

Evidence-Based Structuring and Evaluation of Empirical Research in Requirements Engineering

Fundamentals, Framework, Research Map

Empirical knowledge on the appropriateness and desirability of competing methods and techniques of requirements engineering is rare. At the same time, existing knowledge has been achieved by different research methods which have not yet been consolidated. The paper structures empirical knowledge, evaluates it, and represents it in a research map. It will be demonstrated how this approach allows empirical knowledge to be more consistently transferred into practice. At the same time, we will identify main focuses and gaps in research in the field of requirements engineering.

DOI 10.1007/s12599-010-0101-0

The Authors

Prof. Dr. Matthias Goeken (JP) (✉)
Dipl.-Kfm. Janusch Patas
 IT-Governance-Practice-Network
 Frankfurt School of Finance &
 Management
 Sonnemannstraße 9-11
 60314 Frankfurt am Main
 Germany
m.goeken@frankfurt-school.de
j.patas@frankfurt-school.de

Received: 2009-08-02
 Accepted: 2010-02-10
 Accepted after one revision by the editors of the special focus.
 Published online: 2010-04-23

This article is also available in German in print and via <http://www.wirtschaftsinformatik.de>: Goeken M, Patas J (2010) Evidenzbasierte Strukturierung und Bewertung empirischer Forschung im Requirements Engineering. Grundlagen, Ordnungsrahmen, Forschungslandkarte. WIRTSCHAFTSINFORMATIK. doi: 10.1007/s11576-010-0223-5.

© Gabler Verlag 2010

1 Introduction

Research and literature on Requirements Engineering (RE) generally consist of conceptual publications based on design-

oriented methods and thus applicable to the design science paradigm (Wieringa and Heerkens 2006, p. 305). In recent years, many different approaches have been developed in RE. They are partially based on different principles and definitions of terms and provide for different phases of the RE process (e.g., Hickey and Davis 2003b; Kotonya and Sommerville 1998; Pohl 1997). Apart from more or less comprehensive approaches, a large number of methods, techniques, languages, and tools (hereinafter referred to as “artifacts” in terms of Hevner et al. 2004) has been developed each of which are designed to support specific tasks of the respective REs (Ivarsson and Gorschek 2009, pp. 172–173).

This diversity involves practical problems and thus encourages a scientific discussion. In RE practice, the lack of knowledge about the appropriate use of artifacts in the respective case turns out to be problematic if, for example, techniques for the collection of requirements are chosen intuitively or based on the requirement engineer’s experience (Rupp et al. 2009, p. 98). Contingency models, while available for the choice of techniques, are essentially based on considerations of plausibility (Davis 1982; Goeken 2006; Macaulay 1996). It is safe to assume that decisions on a more appropriate basis can clearly enhance the success of RE and thus of software and systems development (Hickey and Davis 2003a). This also applies not only to the choice of techniques, but equally to the use of other artifacts such as methods, languages, or tools.

Following the approaches of “empirical software engineering research” (Jarke 2009, p. 73) it is assumed in this paper that empirical research can make a useful contribution towards the foundation of decisions in RE and towards dealing with the variety of artifacts, the aim being to empirically underpin design-oriented research results and to consistently transfer empirically validated knowledge into practice (Juristo and Moreno 2001; Wohlin et al. 2000).

The approach taken herein consists of a structuring and evaluation of the existing empirical research on RE as found in literature on a scientific basis while preparing empirical knowledge in a way that allows “evident knowledge” to successfully be transferred into practical system development. Further, this supports the progress of scientific knowledge since empirical knowledge can be analyzed as to its main focuses, and gaps in research can be identified.

The foundations and the methodical research approach for this contribution will be outlined in the following section with parallels to evidence-based medicine (EbM), the basic ideas of which serve as a starting point for our approach. In the following Sect. 3, a framework for the structuring of the empirical results of RE will be developed. Sect. 4 presents the search strategy empirical studies in RE are detected with, explaining the structuring and evaluation of research results by means of a so-called research map for empirical RE research (in the following shortly referred to as “research map”). Subsequently, we will discuss the results,

outlining the need for further research as well as the planned enhancement of the approach in the following section, “Conclusion and Future Work”.

2 Foundations and Methodical Research Approach

RE is considered an essential component of the process of software or systems development (Pohl 2008, p. 30). According to Kotonya and Sommerville (1998, p. 8), RE covers all activities including the discovery, documentation, and maintenance of requirements for a computer-based system. Directed towards a direct problem solution, design-oriented research usually predominates; hence the constructed artifacts are frequently not or insufficiently validated and evaluated (Wieringa and Heerkens 2004; Zerkowitz and Wallace 1997, p. 742). This lack of empirical research leads to confusion as to whether a constructed artifact contributes to the solution of a problem on the one hand and, on the other hand, whether it is suitable for practical use. Comparative studies on alternative and competing artifacts that go beyond evaluation are even more difficult to detect.

A similar challenge was increasingly posed to medicine in the 1990s, namely to have its procedures and the large amount of (action-based) knowledge more rigorously validated and to justify decisions in a verifiable way. This led to the development of EbM as an approach to evidence acquisition and improvement of medical practice. The approach effected a significant “leap in quality” and a re-evaluation of existing knowledge, basically assuming that the systematic preparation of empirical research would enhance the success of treatments (Sackett et al. 1996). Given the evidence-based approach, it is no longer individual experience, medical authority, deduction, or biological plausibility that exclusively determine the “state of the art of medical practice”, but also “the critical appraisal of empirical evidence compiled in scientific studies” (Raspe 1996, p. 560; translated into English).

While EbM practice is focused on healthcare decisions based on a “conscientious, explicit, and judicious use of current best [scientific] evidence” (Sackett et al. 1996, p. 71), the character of EbM as “a method for acquiring and consolidating knowledge that sets up clearly defined criteria for the evaluation of scientific studies and their synthesis” (Bayerl et

al. 2009, p. 120; translated into English) is elsewhere emphasized. It aims to reassess the situation of research and medical procedures by means of evidence criteria (Bayerl et al. 2009, p. 121) and subsequently “to develop, provide, and distribute systematic reviews” (Raspe 1996, p. 558; translated into English).

For a critical appraisal of studies in terms of an evaluation, in EbM classification systems are used which assume a different “rating” (due to higher reliability and representativeness) for different study designs (an overview can be found in Cochrane n.d.). Despite the differences in detail, systematic reviews (meta analyses) of double-blind randomized studies are usually accepted as a “gold standard”, possessing the strongest conclusiveness and thus the highest rating. In contrast, “opinions and beliefs of respected authorities” possess less evidential value.

Critical for the success of evidence-based research, the existence of functional and technical infrastructures allows the structuring of results of existing studies, efficiently and effectively supporting the search for relevant evidence (Raspe 1996, p. 558).

Based on the fundamental idea of EbM as a “method for acquiring and consolidating knowledge”, the analogy presented here starts with the design of a functional infrastructure.

Extensive and widely spread classifications, nomenclatures, and thesauri are available in medicine (e.g., DIMDI 2009), allowing the development of a functional and technical infrastructure. In the fields of computer science, business and information systems engineering (BISE), and especially RE, in contrast, there are hardly any comparable, generally accepted classifications of the subject matter that make it possible to structure empirical results and to create systematic reviews. Therefore, this contribution first concentrates on the construction of a framework for the RE, referring to the “knowledge base” (Hevner et al. 2004), i.e. to existing and proven approaches. We will apply relevant components of these approaches that are represented in a conceptual metamodel (Fig. 1).

By means of this framework, we will structure the empirical studies identified

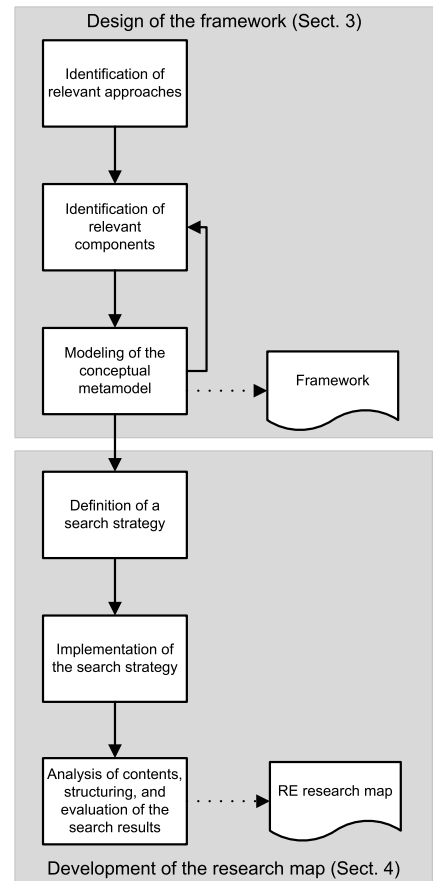


Fig. 1 Research process of this contribution

by a comprehensible search strategy in Sect. 4. This is done by assigning the variables of the studies to the components of the framework. For an evaluation of the structured studies, we refer to a classification system derived from Cochrane (n.d.) and Wilde and Hess (2007) in which the usual methods of empirical research in BISE are ranked according to evidence levels which are geared to the ranking order in medicine (Table 1).

Systematic reviews have gained particular importance in EbM as they accumulate evidence from various individual studies (Cochrane n.d.; Maier and Möller 2007, p. 1028). However, the accumulation of individual evidence involves certain difficulties concerning methodology as well as contents, in particular when statistical methods, such as in meta-analyses, are used. Maier and Möller (2007) distinguish meta-analyses and “narrative reviews”, which represent a weaker way of accumulation of existing evidence in terms of qualitative conclusions. Section 4 of this contribution is

aimed at providing a qualitative review in which the results will be additionally visualized in a research map. (1) The framework must comprise the essential components of RE.

Table 1 Classification system for evaluating empirical work in RE

Level	Type of evidence
Ia	Meta-analyses
Ib	Narrative reviews
II	Qualitative/quantitative cross-sectional work
IIIa	Case studies
IIIb	Experiments (field experiment, laboratory experiment)
IV	Expert interviews

3 Construction of a Metamodel as a Framework

3.1 Preliminaries

The first steps of the research process aim at constructing a framework that represents the field of RE. These steps are a design-science-oriented preparation and therefore based on proposals for its implementation and evaluation.

While Hevner et al. (2004) formulate more general requirements for design science (Frank 2007, pp. 168–169), concrete instruction for a stricter approach can be found in Gehlert et al. (2009). They argue that the researcher should be guided by theories, firstly in order to increase the intersubjectivity of the artifact, and secondly in order to enhance the quality of the artifact to be constructed. In case of a lack of theories for the support of decisions in the design process, Gehlert et al. (2009, p. 449) claim “to explicate his or her background knowledge as hypotheses”. They propose to specify the objective of the research process in requirements and to operationalize these by means of indicators in order to facilitate the traceability of design decisions and to be able to assess the artifact.

Similarly, Frank (2007, pp. 179–181) argues that the objectives, requirements, assumptions, and design decisions are to be explicated.

Due to a lack of theories in the field of RE that could guide the design process as defined by Gehlert et al. (2009), the clarification of the aforementioned aspects remains in order to establish transparency within the design process: In this article, we pursue the aim of “constructing a framework that appropriately represents the scope of RE”. This objective will be decomposed into the following requirements:

- (2) Relevant relations between the components must exist within the framework.
- (3) The relations between the individual components must be evaluable within the framework.

The usefulness of the framework results from the possibility of structuring empirical studies based on which empirical evidence can be integrally represented and meaningful comparative analyses can be created.

Therefore and because of the first requirement, the degree of specialization of the framework is considered significant. Thus, it has to be decided whether it should rather consist of a variety of specialized components or – on the contrary – of generic components. The statements in Sect. 3.3 are based on the assumption that a rather generic framework with the corresponding generic components is appropriate for the integrated representation and evaluation of empirical evidence. Among other reasons, it also has to be generic in order to provide a structure for past and future research.

Related to this, it is assumed for the design of the framework that conceptual metamodels constitute suitable artifacts for representation. By use of metamodels, the relevant components and relationships can be described on a generic level, which has proven to be appropriate for comparisons and analyses (Kühne 2006; Strahringer 1996). In addition, metamodeling allows a consistent presentation of different formalisms so that different approaches can be integrated (Alter and Goeken 2009). This seems purposeful as the framework – as recommended by Hevner et al. (2004) – is derived from existing RE approaches and other relevant approaches. The following Sect. 3.2 explains which approaches are regarded as relevant and appropriate.

Referring to current approaches of both RE and systems development, we

first aim at achieving a consistent embedding of the framework into accepted knowledge according to the coherence theory of truth. Secondly, following Frank (2007, p. 175), we assume an at least partly existing consensus (consensus theory of truth).

Nevertheless, the framework does not claim to be a universal reference model of RE. As design-oriented preliminary work, it intends to offer a “satisfactory solution” (Hevner et al. 2004 referring to H.A. Simon) for the abovementioned objectives and requirements, and must thus be regarded as provisional knowledge according to Popper (1959). For this reason, it is generally considered extendable. An extension may result from both empirical and conceptual reasons. If, for example, future empirical studies investigate variables that cannot be structured by means of the framework, it will have to be extended by appropriate components. A conceptually motivated extension is recommended for an identification of new research questions and research needs.

3.2 Identification of Relevant Approaches

In the following, we will aim to achieve coherence with existing knowledge and at least a partial consensus on the framework by following approaches which have attained a certain universality and dissemination. The latter is assumed if they are referred to as “reference models” or “frameworks”, if they are published in textbooks, if they evolved from standardization projects, or are established as (quasi-)standards while being used e.g., for certification. In addition, we will refer to method engineering as a field of research which explicitly deals with the description of structures of methods.

We will analyze method engineering, unified method architecture (UMA) (Shuja and Krebs 2008), the Software Engineering Body of Knowledge (SWE-BOK) (Abran et al. 2004), the requirements engineering framework according to Pohl (2008), the requirements engineering reference model (REM) by Broy et al. (2007), the ReqMan process framework of IESE (Dörr et al. 2006), and the V-Modell XT (V-Modell XT 2006) regarding their possible contribution to the construction of the framework.

As a generic framework is to be developed at this point, it will disregard specific approaches such as VOLERE

(Robertson and Robertson 2006), RUP (Kruchten 1999), CMMI (SEI 2007), or RM&E. Unlike those approaches, the framework does not define specific components for RE (such as a “requirements specification” which integrates accepted standards subsequent to “documentation” and “quality control”), but generic components for the RE on a meta level (e.g., “RE results” as a result of “RE activities”; **Table 2**).

The abovementioned approaches have been chosen as they comprise a structure of RE-specific knowledge that is superior to specific approaches or include an appropriate structure. SWEBOK offers a “Guide to the Software Engineering Body of Knowledge”, which dedicates a separate chapter to RE. V-Modell XT, CMMI, and REM contain an explicit metamodel that provides a structuring of the relevant components. The requirements engineering framework was chosen for its inclusion of a large part of the fundamental knowledge for certification under the standard IREB as a Certified Professional for Requirements Engineering

(CPRE; Pohl and Rupp 2009). By certification it may achieve the role of a quasi standard for obtaining RE.

Not directly related to RE, the approaches of method engineering and UMA are deliberately generically designed. As their objective is the description of methods and their components, the development of the framework can be based on these explicit structural descriptions. They also serve as a starting point to ensure that the framework is sufficiently generic according to the considerations of expediency stated above.

Obviously, there are a number of other approaches in the literature that could be used additionally or alternatively for the construction of the framework. The detailed analysis of the approaches in the following section shows, however, that saturation is already achieved with these approaches. Comparable with the stopping rules in RE (Browne and Pitts 2004), this concept from qualitative research describes a state in which new conclusions cannot be obtained by further

approaches (Bryant and Charmaz 2007, p. 611).

3.3 Presentation of Approaches and Identification of Relevant Components

The basis for the description, development, and comparison of methods in *method engineering* is constituted by a common description model (Heym 1993). Subsequently, we will consider the approach by Brinkkemper as well as approaches according to St. Gallen and Karlsson in order to identify model components for the framework (a broad overview of other method engineering approaches can be found in Becker et al. 2007).

Brinkkemper's approach provides a framework for method engineering to structure the terminology and the fragmented knowledge of application development (Brinkkemper 1996). Relevant components according to this approach are techniques (consisting of procedures

Table 2 Potential components identified for the framework

	Technique	Result	Role	Activity	Purpose	Notation	Principle	Tool	Context
Generic method descriptions	Brinkkemper	=	= Development products	=	=	=	= Directions & rules	=	
	St. Gallen	=	=	=	=				
	Karlsson	Part of activity	Artifact	=	Action	=	=		Business context
	UMA	14 Guidance elements	3 different work products; Milestones	=	Activities; Task Step		Key principles (from RUP)	=	
Generic methods and models of RE	RE-Framework		Requirements artifact		Core & cross-sectional activities				System context
	REM	Methods	Artifact; Milestones	=	= Process definition				
	ReqMan	=	=		Activities; Practices				
	V-Modell XT	Methods	Decision points & products	=	=	=		=	
Specific RE approaches	VOLERE	=	Deliverables	Stakeholder	=				
	RUP		Artifact	Discipline; workflow details; activity step;		Vision	Vision		
	CMMI		Typical work products		Process area	Goal			

and notations), tools, and methods (directions and rules, development activities and development products).

According to the method engineering approach from *St. Gallen*, a method represents the guideline for the entire development process from the requirements to the implemented design result (Heym 1993). The basis for the description and development of methods is a consistent description model consisting of five components (Gutzwiller 1994). These are activities that produce a defined result. The providers of the activities are roles. Techniques, in contrast, describe how to create an activity as a result. Finally, the component metamodel structures the results achieved in the *St. Gallen* approach as a conceptual data model.

Focusing on the activity component (prescribed action), the approach of method engineering according to Karlsson (2002) presents a concept similar to the *St. Gallen* approach. Roles can be found therein as well. In relation to a specific actor and the action, the context is also included. By use of a notation, results (artifacts) that are incorporated into the concepts are created by means of activities that need to have a purpose.

UMA is a metamodel by IBM for the abstract description of process models, including those of RUP, which was issued by the OMG as SPEM 2.0. It contains very specialized components, in particular 14 so-called “guidance elements” and several “work products”. At this point, these are understood as techniques or results, as are “milestones”. In addition, *UMA* contains several so-called “categories” (discipline, domain, role set, and tool category) which arrange the other components into groups; however, these groups are not included in the framework.

The objective of the *SWEBOK project* is to organize the knowledge from the field of software engineering which is generally accepted and to provide access to this knowledge; an approach similar to the objective pursued herein. However, in *SWEBOK* the existing knowledge is consistently organized in a hierarchic way by forming “knowledge areas” (e.g., requirements engineering), “sub areas” (e.g., requirements process and requirements elicitation), and “topics” (e.g., elicitation techniques). Therefore, *SWEBOK* cannot be used directly for the derivation of the framework since it does not deal with generic structural knowledge

but conducts too general a hierarchical structuring of knowledge.

The *requirements engineering framework* by Pohl (2008) is about the specification of structural elements of the RE process, which are system context, core activities, requirements artifacts, and cross-sectional activities. The system context is divided into four facets. The three core activities consist of elicitation, negotiation, and documentation. They form a possible separation of phases of the RE process. The defined requirement artifacts goals, scenarios, and solution-oriented requirements provide various forms of the component results. Finally, the cross-sectional activities (validation and management) represent nothing else than further instances of the component activity.

REM “is a first step towards a synthesis between formal techniques, in particular modeling techniques and a broad-based methodical approach to requirements engineering” (Broy et al. 2007, p. 134; translated into English. For a detailed overview of the *REM* model see Geisberger et al. 2006). The authors focus on development results (artifacts) and notations to represent these results to a greater extent than in the aforementioned framework. The *REM* describes the structure and results of RE in the artifact model and its tailoring. In addition, it contains artifact-based process definitions and describes milestones. The RE artifacts are divided into business needs, requirements specification, and system specification. Levels of completeness and quality serve as milestones in the process of product development. The RE artifacts are created with the help of methods and by assignment of roles through activities. The result is a specification document which defines a milestone (decision gate; Geisberger et al. 2006).

The *ReqMan process framework* of IESE is a “pragmatic requirements engineering building set for SMEs” (Dörr et al. 2006, p. 7) which aims to “describe the main tasks and practices of requirements engineering and ... to collect related techniques” (Dörr et al. 2006, p. 1; translated into English). In the framework, activities (synonymous: process phases), practices, and techniques can be distinguished. Activities and practices describe the “what”, i.e. what is done, while techniques support the “how”. On the project website (<http://www.re-wissen.de>), however, utilities and tools are distinguished.

The latter are consistent with the understanding underlying herein. Utilities are, among others, “templates and instructions that help to correctly collect and document requirements” (translated into English). The framework includes “utilities”. Templates are techniques that support the production of results. “Instructions” are to be understood as activities and techniques in accordance with the abovementioned sense.

Much like the *REM*, the *V-Modell XT* is artifact-oriented (product-oriented). Placing a further point of emphasis on the “risk-minimizing project control” by including “decision gates” (similar to milestones), it deals with the division of accountabilities and responsibilities more clearly than the other approaches do within an extensive role model. There are different types of projects with different project type variants, which, in turn, determine the project’s implementation strategy. Depending on this, process modules are selected. These, in turn, encapsulate activities and products with responsibility for specific roles. An activity determines how a product has to be provided, applying various methods and tools that are specified in the *V-Modell XT*. The respective project implementation strategy determines specific decision gates where, e.g., decisions on the continuation of the project are made. The decision gates include one or more products, the completion of which documents the project’s progress (*V-Modell XT* 2006). Components relevant to the project organization and tailoring are neglected in the further course.

Furthermore, relevant components from *VOLERE*, *RUP*, and *CMMI* can be assigned to the previously identified components (Table 2). It turns out that the generic methods and models of RE as well as the specific RE approaches contain specialized components. However, they do not add any components not already given by the generic method descriptions. Thus, the criterion of saturation seems to be satisfied.

3.4 The RE Framework as a Conceptual Metamodel

The consolidation into an RE metamodel is based on the previous presentation and analysis of relevant components of the approaches. In the following, the relevant components identified in the publications discussed are extracted for the

framework (Table 2). By this comparison, it becomes obvious that in some cases the presented approaches differ significantly in terms of the components provided. Therefore, it is necessary to make choices and explicitly and unambiguously describe and model the importance of the individual components and their interactions.

RE activities are defined cyclically, i.e. sub-activities are activities in themselves. Consequently, the entire RE process can be seen as an activity, as can a phase, sub-phase, or a step. In addition, they can be in a sequence, and thus in defined, e.g., sequential, relationships. RE activities take place in a certain context which can be characterized further by means of the various facets by Pohl (2008). Furthermore, as seen in the model by Karlsson (2002), RE activities assume a goal (or a purpose). Activities and their arrangement are influenced by principles, such as user participation, iterative, or agile approaches, etc. Additionally, activities can be supported by tools.

The components purpose and vision, however, are often subsumed by goal (generalization). This component is applied in few approaches; however, it is at least implicitly provided in a range of approaches (Table 2).

The core of the framework is the ATRN relationship (ATRN-Rel.) which clarifies the interaction of RE activities, RE techniques, RE roles, and notations. This relationship type composed of four components is conceptualized as an entity type and can itself be a starting point for a relation – namely, the RE result. This shows that the development of results in RE by an activity is carried out by use of a language. This is done by means of a specific role and is supported by technology. The relationship “activity uses development result” remains unaffected by the ATRN-Rel.

Concepts (Karlsson 2002) are not included. Referring to extralinguistic phenomena, they provide a reference to the real world in the development. They have significance especially for conceptual modeling. However, they will not be

considered as a component of the framework at this point. Milestones are defined as the presence of RE results, which in turn can be aggregated or decomposed into individual partial results. This is characterized by the recursive relationship (Fig. 2).

4 Development of the Research Map

4.1 Definition and Implementation of the Search Strategy

A critical requirement for a narrative review as representative as possible is a transparent and systematic literature review based on a previously defined search strategy. For this purpose, the primary steps by Kitchenham et al. (2004), who provide a five-stage approach, and Brereton et al. (2007), who give detailed guidelines, serve as the basis in this paper.

The first fundamental prerequisite is to establish a research question, which is necessary for the implementation of

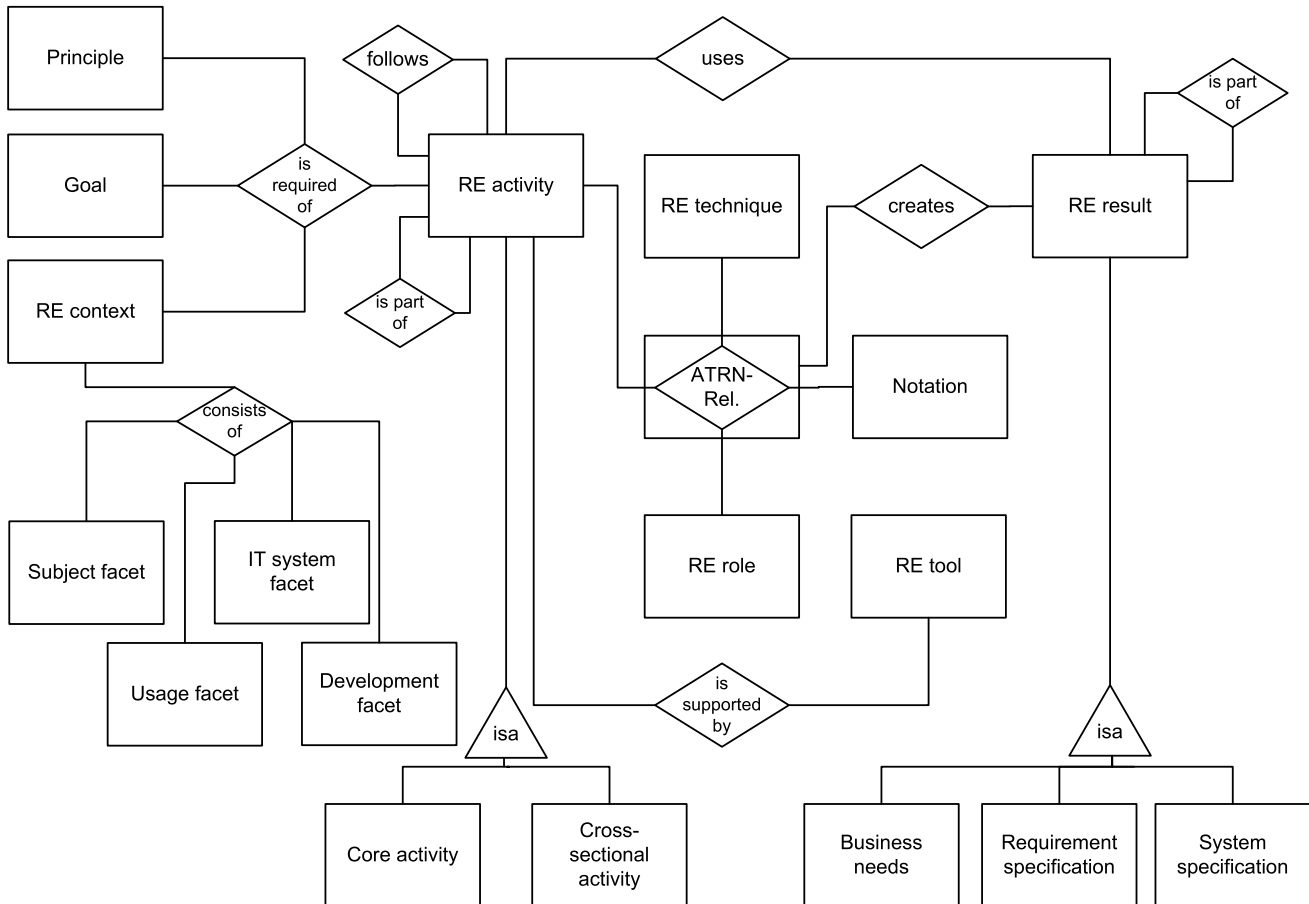


Fig. 2 Framework to structure empirical research in RE

Table 3 Composition of responses after carrying out the search strategy

Database	Composition of the search results	Requirements 'AND' Empirical	Requirements 'AND' Case Study	Requirements 'AND' Survey	Requirements 'AND' Experiment	Requirements 'AND' Field Study	Requirements 'AND' Meta Analysis
ScienceDirect	Total responses	196	468	284	626	9	2
	Preselection	13	25	6	17	1	0
	Final selection	9	15	0	11	0	0
SpringerLink (REJ)	Total responses	71	99	52	30	6	0
	Preselection	25	13	3	2	1	0
	Final selection	24	10	3	2	0	0
IEEE	Total responses	100	100	100	100	100	40
	Preselection	10	13	8	17	12	0
	Final selection	10	8	6	11	7	0
EBSCOHost	Total responses	1346	1109	1326	526	48	10
	Preselection	25	32	8	20	2	0
	Final selection	11	17	1	7	2	0

the search strategy and which narrows down the field to be considered. For this purpose, we derive the following question from the initially defined objectives: “What kind of empirical knowledge exists for requirements engineering?” This question must now be broken down into keywords. The decomposition represents a critical step since the keywords determine the quality of the search results (vom Brocke et al. 2009). The search terms used herein were “requirements”, “empirical”, “meta analysis” as well as some behavioral research methods according to Wilde and Hess (2007). The next step is to determine appropriate publication databases that serve as sources. According to Dybå et al. (2005, p. 61) and Dieste et al. (2009, pp. 518–521), the IEEE Computer Society Digital Library, EBSCOhost, ScienceDirect, and SpringerLink (Requirements Engineering Journal only) were selected.

The search was limited to abstract, title, and keywords for ScienceDirect. All other databases were searched with the settings “all fields” since these databases do not allow a more precise configuration within a search query. Moreover, IEEE is not able to deliver more than 100 hits per search. A total of 6,748 responses were obtained. The preselection and reduction of the responses to 253 is based on title and abstract. An article appears to be relevant if it has thematic connection with RE as defined above (Sect. 2) and includes at least one reference to an empirical research method in its abstract or keywords. After sorting out all the redundant items or, e.g., those with fictional case studies, the

final selection for the analysis includes a total of 154 responses (Table 3).

4.2 Analysis and Classification of Empirical Studies into the Research Map

The structuring and evaluation of the identified empirical studies and thus the representation of the results are done by means of a research map. For this purpose, the studies are structured via the framework by assigning the investigated variables to the components of the framework. The evaluation of the studies is based on the classification system (Table 1).

Overall, we identified 68 case studies, 59 experiments, 14 surveys, 5 field studies, 3 expert interviews, 2 grounded theory, 2 narrative reviews, and 1 combination of experiment and case study as research methods. Following Wilde and Hess (2007, p. 285), we refrained from an intensive examination of the rigor of the applied research methods as these are usually already subject of the review procedure.

The presentation of the results is carried out in two stages: the first stage shows components that represent relations as an independent/moderating variable, or as a dependent variable for which individual evidence exists (Table 4). This approach allows us to represent which components and relationships in empirical studies are analyzed in which absolute frequencies.

In the second stage, the structured, evaluated, and refined studies are sorted

into the research map. Since an aggregation in terms of a statistical summarization is not possible, this represents a qualitative overview of the studies found. By means of a tabular presentation of the research map we try to graphically visualize the results. Due to the large number of studies, a complete presentation of all results is not possible, hence only two samples can be provided at this point (Fig. 3). Therefore, the assessment carried out as part of this paper can be accessed on a website (<http://www.frankfurt-school.de/re-studie>).

In addition to the authors, the research methods and the resulting evidence levels are specified. The lettered columns represent the components of the framework. A mark (●) in the left half of the figure indicates that the component is the subject of investigation in the particular study. The right side offers a view on the variables of the study, however no longer at the component level, but at the instance level. The arrow symbol (►) illustrates the relation between independent/moderating variables and the dependent variable. No distinction was made between independent and moderating variables because such a distinction is usually not explicitly made, especially in case studies. Thus, it is sometimes unclear whether, for example, the RE activity “elicitation” has a moderating effect on the relationship between “RE technique” and “RE result” or whether it should also be regarded as independent. The upper half of the figure shows a detail of the research map; in the bottom half, the selection of the components “RE

Table 4 Absolute frequencies of the relationship of the components

		Dependent variables or components								
Independent/ moderating variables or components		A	T	R	N	E	P	K	W	Σ
	A		2	1	1	68				72
	T	32		2		68				102
	R	6			1	7				14
	N	1		3		14				18
	E	1		1						2
	W	9	2			14				25
	P	6				2				8
	K	17		2		20				39
	Σ	72	4	9	2	193	0	0	0	

A = Activity; T = Technique; R = Role; N = Notation; E = Result; P = Principle; K = Context; W = Tool

activity”, “Re technique”, and “RE result” is presented for the instance “elicitation” (Fig. 3).

5 Discussion of the Results and State of Empirical Research in RE

Some tendencies of the state of empirical research in RE can be derived from the two-stage representation of the results. The first stage indicates that the component result (E = Result in Table 4) is the dependent variable most frequently analyzed; the independent or moderating variables most frequently studied are technique followed by activity (Table 4). This result is not surprising considering that the current RE approaches are product- or process-based. Apart from techniques, tools are most frequently used as independent variables to highlight their usage in activities and their impact on results.

It is noteworthy that the significance of notations is examined to a much lesser extent, even though the role of languages in the elicitation, documentation, and validation of requirements is often considered within the research field of conceptual modeling which is closely related to RE. The analysis performed indicates that there is need for research.

Moreover, the significance of roles that are important for the assignment of responsibilities is rarely examined, even though prominent approaches to RE, such as VOLERE and the V-Modell XT, pay high attention to the role concept. Furthermore, principles are investigated only to a small extent, although discussions on this subject are currently in-

tensively conducted. Thus, in comparison to the most frequently cited components, there are relatively few studies that examine principles such as “agility” (two studies), “iterative approach” (one study), and “continuous RE” (three studies).

The results of the literature review also suggest that principle and context are less frequently considered. This clearly indicates a need for further research.

The evaluation of the second stage shows that meta-analyses and narrative reviews with the evidence of Ia and Ib can hardly be found in RE research. Only Davis et al. (2006), who compare the results from experiments for requirements elicitation techniques during the activity “elicitation” in terms of the quantity/quality of requirements, and Ivarsson and Gorschek (2009), who have investigated evaluated “technologies” (techniques, RE approaches, notations) for RE, provide such narrative reviews. Therefore, the strongest evidence exists for the relation of the components activity and technique to result.

In 2002, Kitchenham et al. (2002, p. 721) drastically stated: “In our view, the standard of empirical software engineering research is poor”, which might indirectly apply for RE as well. According to the study carried out herein, this applies today for the quantity of empirical research in RE at least to a mitigated extent. However, for the quality in terms of the “rating” of empirical studies, measured by the underlying classification system, it still seems valid. A closer analysis of the data shows that the evidence levels almost exclusively range in the areas of IIIa and IIIb. This indicates a clear preference for experimental and case-study-based methods. Other research methods,

such as structural equation models, success models, or studies on success factors, which are frequently used in IS research, could not be identified by the search strategy pursued. This can be interpreted as an indication that RE research at this point is not or only just visibly concerned with these research methods and questions, even though analyses as are regularly submitted by the Standish Group demonstrate that there is an interest and a need for them in practice.

On a critical note, the evaluation and very limited further development and empirical (re-)evaluation of artifacts are always carried out by the author of the artifact, possibly in different constellations. Although customary in other sciences such as medicine, a fertile validation, such as by repeated testing by other scientists, is hard to detect (El Emam et al. 1996, p. 13). Consequently, there is no negative evidence, i.e., any negative utility of the investigated artifact usually remains unpublished.

6 Conclusions and Future Work

The objective of this paper is to develop an approach for the structuring of empirical RE research and to undertake an evaluation of existing work. For this purpose, an analogy to EbM is drawn, an approach the basic ideas of which have only been sporadically used in computer science and BISE so far (Davis and Hickey 2004; Dybå et al. 2005; Kitchenham et al. 2009). While those approaches focus on the evidence-based support in practice, we refer to the idea of evidence-based research as a method for acquiring and consolidating knowledge.

Reference	Method	Evidence	A	T	R	N	E	P	K	W	Relations at instance level	A	T	R	N	E	P	K	W
Agarwal et al. 1996	Experiment	IIIb	•	•			•			•	W; A; T ► E	Modeling	Process-oriented, object-oriented			Requirements			Tool support
Al-Karaghoul et al. 2000	Case study	IIIa		•			•				T ► E		Set diagrams			Requirements			
Al-Salem and Samaha 2007	Case study	IIIa	•	•			•		•		K; A; T ► E	Elicitation	VORD			Requirements		Web application	
Alspaugh et al. 2007	Expert interviews	IV			•	•					N ► R			Stakeholder	ScenarioML, use cases, sequence diagrams				
Anda et al. 2009	Case study	IIIa	•			•	•		•		K; A; N ► E	Analysis			Use case	Functional requirements		Large safety-critical software project	
Aoyama 2005	Field study	IIIa	•	•	•				•		K; A; T ► R	Elicitation	Hanako	User				Software embedded in digital consumer products	
Arthur and Gröner 2005	Case study	IIIa	•				•				A ► E	Elicitation				Requirements			
Azar et al. 2007	Case study	IIIa	•	•							T ► A	Triage	Value-oriented prioritization						
Bajaj 2004	Experiment	IIIb				•					N				ERM concepts				



Reference	Method	Evidence	A	T	R	N	E	P	K	W	Relations at instance level	A	T	R	N	E	P	K	W
Browne and Rogich 2001	Experiment	IIIb	•	•			•				A; T ► E	Elicitation	Task characteristic prompts			Requirements			
Cox and Phalp 2007	Case study	IIIa	•	•			•				A; T ► E	Elicitation	Elicitation questions			Use case classes			
Davis et al. 2006	Narrative review	IIb	•	•			•				A; T ► E	Elicitation	RE technique			Requirements			
Fuentes-Fernández et al. 2009	Case study	IIIa	•	•			•				A; T ► E	Elicitation	ATRET			Requirements			
Haumer et al. 1998	Case study	IIIa	•	•			•				A; T ► E	Elicitation	Rich media			Conceptual model			
Leifer et al. 1994	Experiment	IIIb	•	•			•				A; T ► E	Elicitation	Focus group			Requirements			
Mich et al. 2005	Experiment	IIIb	•	•			•				A; T ► E	Elicitation	EPM			Requirements			
Moody et al. 1998	Experiment	IIIb	•	•			•				A; T ► E	Elicitation	Interview, cognitive interview			Requirements			
Pitts and Browne 2007	Experiment	IIIb	•	•			•				A; T ► E	Elicitation	Procedural prompts			Requirements			
Salinesi et al. 2008	Experiment	IIIb	•	•			•				A; T ► E	Elicitation	RITA			Requirements			
Zmud et al. 1993	Experiment	IIIb	•	•			•				A; T ► E	Elicitation	Positive or negative mental imagery, goal-orientation, critical success factors			Requirements			

A = Activity; T = Technique; R = Role; N = Notation; E = Result; P = Principle; K = Context; W = Tool

Fig. 3 Excerpt of the research map and a selection of components and instances

A functional and a technical infrastructure can be considered a crucial prerequisite for evidence-based research. With the framework, we have presented a functional infrastructure intending to support the structuring of empirical research. The results represented in the research map cover a wide range of the components and their relations as derived in Sect. 3. In this sense, the empirical findings can be seen as evidence for the usefulness and relevance of the framework. This means that the first

two of the three requirements defined in Sect. 3.1 are fulfilled.

Further works are required to more scientifically and methodically discuss whether the used classification system for the evaluation of RE research is also appropriate for other questions in BISE/computer science. Especially, the “rating” and thus the superiority of quantitative methods compared to qualitative research methods certainly cannot be applied directly and without contradiction to disciplines which are

more social-science-oriented respectively engineering-oriented.

A third aspect concerning the infrastructure is the support by information systems. Available literature databases only partially support an evidence-based approach as the search options clearly collide with a sound search strategy (Dieste et al. 2009). The preparation of the results for the search of empirical evidence cannot be effective unless when supported by technology. Therefore, the evaluation presented in this paper will be

Abstract

Matthias Goeken, Janusch Patas

Evidence-Based Structuring and Evaluation of Empirical Research in Requirements Engineering

Fundamentals, Framework, Research Map

The objective of the contribution is to develop and motivate an approach of structuring, evaluating, and representing empirical research results regarding requirements engineering. Therefore, the authors develop a framework in order to organize the area of interest. The use of this framework and an evidence-based classification system allows us to develop a research map which helps to structure identified empirical research while enabling the derivation of further research needs. Additionally, it supports the selection of methods, techniques, etc. in requirements engineering practice.

Keywords: Empirical research, Requirements engineering, Evidence-based research, Literature review, Overview

made available as an internet application so that those interested in the matter can fully comprehend the analysis and results as well as search for evidence for current problems on their own (<http://www.frankfurt-school.de/re-studie>).

In further work, it will be necessary to more firmly include the instance level since the results presented herein remain quite abstract and little detail can be shown. Because of the limited space, the concrete results of the studies included in the research map cannot be presented. For this purpose, we are planning to represent the interwoven structure of the framework in a semantic network and to assign the identified relevant empirical work to its edges and nodes. In this sense, a technical infrastructure is aimed at developing an adequate tool of support for evidence-based, empirical research in the field of BISE.

References

- Abran A, Bourque P, Dupuis R, Moore JW (2004) Guide to the software engineering body of knowledge (SWEBOOK). <http://www.computer.org/portal/web/swebok/htmlformat>. Accessed 2009-11-30
- Alter S, Goeken M (2009) Konzeptionelle Metamodelle von IT-Governance-Referenzmodellen als Basis der Kombination und Integration in einer Multi-Model-Umgebung. In: Hansen HR, Karagiannis D, Fill HG (eds) Tagungsband der 9. Internationalen Tagung Wirtschaftsinformatik: Business Services: Konzepte, Technologien, Anwendungen, Wien
- Bayerl B, Alber K, Wohlgenuth WA, Freitag MH, Nagel E (2009) Evidenzbasierte Medizin und gesundheitsbezogene Lebensqualität als potentielle Priorisierungskriterien medizinischer Leistungen am Beispiel der peripheren arteriellen Verschlusskrankheit – Identifikation relevanter Stakeholder und Interviewleitfadenentwicklung. In: Wohlgenuth WA, Freitag MH (Hrsg) Priorisierung in der Medizin – Interdisziplinäre Forschungsansätze. Medizinisch Wissenschaftliche Verlagsgesellschaft, Berlin, S 119–157
- Becker J, Knackstedt R, Pfeiffer D, Janiesch C (2007) Configurative method engineering – on the applicability of reference modeling mechanisms in method engineering. In: Proc 13th Americas conference on information systems, Keystone
- Brereton P, Kitchenham BA, Budgen D, Turner M, Khahil M (2007) Lessons from applying the systematic literature review process within the software engineering domain. *Journal of Systems and Software* 80(4):571–583
- Brinkkemper S (1996) Method engineering: engineering of information systems development methods and tools. *Information and Software Technology* 38(4):275–280
- Browne GJ, Pitts MG (2004) Stopping rules use during information search in design problems. *Organizational Behavior and Human Decision Processes* 95(2):208–224
- Broy M, Geisberger E, Kazmeier J, Rudorfer A, Beetz K (2007) Ein Requirements-Engineering-Referenzmodell. *Informatik-Spektrum* 30(3):127–142
- Bryant A, Charmaz K (2007) Discursive glossary of terms. In: Bryant A, Charmaz K (eds) *The SAGE Handbook of Grounded Theory*. Sage, London, pp 603–612
- Cochrane (nd) Klassifikationssysteme. <http://www.cochrane.de/de/gradesys.htm>. Accessed 2009-07-30
- Davis GB (1982) Strategies for information requirements determination. *IBM Systems Journal* 21(1):4–30
- Davis AM, Hickey AM (2004) A new paradigm for planning and evaluating requirements engineering research. In: 2nd international workshop on comparative evaluation in requirements engineering, Kyoto
- Davis A, Dieste O, Hickey A, Juristo N (2006) Effectiveness of requirements elicitation techniques. In: Proc 14th IEEE international requirements engineering conference
- Dieste O, Grimán A, Juristo N (2009) Developing search strategies for detecting relevant experiments. *Empirical Software Engineering* 14(5):513–539
- DIMDI (2009) Klassifikationen im Gesundheitswesen. <http://www.dimdi.de/static/de/klassi/index.htm>. Accessed 2009-07-30
- Dörr J, Koenig T, Olsson T, Adam S (2006) Das ReqMan Prozessrahmenwerk. IESE-report Nr. 141.06/D
- Dybå T, Kitchenham BA, Jørgensen M (2005) Evidence-based software engineering for practitioners. *IEEE Software* 22(1):58–65
- El Emam K, Quintin S, Madhavji NH (1996) User participation in the requirements engineering process: an empirical study. *Requirements Engineering Journal* 1(1):4–26
- Frank U (2007) Ein Vorschlag zur Konfiguration von Forschungsmethoden in der Wirtschaftsinformatik. In: Lehner F, Zelewski S (eds) *Wissenschaftstheoretische Fundierung und wissenschaftliche Orientierung der Wirtschaftsinformatik*. GITO, Berlin, pp 158–185
- Geisberger E, Berenbach B, Broy B, Kazmeier J, Paulish D, Rudorfer A (2006) Requirements Engineering Reference Model (REM). Technischer Bericht TUM-I0618, TU München
- Gehlert A, Schermann M, Pohl K, Krcmar H (2009) Towards a research method for theory-driven design research. In: Hansen HR, Karagiannis D, Fill HG (eds) Tagungsband der 9. Internationalen Tagung Wirtschaftsinformatik: Business Services: Konzepte, Technologien, Anwendungen, Wien
- Goeken M (2006) Entwicklung von Data-Warehouse-Systemen – Anforderungsmanagement, Modellierung, Implementierung. DUV, Wiesbaden
- Gutzwiller TA (1994) Das CC-RIM-Referenzmodell für den Entwurf von betrieblichen, transaktionsorientierten Informationssystemen. Physica, Heidelberg
- Heym M (1993) Methoden-Engineering: Spezifikation und Integration von Entwicklungsmethoden für Informationssysteme. Rosch-Buch, Hallstadt
- Hickey AM, Davis AM (2003a) Elicitation technique selection: how do experts do it. In: Proc 11th IEEE international conference on requirements engineering, Monterey Bay
- Hickey AM, Davis AM (2003b) Requirements elicitation and elicitation techniques selection: a model for two knowledge-intensive software development processes. In: Proc 36th Hawaii international conference on systems science, Big Island

- Hevner AR, March ST, Park R, Ram S (2004) Design science in information systems research. *MIS Quarterly* 28(1):75–105
- Ivarsson M, Gorschek T (2009) Technology transfer decision support in requirements engineering research: a systematic review of REJ. *Requirements Engineering Journal* 14(3):155–175
- Jarke M (2009) Perspectives of the interplay between business and information systems engineering and computer science. *BISE* 51(1):70–74
- Juristo N, Moreno AM (2001) Basics of software engineering experimentation. Kluwer, Boston
- Karlsson F (2002) Meta-method for method configuration – a rational unified process case. <http://liu.diva-portal.org/smash/get/diva2:21357/FULLTEXT01>. Accessed 2009-07-14
- Kitchenham BA, Brereton OP, Budgen D, Turner M, Bailey J, Linkman S (2009) Systematic literature reviews in software engineering – a systematic literature review. *Information and Software Technology* 51(1):7–15
- Kitchenham BA, Dybå T, Jørgensen M (2004) Evidence-based software engineering. In: *Proc 26th international conference on software engineering*, Edinburgh
- Kitchenham BA, Pfleeger SL, Hoaglin DC, El Emam K, Rosenberg J (2002) Preliminary guidelines for empirical research in software engineering. *IEEE Transactions on Software Engineering* 28(8):721–734
- Kotonya G, Sommerville I (1998) Requirements engineering: process and techniques. Wiley, New York
- Kruchten P (1999) *Der Rational Unified Process: Eine Einführung*. Addison-Wesley, München
- Kühne T (2006) Matters of (meta-)modeling. *Journal on Software and Systems Modeling* 5(4):369–385
- Macaulay LA (1996) Requirements engineering (applied computing). Springer, Heidelberg
- Maier W, Möller HJ (2007) Metaanalysen. Methode zur Evidenzmaximierung von Therapiestudien? *Der Nervenarzt* 78(9):1028–1036
- Pohl K (2008) Requirements Engineering – Grundlagen, Prinzipien, Techniken. dpunkt.verlag, Heidelberg
- Pohl K (1997) Requirements engineering: an overview. In: Kent A, Williams J (eds) *Encyclopedia of computer science and technology*, 36th edn. Dekker, New York, pp 345–386
- Pohl K, Rupp C (2009) Basiswissen Requirements Engineering. Aus- und Weiterbildung nach IREB-Standard zum Certified Professional for Requirements Engineering Foundation Level. dpunkt.verlag, Heidelberg
- Popper KR (1959) *The logic of scientific discovery*. Harper and Row, New York
- Raspe HH (1996) Evidence based medicine: Modischer Unsinn, alter Wein in neuen Schläuchen oder aktuelle Notwendigkeit? *Zeitschrift für ärztliche Fortbildung* 90:553–562
- Robertson S, Robertson J (2006) *Mastering the requirements process*, 2nd edn. Addison-Wesley, Amsterdam
- Rupp C, Simon M, Hocker F (2009) Requirements engineering und management. *HMD – Praxis der Wirtschaftsinformatik* 48(267):94–103
- Sackett DL, Rosenberg WMC, Gray JAM, Haynes RB, Richardson WS (1996) Evidence-based medicine: what it is and what it isn't. *British Medical Journal* 312(7023):71–72
- SEI (2007) Software Engineering Institute: Capability Maturity Model Integration (CMMI), 2007. <http://www.sei.cmu.edu/cmmi/cmmi.htm>. Accessed 2009-12-15
- Shuja AK, Krebs J (2008) IBM rational unified process reference and certification guide: solution designer: solution designer (RUP). Addison-Wesley, Amsterdam
- Strahringer S (1996) *Metamodellierung als Instrument des Methodenvergleichs*. Shaker, Aachen
- V-Modell XT (2006) V-Modell® XT 1.3. <http://ftp.tu-clausthal.de/pub/institute/informatik/v-modell-xt/Releases/1.3/V-Modell-XT-Gesamt.pdf>. Accessed 2009-07-08
- vom Brocke J, Simons A, Niehaves B, Riemer K, Plattfaut R, Cleven A (2009) Reconstructing the giant: on the importance of rigour in documenting the literature search process. In: *Proc 17th European conference on information systems 2009*, Verona
- Wieringa RJ, Heerkens JMG (2006) The methodological soundness of requirements engineering papers: a conceptual framework and two case studies. *Requirements Engineering Journal* 11(4):295–307
- Wieringa R, Heerkens H (2004) Evaluating the structure of research papers: a case study. In: *Second international workshop in comparative evaluation of requirements engineering*, Kyoto
- Wilde T, Hess T (2007) *Forschungsmethoden der Wirtschaftsinformatik*. WIRTSCHAFTSINFORMATIK 49(4):280–287
- Wohlin C, Runeson P, Höst M, Ohlsson MC, Regnell B, Wesslén A (2000) *Experimentation in software engineering: an introduction*. Kluwer, Boston
- Zelkowitz MV, Wallace D (1997) Experimental validation in software engineering. *Information and Software Technology* 39(11):735–743