem solutions is linked to applying the method, we suggest evaluating the contributions to usefulness from each construct individually. This is done by comparing the solutions with and without the construct, allowing a quantitative evaluation. In addition, solutions should be compared to those found with existing design approaches. In terms of Yin's case studies, Step 5 is a matter of evaluating rival theories, in which alternative explanations for the usefulness of a case are investigated.

(6) Accepting usefulness of method beyond example problems: To build confidence in generality, we suggest using induction that entails the following:

In (1), we demonstrate that the individual constructs are generally accepted for some limited applications.

In (2), we demonstrate the internal consistency of the way the constructs are put together in the method.

In (3), we demonstrate that the constructs are applied within their accepted ranges.

In (4), we demonstrate the usefulness of the method for some chosen example problems, which in (3) are demonstrated to be appropriate for testing the method.

In (5), we demonstrate that the usefulness achieved is due to applying the method.

Based on this we claim generality, i.e., that the method is useful beyond the tested example problems. However, as shown in Section 1.3, every validation rests ultimately on socially justifiable belief, that is, faith. Hence, the purpose of going through the Validation Square is to present 'circumstantial' evidence to facilitate a leap of faith, i.e., to produce belief in a general usefulness of the method with respect to an articulated purpose.

The greatest impact of applying Yin's case study approach to validation is in establishing the external validity of a design method. *External validity* refers to the validity of an approach to situations other than the individual cases studied in Steps 4 and 5 of the Validation Square. By avoiding the temptation to treat case studies as a sampling units in a statistical experiment, the method of generalization is 'analytic generalization,' in which a previously developed theory is used as a template with which to compare the empirical results of the case study. If two or more cases are shown to support the same theory, replication may be claimed (Yin, 1994, pg. 31). Each case is treated as a different experiment (where analytic generalization is supported with a small number of cases), not as different points within the same experiment (where statistical generalization requires a large number of samples). Generalization (i.e., support for external validity) is not accomplished with a handful of data points, but instead with a handful of case studies which assess a theoretical proposition. Through connecting the example problems to analytic generalization and theory development, the "Leap of Faith" depicted in Figure 1 as part of the step from empirical to theoretical performance validity is not necessarily a large leap. If the method is deemed useful for some limited instances (4) and (5), we denote this as Empirical Performance Validity. Similarly, if the method is deemed useful beyond some limited instances (6), i.e., useful in a more general sense, we denote this as Theoretical Performance Validity.

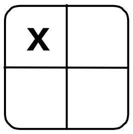
Having proposed a framework for validating design methods, namely the Validation Square, this framework itself needs to be validated.

3. VALIDATING THE VALIDATION SQUARE

In this section, we provide a brief albeit cursory overview of our thinking on this topic and the status of validation. Further details are presented in Pedersen (1999). Much more needs to be done and we look forward to other members of the design research community joining us in this effort.

We validate the 'Validation Square' by validating its internal consistency in addition to its external relevance, i.e., we apply the 'Validation Square' to validate itself. For those who regard this as circular argumentation, we refer them to mathematics - mathematics is validated by means of mathematics.

3.1 The Theoretical Structural Validity of the Validation Square



The Theoretical Structural Validity (TSV) of the Validation Square refers to accepting the structural/logical soundness of its constructs, both individually and integrally. The constructs of the Validation Square are represented by the four internal squares as illustrated in Figure 1: Theoretical Structural Validity (TSV), Empirical Structural Validity (ESV), Empirical Performance Validity (EPV) and Theoretical Performance Validity (TPV). Hence, accepting the TSV of the Validation Square implies accepting the structural/logical soundness of each of these 'validities' in addition to accepting that they are put together in a logical and consistent manner.

(1) Accepting the individual 'validities': It should not be hard to accept that any tools/constructs/etc. to be used in a design method must be individually valid, and that the tools/constructs/etc. must be organized in a way that creates internal consistency. Hence, we assert that TSV as outlined in Section 2.1 is a structurally (logically) valid proposition.

This view is supported in (Hazelrigg 1999). However, Hazelrigg asserts that "validation of a design alternative selection method can be done only mathematically, and only through validation of the procedure, not by verification through results", a view based on (Barzilai 1998). We find it difficult to accept the preceding statement. Both Hazelrigg and Barzilai must be viewed as advocating 'rigorous, formal, and quantifiable' validation, an approach based on fundamental assumptions which we have questioned. That is, they are only considering the closed, objective and scientific part of engineering design. Further, Hazelrigg argues that "only if all steps in a procedure are valid, that is, rational, self-consistent, and derivable from axioms, is it possible for a method [as a whole] to be valid". This is based on the view that rational assessments yield universal and necessary results; hence, if methods produce significantly different results with identical input, at most one of the methods can be logically valid. Hazerigg's approach is further challenged by bounded rationality, where human cognitive limitations are viewed to yield a considerable scope for rational disagreement (Honderich 1995). Hence, we suggest that the presented axioms in (Hazelrigg 1999) can be challenged and logically refuted by using a different set of rules.

Finally, we contend that the total disregard for the usefulness of the results as part of validation, as suggested by Hazelrigg, is illogical. Let us explain. The purpose of a typical 'alternative selection method' is to recommend the best alternative based on some criteria. Hence, what is really important is to have a good set of alternatives, and the right set of criteria for what is best. Both of these aspects cannot be addressed by rational assessments, hence, the usefulness of having a totally rational selection process becomes limited; the process with which an alternative is recommended is valid, but the alternative may not be valid. This is very well demonstrated by 'the battle' between Windows 3.1 and IBM's OS2 operative system; most experts deemed OS2 to be a better operative system, however, the OS2 lost the competition. Non-quantifiable, social issues prevented the superior OS2 from winning the support of computer users.

Based on this observation we assert that internal consistency alone does not assure external relevance, hence, we feel that validation has to be augmented to address external relevance by evaluating usefulness. In this context, it should not be hard to accept that any example problem used to verify a method's usefulness has to:

a) be similar to the problems for which the tools/constructs/etc. are generally accepted;
b) be representative of the problems for which the method is intended; and
c) provide sufficient data to support a conclusion.

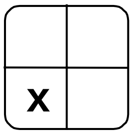
Hence, we assert that ESV as outlined in Section 2.1 is a structurally valid proposition.

Accordingly, it should not be hard to accept that unless a design method is able to produce useful results (i.e., design solutions) for a particular problem it cannot be deemed valid for that particular problem. (It does not mean, however, that the method is invalid in general.) Hence, we assert that EPV as outlined in Section 2.2 is a structurally valid proposition.

Finally, the TPV-validity is based on the other three internal 'validities' being accepted, as well as accepting that the induction presented in Section 2.2 (6) is logically valid. The first is given in the previous paragraphs, while the latter is dealt with next.

(*2) Accepting the Validation Square consistency:* We assert that the induction presented in Section 2.2 (6) is logically valid. To substantiate this claim we turn to the literature. The consistency of viewing validation as a process of building confidence in usefulness (with respect to a purpose) is argued in Section 1.3. Here we argue that a conversational, contextual and subjective validation approach is logical, and therefore formal, since it does not refute its own axioms, i.e., that subjectivity is unavoidable. This view is fully supported in Emblemsvåg 1999, and is based on a tradition of building confidence based on posits, (for details see Chen 1995; Lewis, 1996; Koch 1997; Peplinski, 1997; Simpson 1998). Further, splitting this process of building confidence in one structural (qualitative) and one performance (quantitative) part, and in one theoretical (general) and one empirical (particular) part comes from the area of system dynamics (Richardson and Pugh 1981). This area has been heavily criticized for not employing 'formal, rigorous, objective and quantitative' model validation procedures (Barlas and Carpenter 1990; Barlas 1996), which makes this approach even more interesting from the perspective of design method validation. The first time we have noticed these validation aspects used in the area of engineering design is in Bailey 1997, where they are arranged in a square. This was subsequently developed by Bailey et al., 1999. However, the prescriptive and comprehensive Validation Square as presented in this paper is first presented Pedersen 1999, where it was used to validate that the doctoral research contributed new scientific knowledge to the field of engineering design. Finally, it has been used in Siddique 1999 to validate doctoral research as well. Thus we assert that the Validation Square is Theoretically Structurally Valid.

3.2 The Empirical Structural Validity of the Validation Square



As stated in Section 2.2, any example problem intended for method testing, has to be validated itself and deemed appropriate. As mentioned in the previous paragraph, the Validation Square has been used to validate a design method, namely, the Hierarchical Product Platform Realization Method, or the HPPRM for short (Pedersen 1999). Hence, we use the HPPRM as our example problem to test the usefulness of the Validation Square and show in the following its appropriateness.

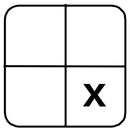
(3) Accepting the HPPRM as an example problem: The HPPRM is a design method intended for realizing large and complex made-to-order systems that are very expensive and produced in small numbers. The outcome of the HPPRM is a so-called Hierarchical Product Platform (or HPP for short), which is a product platform serving as a basis for products addressing different market segments. The HPPRM consists of three phases, namely, Define, Model, and Solve. Each of these phases is centered around an independent construct, namely, Numerical Taxonomy, Technology Diffusion, and the compromise DSP, see Figure 2.

- Numerical Taxonomy (Sneath and Sokal 1973) is used to identify the potential for standardization in an existing design portfolio by means of clustering. The clustering itself is ‘objective’ whereas interpreting the clusters is based on subjective judgements.
- Technology Diffusion (Silverberg, Dosi et al. 1988; Silverberg 1991; Hall 1994) is used to discount the performance of alternative technologies according to their maturity and leverage potential to existing technology. Application of the discounting factor is objective whereas decisions regarding learning rates and leverage potentials are based on subjective judgements.
- The compromise DSP (Mistree, Hughes et al. 1993) is used to enable designers to minimize the distance to their goals for the total system with respect to operational performance, time and cost. Solving the compromise DSP is objective in the mathematical sense whereas deciding on scenarios as well as making the final decision is based on subjective judgements.

Applying the HPPRM is preparing subjective input to objective mathematical constructs that produces output which is judged subjectively; hence, it complies with the kind of problems for which the Validation Square is intended.

Further, in (Pedersen 1999) HPP’s were developed for a family of gravitational separators and a family of marginal field vessels, which constitutes two instances for which the HPPRM is tested. Hence, we claim that this is sufficient data to support a conclusion regarding the usefulness of the HPPRM, and thus, regarding the usefulness of the Validation Square. Based on the following we assert that the HPPRM is an appropriate example problem to test the usefulness of the Validation Square, i.e., using the HPPRM is Empirically Structurally Valid.

3.3 The Empirical Performance Validity of the Validation Square



The purpose of applying the Validation Square to the HPPRM is to build confidence in its validity. Hence, the usefulness of the Validation Square is linked the degree to which confidence is built (4) and to whether this confidence is due to applying the Validation Square (5).

(4) Accepting usefulness of the Validation Square: In order to build confidence in HPPRM validity we applied the Validation Square in the following manner.

- We demonstrated by means of the literature that the core constructs of the HPPRM are generally accepted for their intended applications. Further, we demonstrated by means of flow chart representation that there is no redundant information being generated and the underlying assumptions are valid. Based on this we asserted Theoretical Structural Validity.
- We demonstrated by means of induction that a (1) *Gravitational Separator* example problem was appropriate for exemplifying and illustrating the HPPRM in detail, and (2) that the *Marginal Field Vessel* example problem was appropriate for testing the usefulness of the HPPRM. Based on this we asserted Empirical Structural Validity.
- We demonstrated by means of the example problems that (1) the resulting HPPs represented feasible solutions realizable in less time with less cost, and (2) that each construct contributed to usefulness: Numerical Taxonomy by reducing the combinatorial problem to a manageable size, Technology Diffusion by advocating a continuous improvement approach, and compromise DSP by advocating robust and adaptive solutions. Based on this we asserted Empirical Performance Validity.
- Based on the outcome of the previous steps in the Validation Square we inferred general validity of the HPPRM based on the following. (1) The key method constructs are applicable for problems beyond the example problems, (2) the example problems are representative of the general problem, and (3) the HPPRM is useful for the representative example problems. Based on this we asserted Theoretical Performance Validity.

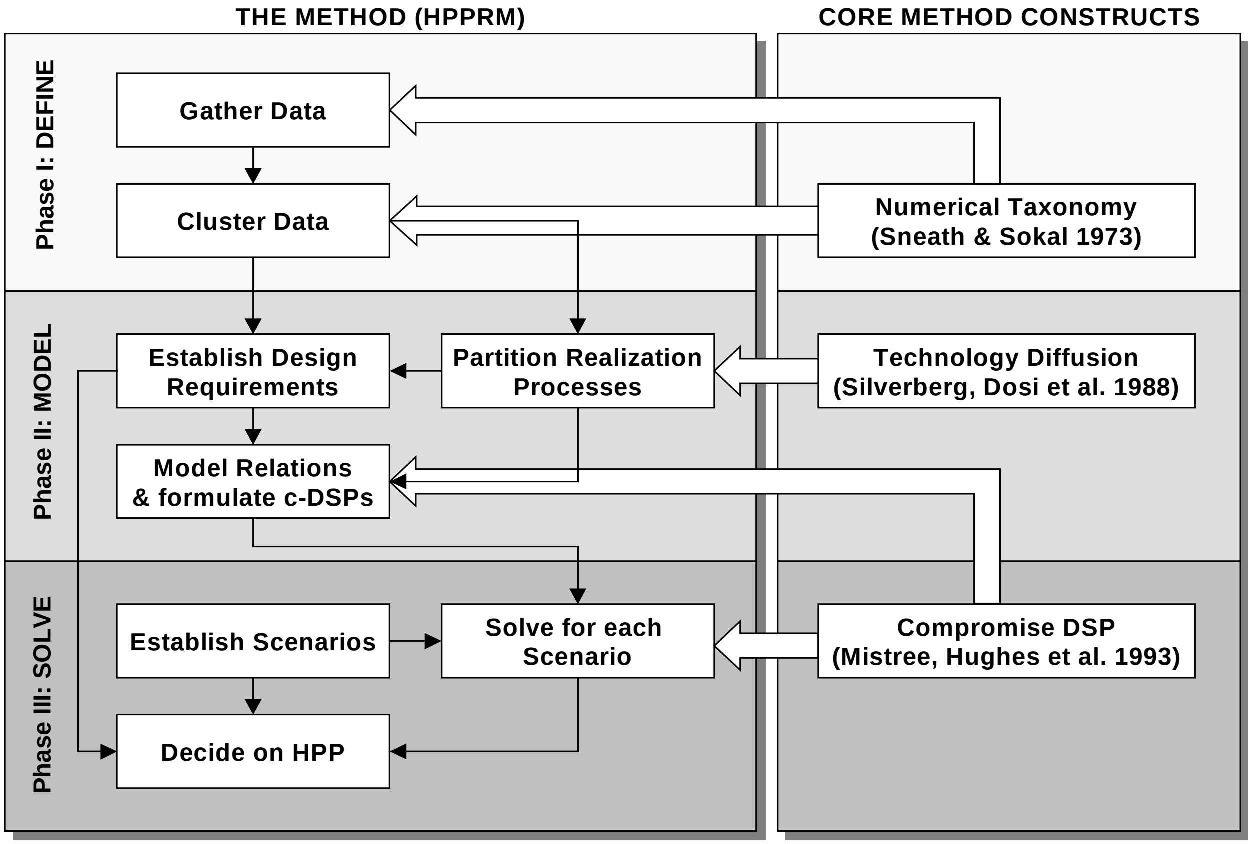


Figure 2 --
The HPPRM and its Constructs: An Overview

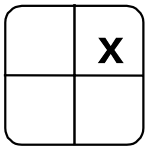
We assert that the Validation Square as applied in Pedersen 1999, built confidence in usefulness with respect to a purpose. What remains is to evaluate whether this confidence was due to the Validation Square.

(5) Accepting that usefulness is linked to applying the Validation Square: Theoretical Performance Validity implies that all other 'validities' are accepted, and hence, each of the 'validities' is necessary but not sufficient. However, confidence can be built with very little evidence presented**;** it all comes down to a socially justifiable belief as pointed out in Section 1.3. Nevertheless, we claim that the likelihood of accepting something as valid increases with the amount of relevant and accepted evidence.

We have now demonstrated that applying the Validation Square provided a sufficient amount of evidence to build confidence in validity for at least one instance, namely, the HPPRM. Based on this we assert that the Validation Square has Empirical Performance Validity.

3.4 The Theoretical Performance Validity of the Validation Square

The general validity of the Validation Square is based on accepting its usefulness beyond the example problem, which is presented next.



(6) Accepting usefulness of the Validition Square beyond the HPPRM example: We assert that the Validation Square is useful to validate design methods with mixed subjective and objective statements beyond the HPPRM example, and we substantiate this claim as follows.

- As long as a design method can be evaluated for its likelihood of 'producing' the wanted outcome, it can be deemed Theoretically Structurally Valid.
- As long as an example problem can be evaluated for its appropriateness regarding method testing, it can be deemed to have Empirical Structural Validity (validating the structure of the example problems).
- As long as a design method can be evaluated in terms of how well the resulting design solution performs compared

to design solutions arrived at by other methods, it can be deemed to have Empirical Performance Validity (validating the performance of the method for the examples).

- And finally, as long as inference can be made towards a design method's general usefulness, it can be deemed to have Theoretical Performance Validity (validating the performance of the method in general).

Based on this, we assert that the Validation Square is generally applicable for validating design methods in particular, and for validating research in general. The latter is based on expanding the previous induction in the following way.

- As long as any research result can be evaluated in terms of the likelihood of fulfilling its intended application, it can be deemed Theoretically Structurally Valid.
- As long as any test of research usefulness can be evaluated for its appropriateness, it can be deemed Empirically Structurally Valid.
- As long as any research results can be evaluated in terms of its usefulness for some particular applications, it can be deemed to have Empirical Performance Validity.
- And finally, as long as inference can be made towards the general usefulness of any research results, it can be deemed to have Theoretical Performance Validity.

Based on this we assert that the Validation Square has Theoretical Performance Validity, hence, it is deemed valid for validating new knowledge associated with heuristics and non-precise representations. From this it follows that we have achieved the principal objective of this paper, namely, to synthesize a framework to validate design methods in particular and research results in general. And we have done so according to the relativist/holistic/social school of epistemology, where scientific knowledge is defined as socially justifiable belief, and research validation is viewed as a process of building confidence with respect to a purpose.

4. CLOSURE

In this paper, we have questioned the fundamental assumptions upon which 'formal, rigorous and quantitative' validation rest, suggested a new set of assumptions and postulated a new hypothesis for knowledge validation in engineering design, namely, a relativist/holistic/social view (see Table 1). Through asserting this view on validation in engineering design, we recognize engineering design as having both scientific (i.e., objective) and artistic (i.e., subjective) components.

Based on the changed view, we assert that validating a design method is a contextual process of demonstrating usefulness with respect to a purpose. Based on this assertion we present a framework for guiding this process, namely, the Validation Square (see Figure 3). This framework builds on research in systems dynamics, and a tradition of using posits in engineering design. However, the Validation Square as presented in this paper extends all these efforts by offering a prescriptive approach that is more comprehensive and systematic.

The Validation Square has been used to validate a design method for realizing large and expensive made-to-order systems, namely, the HPPRM (Pedersen 1999), and to validate a method for product platform configuration (Siddique 1999). Based on this paper we assert that the Validation Square is appropriate for validating research results in general, as long as it can be subjected to qualitative and quantitative evaluation as outlined in Section 2.

By introducing *usefulness* and *purpose* into the validation procedure, we absolutely are not advocating that "anything goes." The validation square is specifically designed to bring rigor to a validation process that is both *quantitative* and *qualitative*. When dealing with open constructs such as design methods, open validation procedures are imperative. It is only in a relativistic validation procedure that the validity of an open construct can be fully assessed.

As we note in the abstract, we recognize that no one has the answer. We trust that you enjoyed thinking aloud with us. We now invite you to comment upon what we have presented so that we together can create something of value not only to us but for our student colleagues - the next generation researchers!

Table 1 -- Summary of Foundations for the Validation Square

Old View of Knowledge Validation	Fundamental Assumptions	Basis for Refutation	New Emerging Assumptions	New View on Knowledge Validation
Foundationalist	Knowledge is absolute/innate	Kant, Hegel, Sellars, Quine, Kuhn	Knowledge is socially justifiable belief	Relativist
Reductionist	Rationality only valid basis for knowledge	Honderich, Einstein	Intuition valid basis for defining purpose for application of knowledge	Holistic
Formalist	Objectivity exists	Hegel, Kuhn, Wittgenstein, Gødel, Einstein	Research validation linked to usefulness	Social and conversational

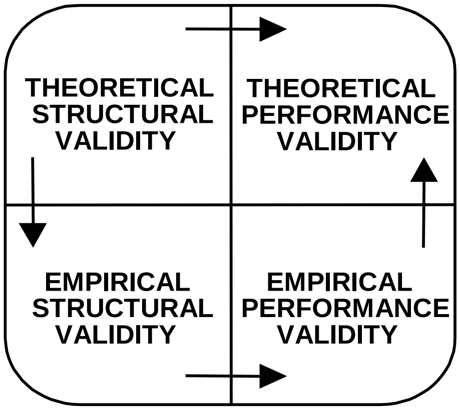


Figure 3 - - The Validation Square

The Validation Square has been used to validate a design method for realizing large and expensive made-to-order systems, namely, the HPPRM (Pedersen 1999), and to validate a method for product platform configuration (Siddique 1999). Based on this paper we assert that the Validation Square is appropriate for validating research results in general, as long as it can be subjected to qualitative and quantitative evaluation as outlined in Section 2.

By introducing *usefulness* and *purpose* into the validation procedure, we absolutely are not advocating that “anything goes.” The validation square is specifically designed to bring rigor to a validation process that is both *quantitative* and *qualitative*. When dealing with open constructs such as design methods, open validation procedures are imperative. It is only in a relativistic validation procedure that the validity of an open construct can be fully assessed.

As we note in the abstract, we recognize that no one has the answer. We trust that you enjoyed thinking aloud with us. We now invite you to comment upon what we have presented so that we together can create something of value not only to us but for our student colleagues - the next generation researchers!

ACKNOWLEDGMENTS

Kjartan Pedersen's doctoral study was supported both financially and by way of information by Kvaerner ASA, Norway. We acknowledge the financial support. We gratefully acknowledge the encouragement that we have received from Rolf Kvamsdahl (Kvaerner ASA) and Stian Erichsen (NTNU, Trondheim) to think out of the box and in making change a handmaiden of success. Reid Bailey has been supported by a Department of Energy Integrated Manufacturing Fellowship. Finally, we acknowledge the contribution of Carolyn Seepersad in identifying and smoothing over some of the rough spots in this paper.

REFERENCES

Bailey, R., 1997, *The Design of Industrial Ecosystems.* M.S. Thesis, The George W. Woodruff School of Mechanical Engineering, Georgia Institute of Technology, Atlanta, GA.

Bailey, R., Bras, B., and Allen, J.K., 1999, "Using Robust Concept Exploration and Systems Dynamics Models in the Design of Complex Systems," *Engineering Optimization*, Vol. 32, No. 1, pp. 33-58.

Barlas, Y., 1996, “Formal Aspects of Model Validity and Validation in System Dynamics.” *System Dynamics Review* Vol. 12, No. 3, pp. 183-210.

Barlas, Y. and S. Carpenter, 1990, “Philosophical Roots of Model Validation: Two Paradigms.” *System Dynamics Review* Vol. 6, No. 2, pp. 148-166.

Barzilai, J., 1998, "Measurement Foundations for Preference Function Modeling." *The 1998 IEEE Conference on Systems, Man, and Cybernetics*, San Diego, CA.

Capra, F., 1991, *The Tao of Physics*. Boston, MA, Shambala.

Capra, F., 1996, *The Web of Life*. New York, Doubleday.

Chen, W., 1995, *A Robust Concept Exploration Method for Configuring Complex Systems.* Ph.D. Dissertation, The George W. Woodruff School of Mechanical Engineering, Georgia Institute of Technology, Atlanta, GA.

Descartes, R., [1641] 1931, *Meditations on First Philosophy. The Philosophical Works of Rene Descartes*. Cambridge, Cambridge University Press.

Einstein, A., 1950, *The Theory of Relativity & Other Essays*. New York, MJF Books.

Emblemsvåg, J., 1999. *Activity-Based Life-Cycle Assessments in Design and Management.* The George W. Woodruff School of Mechanical Engineering, *School of Mechanical Engineering*, Georgia Institute of Technology, Atlanta, GA.

Gödel, K., 1931. “Über Formal Unentscheidbare Sätze der *Principia Mathematica* und Verwandter Systeme.” *Monatshefte für Math. u. Physik* Vol. 38, pp. 173-198.

Hall, P., 1994, *Innovation, Economics and Evolution - Theoretical Perspectives on Changing Technology in Economic Systems*. Hertfordshere England, Harvester Wheatsheaf.

Hazelrigg, G. A. (1999). “Validation of Engineering Design Alternative Selection Methods.” unpublished.

Hegel, G. W. F., [1817] 1959, *Encyclopedia of Philosophy*. New York, Philosophical Library.

Honderich, T., Ed., 1995, *The Oxford Companion to Philosophy*. New York, Oxford University Press.

Kant, I. [1781] 1933, *Critique of Pure Reason*. London, St. Martins Press.

Koch, P. N., 1997, *Hierarchical Modeling and Robust Synthesis for the Preliminary Design of Large Scale Complex Systems.* Ph.D. Dissertation, The George W. Woodruff School of Mechanical Engineering, Georgia Institute of Technology, Atlanta, GA.

Kuhn, T., [1962] 1970, *The Structure of Scientific Revolutions.* Chicago, University of Chicago Press.

Lewis, K., 1996, *An Algorithm for Integrated Subsystem Embodiment and System Synthesis.* The George W. Woodruff School of Mechanical Engineering, Georgia Institute of Technology, Atlanta, GA.

Locke, J., [1690] 1894, *An Essay Concerning Human Understanding.* Oxford, Clarendon Press.

Mistree, F., O. F. Hughes and B. A. Bras, 1993, "The Compromise Decision Support Problem and the Adaptive Linear Programming Algorithm." *Structural Optimization: Status and Promise.* M. P. K. (Ed.). Washington, D.C., pp. 247-289.

Pedersen, K., 1999, "Designing Platform Families: An Evolutionary Approach to Developing Engineering Systems." Ph.D. Dissertation, The George W. Woodruff School of Mechanical Engineering. Georgia Institute of Technology**,** Atlanta, GA.

Peplinski, J., 1997, *Enterprise Design: Extending Product Design to Include Manufacturing Process Design and Organization.* Ph.D. Dissertation, The George W. Woodruff School of Mechanical Engineering, Georgia Institute of Technology**,** Atlanta, GA.

Quine, W. v. O., Ed., 1953, *Two Dogmas of Empiricism. From a Logical Point of View.* Cambridge, MA, Harvard University Press.

Richardson, G. P. and A. L. Pugh, 1981, *Introduction to System Dynamics Modeling with DYNAMO.* Cambridge, MA, MIT Press.

Sellars, W., Ed., 1963, *Empiricism and the Philosophy of Mind.* Science, Perception and Reality. New York, Humanities Press.

Siddique, Z., 1999, *Common Platform Development: Designing for Product Variety,* Ph.D. Dissertation, The George W. Woodruff School of Mechanical Engineering, Georgia Institute of Technology, Atlanta, GA.

Silverberg, G., 1991, “Adoption and Diffusion of Technology as a Collective Evolutionary Process.” *Technological Forecasting and Social Change*, Vol. 39, pp. 67-80.

Silverberg, G., G. Dosi and L. Orsenigo, 1988, “Innovation, diversity, and diffusion: a self-organisation model.” *The Economic Journal* Vol. 98, pp. 1032-1354.

Simpson, T. W., 1998, *A Concept Exploration Method for Product Family Design.* Ph.D. Dissertation, The George W. Woodruff School of Mechanical Engineering, Georgia Institute of Technology, Atlanta, GA.

Sneath, P. H. A. and R. R. Sokal, 1973, *Numerical Taxonomy*. San Francisco, W.H. Freeman and Company.

Toulmin, S. (1977) "From Form to Function: Philosophy and History of Science in the 1950s and Now," *Daedalus,* Vol. 106, pp. 143-162.

Wittgenstein, L., [1921] 1961, *Tractatus Logico-Philosophicus*. London, Routledge and Kegan Paul.

Yin, R.K., 1994, *Case Study Research: Design and Methods*, 2nd edition, Sage Publications, Thousand Oaks, CA.