# What is this Science called Requirements Engineering?

Hans Akkermans
Free University Amsterdam VUA
Business Informatics (FEW/BI)
Amsterdam, The Netherlands
Hans.Akkermans@akmc.nl

Jaap Gordijn
Free University Amsterdam VUA
Business Informatics (FEW/BI)
Amsterdam, The Netherlands
gordijn@cs.vu.nl

## Abstract

*This vision paper reflects on the nature and status of Requirements Engineering as a science, based on recent RE journal articles on methodological foundations and research classification such as [18]. We put this discussion into the wider perspective of ongoing debates in other areas, including social and natural sciences methodology, IS and design research. We propose a new categorization of the notions of validity and evaluation, and suggest a criterial framework inspired by work on the structure of scientific argument.*

Keywords: methodology, evaluation, theory formation

## 1. Introduction

Quite a few recent papers are concerned with Requirements Engineering as a research discipline. They reflect in various ways upon its scientific, methodological and philosophical foundations (e.g. [18, 17, 7, 6, 4]). It is noteworthy that similar discussions take place in related areas such as (Management) Information Systems (IS) [9] and ontology engineering [8, 11].

Such discussions in information and computer systems research have characteristic recurring themes that one may summarize as follows:

- The proper scientific methodology to evaluate and validate produced claims to knowledge.
- The question to what extent the design of IT artefacts is (or is not) part of scientific research.
- The appropriate relationship between scientific research and societal practice in 'the field'.

In this short paper, our aim is to contribute to the ongoing discussion on two aspects. First, we believe that the discussion is enhanced if disciplines other than Computer Science are taken into account, including perspectives from social and natural sciences on empirical and theoretical research. Second, it seems to us that the concepts of evaluation and validation of research are currently used in an overly diffuse manner and so are in need of a more coherent and rigorous classification, applicable to RE- and IS-related research work.

## 2. Images of Science Influencing RE and IS

Wieringa, Maiden, Mead, and Rolland [18] propose a classification of RE research papers, which we henceforth will call the WMMR classification. It contains as major RE research categories: evaluation research, solution proposals, validation research, philosophical papers, and experience papers. Importantly, each research category has its own distinct evaluation criteria.

Evaluation and validation are clearly important aspects and criteria of scientific research. In the WMMR classification, evaluation is defined as the study of the properties of an RE problem or solution in practice, i.e., in a field setting. Validation is defined as the study of solution proposals (for example, new RE techniques or conceptual modelling methods) that have not (yet) been implemented or tested in field practice. WMMR explicitly state that evaluation and validation may use "both quantitative and qualitative research methods, ranging from controlled experiments to case study and action research" [18].

An interesting, partly implicit, implication of this view is a rather broad-church, inclusive position concerning what is to be accepted as scientific research. The WMMR classification embodies as a shared background two general elements: (i) scientific research is fundamentally concerned with knowledge claims; (ii) it is important that these claims are evaluated and validated in some way. The importance and indeed necessity of both elements is quite explicitly indicated; but both *what* these knowledge claims may refer to and *how* they may be justified in terms of method is circumscribed in a broad way.

To be clear upfront, the present authors are in favour of such a broad and inclusive position with regard to scientific research. However, we think it is useful to uncover and reflect upon the underlying assumptions, for two reasons. First, RE and IS research works at the borderlines of or in collaboration with several different scientific disciplines and fields of professional practice (more specifically, inside as well as outside of Computer Science). It is a common experience that other disciplines have specific (sometimes also narrower) views on what science is. Cross-boundary research activities have to be able to understand these views and relate to them. Second, different (seemingly 'philosophical') views on what science is or is not have important *operational consequences* for research as it is actually organized, carried out, and reviewed. Here, science is no different from business or industry: general strategic considerations ('this is the kind of business we are or want to be in') do influence the nature, style and management of actual operations.

The point we want to make here is that also in RE and IS there are different (possible and actual) positions concerning what our research discipline is. Before trying to achieve any consensus, it is important to clarify first what these different images of science are. As we will show, RE as a research field is influenced by and modelled after several different images of science, but their underlying assumptions and differences are often not clearly recognized. This blurs the foundations of the field.

### 2.1. Inclusive Models of Scientific Research

A first image of science is one that is broad and inclusive, roughly in line with the WMMR classification. A good working definition of this view on scientific research we propose is: *scientific research is a social practice concerned with the production of claims to knowledge through a process of inquiry in a way that is:*

1. *Relevant*: produced knowledge claims are to contribute and have specified added value to the current state of knowledge.

2. *Systematic*: the process of inquiry that leads to the production of knowledge claims is carried out in a critical and rigorous fashion.

3. *Transparent*: knowledge claims are produced and argued for in such a way that they are clear and open to critical scrutiny by others.

This definition of scientific research does state a number of specific epistemic desiderata (which are commonly found in various guises in review and evaluation criteria), but it is inclusive in the sense that it contains no specific ontological commitments regarding the possible or allowed subject matter, goals, nor methods of research.

### 2.2. Natural Science Models

There exist several other images of science, however, that are more specific than the one outlined above and that are now influencing RE and IS research. One influential view has its roots in the natural and 'exact' sciences, as concisely characterized in Figure 1.

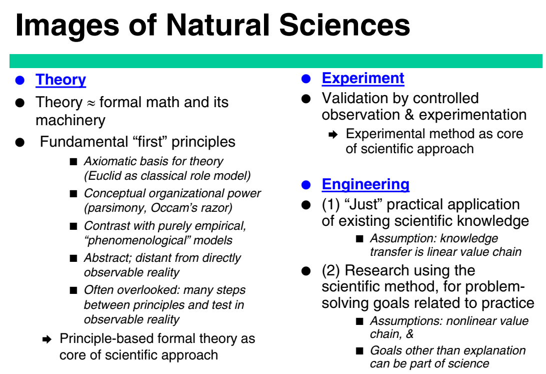


**Figure 1. Images of Natural Science.**

Upon closer inspection, the natural science model of research itself appears to be the co-existence of (at least) two distinct views on science. The first view, often labelled 'Baconian', is that controlled experimentation constitutes the core of the scientific method. The second view (that one may justifiably call, say, 'Einsteinian') is that the scientific method is first of all about uncovering the (abstract) fundamental 'first' principles as the basis for universal theories and laws.

These two views are not necessarily conflicting, but they do have to be recognized as clearly distinct. They operationally lead to emphasis on different kinds of research, ranging from purely theoretical to purely experimental laboratory research. This is also visible in the labour division within natural science in terms of departmental organization, and journal and conference structures. A consequence is a non-trivial distance between theory and experiment: validation of theories (e.g., in elementary particle physics or cosmology) involves a lot of bridging research that provides the construction of the knowledge how first principles ultimately are expressed or reflected in the observable real-world phenomena that we can test.

Although Computer Science is often seen as a branch of the exact sciences, there are quite some differences that especially hold for the RE field. Although mentioned by WMMR, controlled experimentation is not abundantly present in published RE research. Furthermore, conspiciously absent from the WMMR classification are theoretical research papers. This is quite remarkable since in many disciplines (e.g., physics, biology) theory development is recognized as an area in its own right, separate from experimental and empirical research. The 'philosophical papers' of the WMMR classification seem to go some way in this direction, but already the term suggests that there are quite some things still to be desired.

One difference is the nature of formalism that is typically employed: logic-based rather than mathematical. A deeper difference is that it is quite difficult in RE to state what the underlying first principles are. Sometimes this is attributed to the youth of the field as a science (an extreme version of this view is that it is not really scientific after all). We believe that there is a more fundamental issue here. First principles tend to be — and aim to be — universal and therefore context-free. In contrast, RE research is inherently involved in issues of system context, both in terms of technical and even more strongly in organization and business aspects of requirements. Context aspects should not be ignored nor can they easily be abstracted away.

On the other hand, theoretical research heavily depends on *conceptualization* preceding the expression of formal theories and mathematical equations. In science, conceptualization takes place in an only informal way. In contrast, RE and IS have over the years developed novel and rigorous model-based methods for conceptualization, which allow for computational and other forms of validation e.g. by CASE and simulation tools, but that are still much more understandable and accessible to practitioners compared to formal math and logic. Conceptual modelling and graphical diagramming methods have been elaborated into a fine art at a level of sophistication not found in other disciplines. We would argue that this is an important specific contribution to science: making conceptualization itself into something that is scientifically more rigorous.

### 2.3. Engineering: Is Design Research?

The above natural-science image also plays a significant role in the received view of engineering (including RE and Computer Science). Upon one view, also cited in the WMMR classification, engineering is the use of the scientific method (as outlined above) to practical problems. So, engineering is a form of science in which only the goal of research is different — problem solving rather than explanation and prediction — but the method is the same. In contrast, the more traditional view sees engineering as 'just the application' of existing scientific knowledge to practical problems (and so is accorded a lower scientific status). This is a view one encounters both in natural and social scientists, by the way.

The traditional view is logically defensible only if two assumptions are made: (i) knowledge transfer forms a linear value chain; (ii) the possible goals of scientific research are limited, typically to explanation and prediction. Both assumptions are in our view very dubious, and can historically and empirically be shown to be incorrect. The more defensible view is that engineering is a form of science (if done according to the above-stated general criteria of relevance, systematicity, and transparency), but with different goals, namely, related to problem solving and guidance of practice. Accordingly, design or 'solution proposal' studies (as the WMMR classification calls them) can normally be part of scientific research.

This is in agreement with the position taken by e.g. Simon [14], although the present authors believe that Simon's design science paradigm has important weaknesses. First, it does not really consider the issues of problem framing and analysis in context, or the 'why' questions that are key to RE. Second, it frames design itself computationally as search, centred around the question: What is the best next move? In design practice, search plays a minor role: for practical reasons, one even strives to avoid search as much as possible. Mature engineering fields have therefore developed prefab solutions for many types of problems that are then fitted to the problem at hand by forms of parametric or variant design. Software and knowledge engineering put significant efforts in the same direction, such as OO patterns, reusable knowledge templates, etc.

In the end, what makes design scientific is not the act of designing itself, just as it is not the act of writing that makes research papers scientific. It is the ensuing claims to new knowledge and their justification that makes design or solution proposals scientifically interesting. So, in design research it is mandatory that these knowledge claims are stated and argued for explicitly.

### 2.4. Social Science Models

There are still other images of science that directly influence RE and IS but that stem from social research, as summarized in Figure 2. A major distinction here is that between the 'quantitative' and 'qualitative' schools of methodological thought (for extensive and balanced overviews, good sources are Robson [13] and Bryman [3]). Other schools in social research and especially the humanities exist, notably the 'critical' and 'postmodern' schools of thought, but to date they have much less exerted a direct influence on RE research.

The quantitative school bases itself upon what is often called the natural science model of social research. Experimentation is seen as the gold standard of the scientific method, especially in its guise of the randomized controlled trial (RCT) and so-called quasi-experimentation. Characteristic for this school is the reductionist approach from com-

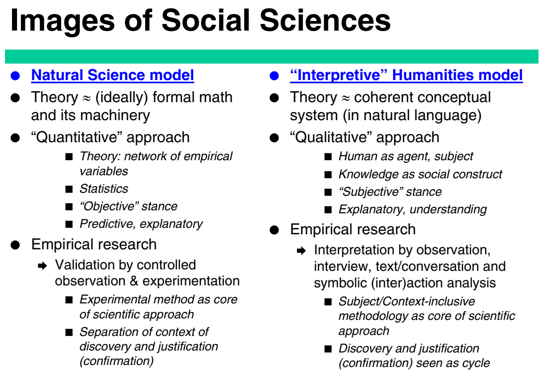


**Figure 2. Images of Social Science.**

plex conceptual frameworks and concepts ('constructs') to *variables* as its theoretical languange and, in evaluation and validation, its preference for statistical methods as the main route to empirical confirmation of hypotheses.

Information Systems research as published in journals such as the MIS Quarterly, an important forum for much business school related research on IT, is strongly influenced by this quantitative school. Due to the traditional emphasis on empirical and confirmatory modes of research, its mainstream is far from considering engineering and design issues as part of scientific research, and as a result shows no tendency to refer to results appearing in the Computer Science literature, and to propose solutions for the IT problems it considers. This explains why doing so [9] is perceived in the (M)IS field as kind of a recent breakthrough.

The qualitative school also has a focus on empirical research but emphasizes that (unlike the natural sciences) it is not so much the outside and context-free view of the researcher/observer that is important in explaining the world, but the context-inclusive interpretation and meaning that people themselves attach to their social world. Characteristic for the qualitative and interpretive approach are methods such as interview, focus group, observation, ethnography, action research, case study — methods also widely used in RE and IS research *and* practice.

In sum, RE and IS research is being influenced by several schools of methodological thought that make very different and sometimes even conflicting assumptions on what scientific research is. These assumptions do shape the general approach to three salient issues : (i) how evaluation and validation is done; (ii) the attitude towards engineering and design and, more generally, the relationship between research and practice; (iii) the nature and role accorded to conceptualization and theory. Therefore, we believe it is important for RE and IS to be aware of and clear in the underlying but often tacit assumptions it makes in researching, publishing and reviewing.

# 3. The Issue of Evaluation

Now, let us clarify this with some examples. In their classification paper, WMMR refer to Davis's and Hickey's suggestion to look at medicine as a model for how to do evaluation and validation research in a way that succeeds in establishing a clear relationship with practice in the field. Undoubtedly, there is much to learn from other disciplines. But it is quite another matter whether they are able to supply the proper research paradigm for RE and IS.

## 3.1. Varieties of Validity I

Evidence-based medical research tends to be quite strongly informed by the quantitative school, witness standard widely used textbooks such as [2]. Apart from the question whether this isn't too narrow as a general position, it is also questionable whether this would adequately cover the issues specific to RE and IS research and practice — including the IT-oriented engineering and design aspects [17, 9], the many conceptual model-based techniques en vogue such as UML [6, 10], or the (often qualitative) ways practitioners work with their customers [4, 12].

Our own experience with the $e^3$*value* approach in collaborative work with medical care research ([1], Ch. 8) suggests that this is not the case. Also the viewpoint article by Davis and Hickey themselves [4] is much less narrow in this respect: positioning "know thy customer" as a first rule, scenarios, human communication analysis, and "situational research" all sit very well with qualitative-interpretive, case study, and action research approaches. The key difference is the context inclusiveness of these approaches, in contrast to the context-free scientific ideal of empirical confirmationist/falsificationist research [19], visible in much evidence-based medicine but also much empirical MIS research.

Central theoretical principles underlying recent RE developments, such as the $i^*$ goal-oriented approach to software requirements, the $e^3$*value* approach to business context and requirements, and viewpoint-oriented requirements engineering *vis-à-vis* diverse stakeholder concerns, all point to the importance of *inherently* taking into account context in IS system development. Certainly if one appreciates a strong connection between RE and IS research and practice, ignoring or avoiding context is minimally a very high price to pay for scientific research.

## 3.2. Varieties of Experimentation

In the WMMR classification, evaluation and validation figure prominently. These terms are used in ways that differ from how they are used elsewhere, for example in social research, where in addition many varieties of validity are distinguished. All these conceptions of validity link to specific evaluative questions one has to ask concerning research designs and results, but they appear to depend on both the goals and the methods used by the research. As the teleology of RE and IS research differs from both social and natural sciences, the field has to think for itself to develop an appropriate characterization of experimentation and empirical validation.

Although it is customary to speak of *the* experimental method, there is a broad variety in experimentation relevant to RE and IS. A useful categorization we suggest is:

(1) *Logico-mathematical demonstration ('proof')*: if a theory is sufficiently rigorously specified, certain desired properties may be strictly mathematically or logically derived. This is seen in natural sciences or economics (e.g. tendency to unique equilibrium), but is also relevant say in safety-critical IS (e.g. deadlock states are impossible to reach).

(2) *Thought experiment*: describes an idealized but not unrealistic pseudo-experimental scenario and then situationally explores the implications of conceptualizations and theories and their coherence (the debate between Einstein and Bohr on the completeness of quantum mechanics is a good example, see the Physical Review of 1936).

(3) *Computational simulation and analysis*: might be considered as the next stage of the thought experiment, whereby the computer has made it possible to run very high numbers of experiment scenarios and explore a large parameter space (a computer use widely found in natural sciences, engineering and social system simulation).

(4) *Laboratory experiment*: the kind of experiment standard in natural sciences but also quantitative social research, based on clearly specified (e.g. causal) hypotheses, strictly controlled conditions, and minimized external context influence so as to come to unambiguous knowledge claims.

(5) *Field experiment*: pilots of interventions (solution proposals), familiar from engineering but also medicine, education, social evaluation research, where one attempts to demonstrate the usefulness of an intervention under realistic (context-inclusive) field conditions. Yin, a foremost author on the case study method [19] positions case study as such a form of field experiment.

(6) *Practice/experience-oriented field study*: if certain practices are established for a longer period, they can be observed and studied as empirical phenomena both via quantitative and qualitative modes of research (e.g. survey, observation, interview, ethnography).

WMMR's 'validation research' is in fact a single umbrella term for all different forms (1)-(3), whereas its 'evaluation research' collapses all forms (4)-(6) into one.

## 3.3. Varieties of Validity II

Given this strong variety of experimental forms, it doesn't come as a surprise to find a similar wide range of differing notions concerning the validity aspects of knowledge claims. In quantitative research, we encounter construct, convergent and divergent validity etc., which are absent in qualitative research where we find instead, say, descriptive and interpretive validity.

Campbell, reputed for his work on experimental methodology in the social sciences, makes some interesting observations here (see his Foreword to Yin's case study methodology book [19]). He states that it is not experimentation *per se* that makes up the core of the scientific method, but rather the *systematic treatment of plausible rival hypotheses, by investigating and testing the network of implications of the theories* from which hypotheses under scrutiny originate. Thus, theories are tested rather than just individual hypotheses ("ramification extinction"). From this perspective, it becomes clear why all above-mentioned approaches (1)-(6) are to be accepted as suitable scientific test methods for knowledge claims.

A step further we make here is the observation that validation of knowledge claims — whether theories, hypotheses, new conceptual methods, computational techniques, practice guidelines, or artefact designs — always boils down to the *construction of rational communicative argument* that must be defended and made credible. In scientific work, available empirical data and theory are systematically brought together such that knowledge claims follow in a step-by-step transparent process of rational reasoning. Therefore, we suggest that a more unifying framework for the variety of validity notions can be found in considering the general structure of argument.

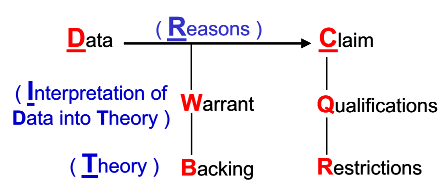


**Figure 3. The structure of argument.**

The general structure and rules of argument have been investigated in past years in philosophy [15], communication theory [16], and critical thinking and informal logic [5]. A simple model of the layout of arguments due to Toulmin [15] is suitable for our present purposes, and depicted in Figure 3. This model suggests a categorization of validity notions, with a checklist of review criteria:

(1) *Descriptive validity* ***D***: do the data provide a truthful description of the situation or problem that is considered?

(2) *Theoretical validity* $\boldsymbol{T}$: are the employed theories or conceptual frameworks explicated and shown to be appropriate for the purpose?

(3) *Interpretive validity* $\boldsymbol{I}$: is the way in which all available data are mapped onto or interpreted in the employed theories or frameworks clear and adequate?

(4) *Reasoning validity* $\boldsymbol{R}$: are all steps in the reasoning sound and, in addition, consistent and coherent with other knowledge that we possess?

(5) *Internal validity* $\boldsymbol{C}_{int}$: are the claims made acceptable 'beyond reasonable doubt' *within* the situation or context (or sample) considered in the study?

(6) *External validity* $\boldsymbol{C}_{ext}$: are any generalization claims that go *beyond* the studied situation sufficiently credible?

Clearly, not all research papers need to satisfy all validity criteria. For example, for the 'personal experience papers' mentioned in the WMMR classification the main criterion will be their descriptive validity if they are to be a basis for further theoretical and/or empirical research work.

## 4. Concluding Remark

Karl Marx once commented that those who are not aware of the course of history are doomed to repeat it. RE and IS should be able to avoid this trap. They can learn a lot from experience, history, and philosophy of other disciplines and associated images of science, as we have sketched in this paper. Whether or not this was the intention of Marx's comment, however, once one knows the history, in the end one has always to think and decide for him/herself.

We have therefore attempted to clarify the foundations of the field in comparison with other disciplines. Investigating context aspects, in other words the 'why' questions, is essential to RE research, both in theory formation and empirical validation. There is a long way to go here, but an encouraging result is that rigorous methods for conceptualization have been developed in the field — and they have shown some capabilities to deal with such matters. Criteria to judge scientific progress in RE and IS also must take into account its specific goals and setting: evaluation, validation, experimentation, and conceptualization all assume their own specific forms. This includes (though not exclusively) real-world domains and practice beyond solely academic concerns.

**Acknowledgments.** This work was partially supported by the BSIK Freeband/FRUX project, the VU* VUBIS centre at VUA, and the European Network of Excellence INTEROP. We are grateful for the many interesting reactions by the Semantic Web and GREETING business informatics meeting participants in Amsterdam, where the present views were discussed by the first author.

## References

[1] Z. Baida. *Software-Aided Service Bundling — Intelligent Methods and Tools for Graphical Service Modeling*. PhD thesis, Free University Amsterdam VUA, Amsterdam, NL, 2006. Also Freeband/FRUX BSIK project deliverable.

[2] A. Bowling. *Research Methods in Health: Investigating Health and Health Services*. Open University Press, Maidenhead, Berkshire, UK, 2002. Second Edition.

[3] A. Bryman. *Research Methods and Organization Studies*. Routledge, London, UK, 1989.

[4] A. Davis and A. Hickey. Requirements researchers: Do we practice what we preach? *Requirements Engineering*, 7(2):107–111, 2002.

[5] A. Fisher. *The Logic of Real Arguments*. Cambridge University Press, Cambridge, UK, 2004. Second Edition.

[6] A. Gemino and Y. Wand. A framework for empirical evaluation of conceptual modeling techniques. *Requirements Engineering*, 9:248–260, 2004.

[7] J. Gordijn, V. Kartseva, J. Schildwacht, R. Wieringa, and J. Akkermans. Developing a domain-specific cross-organizational requirements engineering method. In M. Aoyama, M. Saeki, and N. Maiden, editors, *Proceedings 12th IEEE International Conference on Requirements Engineering (RE04)*, pages 134–143. IEEE Computer Society, Washington, DC, September 2004.

[8] N. Guarino and M. Musen. Applied ontology: Focusing on content. *Applied Ontology*, 1(1):1–5, 2005.

[9] A. Hevner, S. March, J. Park, and S. Ram. Design science research in information systems. *MIS Quarterly*, 28(1):75–105, 2004.

[10] T. Menzies. Model-based requirements engineering. *Requirements Engineering*, 8:193–194, 2003.

[11] P. Mika and J. Akkermans. Towards a new synthesis of ontology technology and knowledge management. *The Knowledge Engineering Review*, 19(4):317–345, 2004. DOI Online 11 November 2005.

[12] R. Offen. Domain understanding is the key to successful system development. *Requirements Engineering*, 7:172–175, 2002.

[13] C. Robson. *Real World Research*. Blackwell Publishers, Oxford, UK, 2002. Second Edition.

[14] H. Simon. *The Sciences of the Artificial*. The MIT Press, Cambridge, MA, 1996. Third Edition.

[15] S. Toulmin. *The Uses of Argument*. Cambridge University Press, Cambridge, UK, 1958. Updated Edition 2003.

[16] F. van Eemeren and R. Grootendorst. *A Systematic Theory of Argumentation — The Pragma-Dialectical Approach*. Cambridge University Press, Cambridge, UK, 2004.

[17] R. Wieringa. Requirements researchers: Are we really doing research? *Requirements Engineering*, 10:304–306, 2005.

[18] R. Wieringa, N. Maiden, N. Mead, and C. Rolland. Requirements engineering paper classification and evaluation criteria: A proposal and a discussion. *Requirements Engineering*, 11(1):102–107, 2006.

[19] R. Yin. *Case Study Research — Design and Methods*. SAGE Publications, Thousand Oaks, CA, 2003. Third Edition.