Requirements Eng (2009) 14:155–175
DOI 10.1007/s00766-009-0080-1

ORIGINAL ARTICLE

Technology transfer decision support in requirements engineering research: a systematic review of REj

Martin Ivarsson · Tony Gorschek

Received: 22 October 2008 / Accepted: 16 February 2009 / Published online: 7 March 2009

Abstract One of the main goals of an applied research field such as requirements engineering is the transfer of research results to industrial use. To promote industrial adoption of technologies developed in academia, researchers need to provide tangible evidence of the advantages of using them. This can be done through industry validation, enabling researchers to test and validate technologies in a real setting with real users and applications. The evidence obtained, together with detailed information on how the validation was conducted, offers rich decision support material for industrial practitioners seeking to adopt new technologies. This paper presents a comprehensive systematic literature review of all papers published in the Requirements Engineering journal containing any type of technology evaluation. The aim is to gauge the support for technology transfer, i.e., to what degree industrial practitioners can use the reporting of technology evaluations in the journal as decision support for adopting the technologies in industrial practice. Findings show that very few evaluations offer full technology transfer support, i.e., have a realistic scale, application or subjects. The major improvement potential concerning support for technology transfer is found to be the subjects used in the evaluations. Attaining company support, including support for using practitioners as subjects, is vital for technology transfer and for researchers seeking to validate technologies.

Keywords Systematic review · Requirements engineering · Technology transfer

M. Ivarsson (✉)
Department of Computer Science and Engineering,
Chalmers University of Technology,
412 96 Göteborg, Sweden
e-mail: martin.ivarsson@chalmers.se

T. Gorschek
School of Engineering, Blekinge Institute of Technology,
Ronneby, Sweden
e-mail: tony.gorschek@bth.se

1 Introduction

To maintain and increase competitive advantages, software organizations must continually strive to improve the processes and practices they use to produce software. This is especially true for requirements engineering (RE) as good RE practices are a decisive success factor in large-scale software development. Following process assessment, any issues identified and subsequent improvement efforts are centered on selecting and introducing new technologies to counter weaknesses, and in this context technologies can be anything from methods, techniques and procedures to models, and tools (or combinations thereof) [1, 2].

When searching for new technologies, practitioners in industry need adequate decision support material on which to base their decision to adopt. In essence, they need to be able to evaluate the evidence offered in relation to the technology to be able to make an informed decision concerning what choice provides the best investment and to minimize risks associated with the introduction of a new technology. Researchers developing new technologies need to provide not only compelling evidence of advantages over already existing ones, but also base information on the extent to which the technology has been validated (tested).

Ultimately, success in an applied engineering field such as requirements engineering is reflected in the level of adoption and use of its research results [3]. Still, research in requirements engineering is suggested to have little impact on requirements engineering practices [4, 5]. Reasons might be found in the criticism received for not being relevant to practice and not providing technologies usable in real environments [6, 7]. In addition, the research that has been done has been found lacking proper evaluation [8, 9] leading to little tangible evidence that can be used by practitioners as decision support. However, the arguments are not one-sided, as requirements engineering research is also claimed to provide practitioners with useful results [10].

This paper is intended to provide an objective view of what technologies are present in requirements engineering research, more specifically published in the Requirements Engineering journal (REj), and to what extent papers describing these technologies provide decision support for practitioners seeking to adopt technologies. A systematic literature review of all research papers published in REj forms the basis for evaluating the evidence presented in a technology transfer perspective. The major differences between a traditional review and a systematic one lies in that systematic reviews aim to minimize error and bias to increase the quality of the review. This is done by using explicit and rigorous methods to identify, appraise and synthesize research on particular research questions established prior to the actual review [11].

For evidence to be convincing to industry, thus facilitating technology transfer, it needs to be realistic. This means that the environment, subjects and scale of the evaluation should be as realistic as possible [3]. Making evaluations in a realistic setting aids technology transfer as practitioners can evaluate the context to see whether the results are transferable to their own environment [12]. It also addresses aspects such as scalability and usefulness, which are often hard to evaluate in a small-scale evaluation [6].

The research method used to produce the evidence also influences the way practitioners perceive it. In this case, practitioners value methods that are most relevant to their own environments, such as case studies and lessons learned [13]. In addition to the evidence itself, the manner in which it is presented is also important. Without detailed descriptions of the study's design, validity and context, an assessment of the evidence produced is difficult. All these aspects are taken into consideration in the systematic literature review presented in this paper.

The paper is structured as follows. Section 2.1 introduces technology transfer in software engineering and outlines a generic technology transfer model and Sect. 2.2 provides an overview of related work. The design of the study is described in Sect. 3. The execution of the study is briefly described in Sect. 4, and a discussion of the validity of the study is presented in Sect. 5. The results of the review are presented and discussed in relation to the research questions in Sect. 6. Conclusions are given in Sect. 7.

2 Background and related work

This section gives an introduction and background to technology transfer and related work.

2.1 Technology transfer in software engineering

A wide gap between what is used in requirements engineering practice and what is proposed in the research has been identified [4, 5]. The technology transfer from academia to industry is at the center of this gap.

Technology transfer is the process of moving new technologies from academia, and a laboratory environment, to industry and an organization, where they are used to perform engineering tasks. There are two main perspectives of technology transfer. In a research (academic) perspective, technology transfer enables researchers to validate (test) technologies in a real setting. This is only accomplished by enabling and attaining industry transfer [14], as usability and usefulness in industry is the ultimate test of a technology. The second perspective has to do with the research impact on industry. RE practice is not likely to improve if no technology is transferred into practice [15].

Little is known about technology transfer in software engineering. Redwine and Riddle [16] investigated the time needed for maturation of technologies developed in the 1960s and 1970s. They found that it takes in the order of 15–20 years to mature a technology to a state where it can be popularized and disseminated to the technical community, which is a long time in a rapidly changing industry [1]. Looking at these numbers as an argument for sustaining technology maturation lead-times has been called a circular argument [17], i.e., there is nothing inherent in software engineering that prohibits shortening the time needed for technology transfer. Zelkowitz [18] studied the infusion processes, i.e., the process a organization undergoes to adopt a new technology. He found that infusion at NASA took 2–4 years, which included training, and two to four pilot projects to tailor the technology to the given environment. He also noted that infusion of software engineering technologies differs from other technologies in that they are not products and that it is crucial to understand software engineering technologies to support technology transfer.

Even though technology transfer has received relatively little interest in software engineering, other disciplines have well-established frameworks. Rogers [19] suggests five technology attributes relevant to the relative speed at which technologies are adopted:

- Relative advantage. The degree to which a new technology is better than one already available.
- Complexity. The degree to which the new technology is easy to understand and use.
- Compatibility. The degree to which a new technology is consistent with existing values, past experiences, and needs of potential adopters.
- Testability. The degree to which the new technology can be tried out and tested on a limited basis.
- Observability. The degree to which the new technology has visible results for adopters.

Based on Rogers' framework, recommendations for software engineering research to facilitate technology diffusion have been suggested [1, 2, 17]:

- Software engineering researchers should provide tangible evidence through empirical studies that show the advantages, risks, and potential benefits of the new technologies.
- Software engineering researchers need to communicate the context in which and the assumptions under which the evidence presented was observed.
- Software engineering researchers must develop technologies that carry with them a relative advantage of using them.
- Software engineering researchers need to package their technologies with training materials and CASE tool support to enable easier adoption by practitioners.

Several technology transfer models in software engineering embody the need for maturing technologies through different kinds of studies that provide different types of evidence [1, 6, 14, 20, 21]. A general model for technology transfer developed for software engineering is shown in Fig. 1 (see Gorschek et al. [14]). The model consists of innovation, static validation, dynamic validation, and finally, the release of the technology for wider use.

The innovation step is where the idea is born and developed into a technology. The technology is then tested and evaluated in different settings, from laboratory to industry, utilizing different types of research methods as

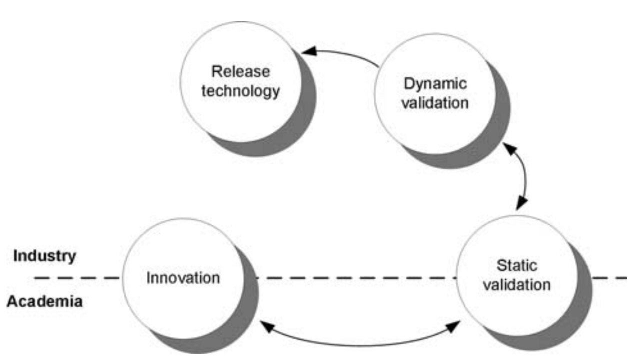

Fig. 1 Technology transfer process (adapted from [14])

needed. Static validation of technologies often involves experimentation to investigate the basic concepts of the technology in order to sort out teething problems before the technology is tested in production projects. Given the results of the static validation, one can either move back to the innovation step and refine the idea or move on to dynamic validation to test the idea in a real life setting. Dynamic validation is carried out in case studies in either "real" software projects or a smaller pilot project aimed at evaluating the technology in question. The last step in the technology transfer process is to release the technology for wider use when it has been shown to be useful and usable.

From an academic research perspective, each step of validation provides different kinds of evidence in relation to the technology being tested. In static validation, variables can be controlled and hypotheses on effectiveness, efficiency, and scalability can be tested initially. Dynamic validation does not offer the same level of control; instead, the results of case studies provide answers having to do with contextual factors and how the technology scales to real life industry use.

From an industry perspective, the staged technology transfer model described here is an example of how evidence can be obtained through a gradual validation (testing) of a technology in several steps. This evidence, together with detailed information about how the validation rounds were performed (e.g. contextual data), can offer rich decision support for technology transfer.

2.2 Related work

A comprehensive, systematic review was made of all papers published in RE journal to evaluate the support offered to technology transfer in current RE research. Systematic reviews have been used in other fields such as medical research and have received increasing attention in software engineering since they were introduced as a means to evaluate and interpret a field's collective evidence related to a particular research question in evidence-based software engineering [21, 22]. To be able to answer the research questions posed in this systematic review, a data extraction form aimed at collecting information is established. Several different data extraction forms have been used in systematic reviews of software engineering research that range from properties for characterizing the research investigated, e.g., the topic investigated, the research method used [23], to properties aimed at synthesizing several studies carried out in a field, e.g., accuracy and variance [24].

Several reviews, both systemic and traditional, have been published that relate to requirements engineering, some of which appear in Table 1. The studies' purpose, scope, literature delimitation, data extraction properties, and major results are summarized. It is also noted whether the review is systematic. It can be seen that the studies utilize a diverse set of data extraction properties (for additional ones see [25]). The main idea of constructing the data extraction form is that the properties chosen address the research questions. Thus, as this paper presents a characterization of RE research and identifies evidence from the perspective of technology transfer, a data extraction form aimed at capturing properties relevant for technology transfer is applied (see Sect. 3.1.3).

In a systematic review, a further factor is the delimitation of the literature that should be included in the review, depending on the purpose of the review. When a particular phenomenon is investigated, all relevant papers should be identified and reviewed, regardless of the kind of source. An example of this type of delimitation is seen in the review by Davis et al. [27] (summarized in Table 1) that investigates empirical studies concerning the effectiveness of elicitation techniques. However, when the purpose is to characterize a field, delimitation with respect to time and source might be appropriate, which is done to limit the effort needed to do the review while still providing useable results. An example of a review characterizing research in software engineering is found in [23] (summarized in Table 1), where papers from six leading journals are reviewed. This is the type of delimitation used in this literature review, as the purpose is to investigate the requirements engineering field from a technology transfer perspective.

3 Study design overview

This section gives a detailed account of the design of the systematic literature review.

3.1 Systematic review design

The process used to conduct the review is shown in Fig. 2. It consists first of identifying papers that should be

Table 1 Reviews related to requirements engineering

	Glass et al. [23]	Wieringa and Heerkens [26]	Davis et al. [27]	Parviainen and Tihinen [28]
Purpose	Examine the state of software engineering research	Investigate the methodological soundness of RE papers	Evaluate the effectiveness of elicitation techniques	To provide an inventory of RE technologies and their coverage
Systematic review	Yes	No	Yes	No
Scope	SE	RE	RE	RE
Literature delimitation	Six leading journals	Research papers submitted to RE'03 conference	Empirical studies found through SCOPUS, IEEEXPLORE and ACM DL	Not mentioned
Properties used for data extraction	Topic, research approach, research method, reference discipline, level of analysis	World problem solved, how the problem was solved, relevance of the solution	Elicitation technique, features of the empirical study, evidence hierarchy, data on effectiveness of each study	Established and well known, understandable and usable, flexible and modifiable, embedded viewpoint, support for different requirements types, up-to-date information, up-to-date experiences, phase coverage, activity coverage
Major results	SE research is diverse regarding topic, narrow regarding research approach and the method inwardly focused regarding reference discipline and technically focused regarding the level of analysis	Lack of problem investigation and solution validation	Interviews appear to be one of the most effective elicitation techniques. Several techniques often cited in the literature tend to be less effective than interviews. Analyst experience does not appear to be a relevant factor. Prototypes have not been shown to have significant positive effects during elicitation	An overview of available technology support for RE

Fig. 2 Review process

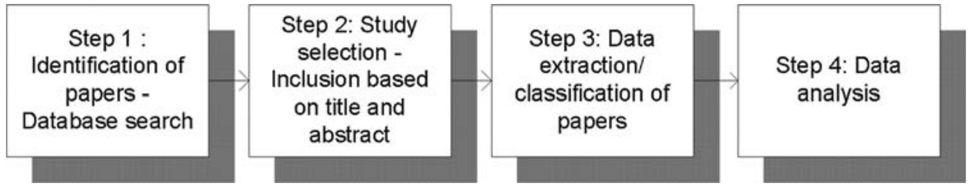

Table 2 Search terms

Population: Requirement*, specification	
Intervention: empiric* OR experience* OR "lesson learned" OR "lesson learnt" OR "lessons learned" OR "lessons learnt" OR evaluat* OR validation* OR experiment* OR stud* OR case* OR example* OR survey OR analys* OR investig* OR demonstrate*	

included in the study based on a database search. The second step is to include papers on the basis of the title and abstract of the papers that were identified in the search. The papers included are then classified according to the data extraction form presented in Sect. 3.1.3 to enable analyzing the research identified. The review process is based on guidelines provided by Kitchenham and Charters [11] with the difference that the study quality assessment is included in the inclusion criteria and scoping, i.e., only papers that present any type of evidence or evaluation related to RE technologies are included in the study.

The following sections describe the details of each step in the review process.

3.1.1 Step 1: Identification of papers

A comprehensive and unbiased identification of research is one of the factors that differentiate a systematic review from a traditional one. The main criterion used in establishing the keywords used is that they identify requirements engineering technologies that have been evaluated. The search term was elaborated over several test searches and trial reviews, which aimed to establish synonyms of terms relevant to the review. The population is requirements technologies and the intervention is that the technology has been evaluated in some form. The search terms used for population and intervention are presented in Table 2 and the resulting search term used is then **population** AND **intervention**. No search terms were used to discern whether the paper presented a technology; this was deemed impossible as the terminology differed too much between the papers.

The database search is done using Inspec and is only applied to the title and abstract as full text searches would yield too many irrelevant results [29].

The next section describes the criteria used to determine whether papers identified in the search should be included in the review, i.e., remain for data extraction.

3.1.2 Step 2: Paper selection

The papers identified in the search were examined so that only the ones relevant for answering the research questions would be included (see Sect. 3.1.3). The main criterion for including papers is that they present an evaluation (of any sort) of a technology that has bearing on requirements engineering. Evaluation is defined to cover a wide range of activities, making the selection "include heavy", i.e., avoiding dismissing papers that have some sort of evaluation. The term evaluation thus covers the range from application (test/illustration) of a technology on a toy example invented by the researchers themselves, to experiments and any sort of empirical evaluation. In addition, only research papers were considered for inclusion, not editorials, news, correspondence, comments, and so on.

The inclusion procedure was applied to the title and abstract, that is no papers were excluded based solely on the title. This is an inclusion-heavy selection, i.e., the selection of papers is inclined towards inclusion rather than exclusion. The motivation for including papers based on reading abstracts is that, if the evaluation presented in the paper is a major point of the paper, it is likely to be mentioned in the abstract. Missing including papers that present some form of evaluation on the grounds of only reading the abstract is not considered a threat to the validity of the study as the evaluation is not likely to be a major point of the paper in these cases. In addition, having an inclusion-heavy selection enables more papers to pass on to a more elaborate scrutiny and limits the risk of missing papers that are relevant. All papers included were entered into a database for further analysis, which is described in the next section.

3.1.3 Step 3: Research questions and data extraction

Before elaborating on data extraction and analysis, it is important to have a clear idea of the research questions posed for the evaluation. Each research question is listed in Table 3, with a description/motivation. Each of these research questions can then be mapped to the data extraction form shown in Table 4.

The papers included in the study are reviewed and data are extracted according to the data extraction form presented in Table 4. The data extraction form is derived from

Table 3 Research questions

Research question	Description
RQ1: What RE technologies exist?	To provide an inventory of all RE technologies that have been subjected to some sort of evaluation. In this study, the minimum level of evaluation is that the technology has been applied to an example. The sub research questions further detail the characteristics of the technologies identified.
RQ1.1: What RE sub-process areas do the technologies target?	The requirements engineering sub-process areas targeted by the technologies.
RQ1.2: What is the timeline focus of the technologies?	In what timeline perspective the technologies are used, i.e., whether the technologies are used before the project has started, in the project or after it has finished.
RQ2: What is the state of technology evaluation?	The state of evaluation is in this study evaluated by combining the aspects considered in the sub research questions. Each aspect is further detailed in the sub research questions.
RQ2.1: What research methods are most common in the evaluation of technologies?	Different research methods offer different kinds of evidence. The evidence produced varies in terms of level of control and realism. Experiments usually offer higher levels of control while case studies offer more realism.
RQ2.2: In what context are the technologies evaluated?	For evaluations to be realistic, they need to be performed in a real setting. This research question investigates the context in which the evaluations are performed. As the purpose is to investigate the amount of realistic evaluation, the only distinction made is between evaluations in real industrial settings and academic settings.
RQ2.3: What subjects are used in the evaluation?	The subjects used in the evaluation will influence the results. A distinction is made between industrial practitioners, researchers and students.
RQ2.4: What scale do evaluations have?	The scale of evaluations is estimated by looking at the applications used. The scale can thus range between toy examples to real industrial applications.
RQ2.5: What degree of realism do the evaluations have?	Combining research method, context, subjects and scale gives the realism of the evaluation.
RQ3: To what extent does the state of research support the actual adoption of technologies?	To support adoption of technologies, evaluations need to present plausible evidence of the benefit of using technologies. The valuation of evidence is based on the state of technology evaluation treated by RQ2. In addition, the presentation of the evaluation needs to provide a basis for appraising it. This is evaluated by looking at the extent to which the study design, context of evaluation and the validity of the study are presented. These three aspects are further examined in the following sub research question.
RQ3.1: To what degree are the evaluations described?	To understand whether the results presented can be transferred to another context, the presentation of the evaluation needs to be understandable. The context of the study needs to be described to be able to understand whether the results presented can be transferred to another environment. Technologies evaluated in a small project might not perform the same in a large project. Description of how a study is set up and executed increases the understanding of the results presented. Finally, the validity of a study concerns how valid the results are for the population of interest. To comprehend the results of a study the validity must be described.

the research questions, and the mapping to these can be seen in the table. For example, RQ1 concerns the technologies presented and is characterized by three different properties (see properties number 1–3 in Table 4). All properties are noted in accordance with what is reported in the papers reviewed (where possible). If the property is not mentioned, it is marked in accordance to the reviewer's understanding; as the aim is to investigate the support for technology transfer, all evidence presented in the papers is considered relevant.

The first three properties (1–3) are marked once for each paper and the rest (4–10) are marked once for each study appearing in the paper. This enables capturing cases where several evaluations of a particular technology are presented in one paper. For the first three properties (1–3), one paper can have several different values, e.g., a paper can present several different technologies. In these cases, the property is marked once for each suitable value.

The remaining seven properties (4–10) can be mapped to the research questions treating state of technology evaluation (RQ2) and support for technology transfer (RQ3). The properties all concern the credibility of the evidence, i.e., how the evidence is produced and presented. This is derived from technology transfer being able to be supported by providing trustworthy, realistic evidence [3] and descriptions of how and where, i.e., in what context the evidence has been produced. Each of these properties is marked only once per study. For example, one study

Table 4 Data extraction form

Nr	Property	Values	Description	Mapping to RQs
1	Technology	From the reviewed literature.	The name of the technology under evaluation. These values are compiled from the reviewed papers.	RQ1
2	Project focus	Pre-project In-project Post-project N/A	Project focus specifies whether the technology is meant to be used before starting the project, e.g. market driven, in the project or after the project is concluded.	RQ1
3	Sub-process area	Elicitation Analysis and negotiation Management Validation Specification N/A	Process area comprises the requirements process areas to which the technology evaluated belongs.	RQ1
4	Research method	Action research Conceptual analysis Lessons learned Conceptual analysis/mathematical Case study Field study Laboratory experiment (human subject) Laboratory experiment (software) Interview Descriptive/exploratory survey Other N/A	Captures the research method used to evaluate the technology.	RQ2, RQ3
5	Context described	Strong Medium Weak	Specifies the degree to which the context of the study is described.	RQ3
6	Study design described	Strong Medium Weak	Specifies the degree to which the design of the study is described.	RQ3
7	Validity discussed	Strong Medium Weak	Specifies the degree to which the validity of the study is discussed.	RQ3
8	Subjects	Practitioner Researcher Student Not mentioned	Specifies who uses the technology in the evaluation.	RQ2, RQ3
9	Context	Academia Industry	Specifies the context in which the evaluation is made.	RQ2, RQ3
10	Scale of evaluation	Toy example Down-scaled real example Industrial Not mentioned	Specifies the scale on which the evaluation is made.	RQ2, RQ3

utilizing different research methods will only be scored as using one. In these cases papers are scored in a favorable way, and the research method providing the strongest evidence is used.

The *Technology* (1) property is scored according to what is presented in the papers. It is also noted whether the technology is a refinement or further development of another technology.

Project focus (2) deals with the perspective of timeline of the technologies in relation to projects, i.e., the technology's main use is either before, during or after the project.

Sub-process area (3) categorizes technologies with respect to what sub-process area of requirements engineering they are used in. The values of this property are adapted from [30].

Research method (4), adapted from [23], captures the research method used to produce the results.

Context described (5), *Study design described* (6) and *Validity discussed* (7) all relate to the paper conveying an understanding of context and the assumptions under which the evidence for a particular technology have been derived. This is one of the aspects that has been identified to support technology transfer, as practitioners can evaluate the similarity to their own environment and thus assess whether the evidence is valid for them [2, 12, 13, 19].

Context described (5) is scored on a three-level scale where *Weak* implies that the context relevant for the study is not mentioned in the paper at all. The property is marked as *Medium* when the organization/company and development effort is presented in brief, while *Strong* involves characterizing the development effort/organization. Characterization involves development mode, e.g., contract driven, market driven etc., development speed, e.g., short time-to-market, company maturity, e.g., start-up, market leader etc.

The *Study design described* (6) property captures the degree to which the design of the study is described. This involves presenting the variables measured, the treatments, the control etc. The *Weak* value indicates no description at all of the design, while *Medium* indicates a brief description of how the study was carried out, e.g. "ten students did step 1, step 2 and step 3". *Strong* study design indicates that the selection/sampling of subjects, variables measured and so on is present. For example, "ten students were randomly assigned to two groups in a one-factor, two-treatments experimental design".

The *Validity described* (7) property captures how the validity of the studies is presented. The validity of a study is important as it divulges how valid the results are for the population of interest [31]. In this case, *Weak* indicates that the validity of the study is not discussed at all, while *Medium* refers to the case in which the author has mentioned the validity but not discussed it in detail, e.g., "The action research approach means that claims for the generalizability of our findings are limited" [32]. The description of validity is considered to be *Strong* when it is discussed according to classification schemes like the ones presented in [31, 33] or the equivalent.

In addition to *Research method* (4), *Subjects* (8), *Context* (9) and *Scale of evaluation* (10) are also related to the RQ2 (the properties are influenced by [3]).

Subjects (8) captures the subjects that use the technology in the evaluation. A case study of a real development effort where practitioners use the technology under evaluation would imply that the practitioners are the subjects. On the other hand, if a researcher takes part in a real development effort and the researcher himself uses a new technology to see whether it can be used in real software development, the subject would be noted as being the researcher.

Context (9) captures the context in which the study has been carried out and is coarse grained in that it distinguishes only between academia and industry.

Finally, *Scale of evaluation* (10) captures what type of application the technology was used on in the evaluation. *Scale of evaluation* (10) ranges from *Toy examples* used for student projects and small examples to *Industrial* scale applications.

4 Execution

The execution[1] of the review involved searching the Requirements Engineering journal (REj) using the Inspec database and manually searching the issues not indexed in the database. The reason for including papers exclusively from REj is that it is the premier publication venue for researchers in requirements engineering. Thus, it should contain the most mature research in the field, implying a higher degree of empirical evidence and higher quality than research presented in workshops or conferences. This makes the selection suitable for answering the research questions posed here. The database search resulted in the identification of 181 papers for possible inclusion in the review. This high number found by the search in comparison to the total in the journal is expected, as papers published in REj most often discuss requirements engineering topics. Approximately half of the papers remained for data extraction when the inclusion criteria had been applied to the title and abstract of the papers. Upon completing the search, it was noted that Inspec did not index all issues in REj. Thus, the non-indexed ones were searched manually.

To test whether the search or inclusion criteria missed relevant papers, 10% of the issues were randomly selected and the full paper was read. No relevant papers were found that had not already been included, indicating that the inclusion/exclusion criteria were robust.

Of the 99 papers considered for data extraction, two were removed from the study when the full paper was read

[1] The search was carried out on June 23 2008 and two of the articles were published as online first articles. Inspec indexed articles from issue 1:1 up until issue 12:2. The rest of the issues were searched manually.

on the basis of their not having presented an evaluation of a requirements engineering technology. In total, 97 papers were classified in which 125 studies were identified and classified.

5 Threats to validity

The main threats to validity of this study are publication and selection bias, and data extraction, each detailed below.

5.1 Publication and selection bias

Excluding papers not published in REj limits the possibility of generalizing the results as it is only one of the forums in which research in requirements engineering is published. Moreover, there is a chance of missing technologies and evaluations when other publication venues are not considered, e.g., other journals, conferences, technical reports etc. However, as stated in Sect. 4, REj is the premier publication venue for researchers in requirements engineering and should thus include the most mature technologies and evaluations.

Another threat to the validity is the selection of papers from the journal. First, the key words used to search the publication database could miss papers relevant for inclusion in the review. As few papers (27) were removed by the search, however, it is reasonable to assume that this threat is limited. Second, the inclusion criteria used to include papers in the review are based on a reading of the abstract. This introduces a threat, as abstracts might not reflect the actual contents of the papers. This threat was investigated, as described in Sect. 4, and found to be limited.

5.2 Data extraction

A potential threat to validity is that the judgments used to include/exclude and to classify papers are biased. To limit this threat, the process used to include/exclude papers and the classification scheme was pilot tested prior to the study. The classification scheme itself is also derived from the research questions and is based on previous studies [23], which limits the threat.

One threat to validity inherent to classifying research presentations is that the classification will be of the presentation instead of the actual research performed. This is not a primary concern, however, as the presentation of the research is of central importance since the support for technology transfer is dependent on the descriptions of the work in the publications. In addition, not presenting the details used in a study limits the possibilities for other researchers to replicate the studies and thus directly affects the credibility of the research.

Another potential problem in reviewing research literature from the perspective of technology transfer is that research papers are written for researchers and not practitioners. However, requirements engineering is an engineering field, speaking towards application, meaning that research presented should not be too far from reality or use. Therefore, we feel that it is feasible to evaluate the research in this way.

With respect to the actual data extraction, the classification of papers is prone to some subjective variations owing to the inherent complexity of classifying software engineering research. There may be several plausible classifications for one paper [23], meaning that, if the study is replicated, the results might vary to some degree. This is primarily because the classification scheme is not mutually exclusive. However, in the cases in which several different classifications were possible, the papers were classified favorably with respect to the subsequent analysis. This means that the results presented in this systematic review are in some sense a best-case scenario, i.e., gives the researchers the benefit of the doubt.

6 Results and analysis

This section presents the results of the classification and is arranged in the order of the research questions presented in Sect. 3.1.3.

6.1 RQ1: What RE technologies exist?

A plethora of technologies was found, presented in Appendix together with the number of papers discussing each technology (given in the column denoted "#"). The indentation in front of the technologies in Appendix indicates that the technology is a refinement of the technology specified on the row above it (e.g. "–Cost-value approach" is a refinement of "AHP"). This division and sorting is not an invention of this review but is rather taken from the papers included in the review.

The papers are classified such that several technologies can be noted for each paper. This was necessary as several papers presented evaluations that used a synthesis of technologies. Only the technologies that are the focus of each paper is noted as being discussed, i.e., counted. This means that technologies that have refinements in Appendix might be marked as N/A (see e.g. Goal-orientation).

In Appendix, it can be seen that few of the technologies are discussed in more than one or two papers. The technologies that have received the most attention are Language extended lexicon (LEL) [34–37] and Use cases

[38–42], with four and five publications on each technology, respectively. In total, 101 technologies were found in a total of 97 papers included in the review.

6.1.1 RQ1.1: What RE sub-process areas do the technologies target?

The requirements engineering sub-process area(s) targeted by the technologies treated were seldom explicitly stated in the papers and thus had to be mapped. Technologies often address several different requirements engineering sub-process areas, which was noted when marking this property. Table 5 shows that technologies aimed at the sub-process area of *Analysis and Negotiation* has by far the largest representation (44 technologies), followed by *Specification* (30 technologies) and *Elicitation* (23 technologies). It is notable that technologies for *Management* have the lowest representation.

6.1.2 RQ1.2: What is the timeline focus of the technologies?

Looking at the perspective of timeline, most technologies are developed for use *In-project* (i.e., for use in already started projects), as can be seen in Table 6. Not a single technology is meant to be used *Post-project* and only nine technologies for use in *Pre-project* RE.

The property of timeline was seldom mentioned explicitly in the papers. The technologies were classified as *In-project* if the understanding of the system under construction was perceived as being rather mature and as *Pre-project* if pre-project work was mentioned in the paper. If neither of these were fulfilled, the paper was classified as N/A as no article addressed requirements engineering technologies used after the project was concluded. Most of the articles that were classified as *Pre-project* had a distinct market driven requirements engineering perspective (see e.g. papers [43, 44]).

Summary and discussion of RQ1: To summarize the findings for RQ1 and its sub research questions, one can notice that researchers in requirements engineering more often than not develop new technologies that they evaluate and then move on to develop new ones. This is noted in Sect. 6.1, where few technologies have several studies that are related to them. This may imply that new technologies are presented once and that replication by other authors or additional evaluations (e.g. reporting industry trials etc.) are not common.

Looking at the sub-process area addressed by the technologies, researchers tend to focus on technologies to be used in a project that is already started (in-project). This could potentially be a problem, as uncovering requirements prior to projects is difficult, and failure to do so makes the choice of subsequent analysis tools irrelevant [4]. From a market-driven perspective, pre-project RE is critical and many companies working with long-term product development (e.g. automotive, automation, telecommunication and so on) are completely dependent on pre-project RE to decide what requirements they should include in a release [45, 46]. This perspective is not the focus at all of the technologies presented. This is confirmed by the low focus on management technologies, as management of large-scale RE is a central part of market-driven RE [47, 48].

Table 5 Requirements engineering sub-process areas addressed

Sub-process area	Number
Analysis and negotiation	44
Elicitation	23
Management	9
N/A	3
Specification	30
Validation	19

Table 6 Timeline focus of the technologies

Timeline focus	Number
Pre-project	9
In-project	64
Post-project	0
N/A	24

6.2 RQ2: What is the state of technology evaluation?

Assessing the state of technology evaluation involves examining several different properties of the evaluations presented in the papers. The properties are *research method used*, *context in which the evaluation takes place*, *subjects that take part in the evaluation* and the *scale of the evaluation*. The results in these four properties are all related to the evidence produced and thus related to decision support enabling technology transfer.

6.2.1 RQ2.1: What research methods are most common in the evaluation of technologies?

Looking at the results for research method (see Table 7), it can be seen that the most common research methods used to evaluate technologies are conceptual analysis (39%) followed by case studies (35%). In relation to earlier studies, conceptual analysis has been found to be the most commonly used research method in software engineering [23]. An interesting point is that our study shows a very

Table 7 Research method used in evaluations

Research method	Number	Percentage
Action research	5	5
Case study	34	35
Conceptual analysis/Assertion	38	39
Interview	2	2
Laboratory experiment (human subjects)	14	14
Laboratory experiment (software)	2	2
Lessons learned	2	2
Other	0	0
Descriptive exploratory survey	0	0
Field study	0	0
Conceptual analysis/mathematical	0	0
N/A	0	0
Total	97	100

large number of case studies. The reason might be that this study focuses on evaluations of technologies.

The results also show a relatively high number of experiments with human subjects in relation to previous studies [23].

The research methods used in the majority of the evaluations all provide evidence with different traits. Conceptual analysis is an ad-hoc research method and thus does not provide realistic evaluations or control of the variables studied. This makes it difficult to evaluate and use the evidence produced to make technology transfer decisions. Experiments with human subjects generally have a high level of control of the variables studied. High level of control often amounts to trustworthy evidence and understandability of cause-effect relationships. However, the price of control is often a lack of realism [3, 6, 7], as the setting of the experiment is artificial (often a research laboratory). Case studies offer the possibility of realism if conducted in a real setting (e.g. in industry).

6.2.2 *RQ2.2: In what context are the technologies evaluated?*

The contexts in which the technologies are evaluated were classified as either *Academia* or *Industry*. This classification is made to distinguish the evaluations performed in industry from the ones performed in academia, as one major point of technology transfer is to transfer results produced in one context to another. If the evaluation presented is done in industry, it increases the possibility of transferring it to another industry context [12].

The results for context (see Table 8) show that a significant number of evaluations take place in industry (37%), even though most of the evaluations are carried out in academia (63%).

Table 8 Context of evaluation

Context	Number	Percentage
Academia	61	63
Industry	36	37
Total	97	100.00

6.2.3 *RQ2.3: What subjects are used in the evaluation?*

The similarity of the subjects participating in the evaluation with those envisioned to use the technology affects the ease of technology transfer [19]. With respect to subjects performing the tasks in the evaluation, Table 9 shows that only 19% of the evaluations involve industrial practitioners.

The most common scenario is that in which the researcher him/herself performs the evaluation (68%), for example uses the technology on either invented data or data from industry.

The reason might be that it is difficult to achieve industry trials (they are associated with risk and cost for the company). However, it may also mean that researchers have a hard time convincing the company of the benefits of using the new technology and thus a hard time gaining commitment to participation in the evaluation.

Cost and risk might also influence researchers not to use practitioners in evaluations, as carrying out evaluations in industry introduces higher risks and lead times for the research itself. That is, researchers might not be able to publish studies that do not show positive results and it takes considerable time to follow a real project.

6.2.4 *RQ2.4: What scale do the evaluations have?*

The scale of the evaluations also affects the evidence produced by the evaluations. Some factors might be impossible to evaluate in small-scale evaluations, e.g., scalability [6]. To provide evidence of usefulness in a real setting, evaluations must have a more realistic scale. In this study, the scale of evaluations is investigated by looking at the scale of the applications used in the evaluations.

Table 10 shows that most evaluations consist of applying technologies to *Toy examples* (46%), followed by *Industrial* applications (41%).

Table 9 Subjects in the evaluations

Subjects	Number	Percentage
Not mentioned	1	1
Practitioner	18	19
Researcher	66	68
Student	12	12
Total	97	100

Table 10 Scale of the evaluations

Scale	Number	Percentage
Down-scaled real example	12	12
Industrial	40	41
Toy example	45	46
Total	97	100

6.2.5 *RQ2.5: What degree of realism do the evaluations have?*

The realism offered by the evaluations is evaluated by considering the combination of research method, context, subjects, and scale of the evaluations. These combinations, together with the number of papers that fit into each combination, are presented in Table 11.

The distribution in the table suggests that:

- The most frequently occurring combination of research method, context, subject and scale is conceptual analysis applied to a toy example by the researcher in academia (shown as A in Table 11).
- Of the 34 case studies given in the table, only 8 are carried out in a real setting, utilizing industrial applications and practitioners as subjects (shown as B in Table 11).

Summary and Discussion of RQ2: There seems to be a lack of in-depth studies on the actual use of technologies in real-life evaluations. Even though several studies are presented as case studies, only a few (8 of 34) use practitioners as subjects in a real environment and with industrial scale applications. This lack of reality in these studies could potentially hamper technology transfer, as it makes evidence more difficult to interpret. If papers presenting an evaluation consisting of a technology applied to a toy example by a researcher in academia provides the researcher with as much merit as a case study using practitioners in industry, there is no clear incentive for researchers to carry out the more time consuming research [49].

This is not to say that all evaluations should be case studies performed in industry but rather that different aspects of the technology may require different types of evaluation. For example, the correctness of the results produced by a method or tool might be validated by proofs of correctness or in an experiment in academia. A case study conducted by researchers using, e.g., interviews with practitioners could be a step in an initial evaluation of a technology.

Issues such as the usability and usefulness of a technology, as perceived by potential users, must ultimately be evaluated in industry using practitioners. This is especially relevant regarding usefulness, i.e., the degree to which the

Table 11 Combining research method with context, subject and scale of evaluation

Research method	Context	Subjects	Scale	#	Percentage	
Action research	Academia	Researcher	Toy example	1	1%	
Action research	Industry	Practitioner	Industrial	2	2%	
Action research	Industry	Researcher	Industrial	2	2%	
Conceptual analysis	Academia	Researcher	Toy example	20	21%	A
Conceptual analysis	Academia	Researcher	Down-scaled	7	7%	
Conceptual analysis	Academia	Researcher	Industrial	3	3%	
Conceptual analysis	Academia	Student	Toy example	1	1%	
Conceptual analysis	Industry	Practitioner	Industrial	1	1%	
Conceptual analysis	Industry	Researcher	Down-scaled	2	2%	
Conceptual analysis	Industry	Researcher	Industrial	4	4%	
Lessons learned	Industry	Practitioner	Industrial	1	1%	
Lessons learned	Industry	Researcher	Industrial	1	1%	
Case study	Academia	Researcher	Toy example	8	8%	
Case study	Academia	Researcher	Down-scaled	1	1%	
Case study	Academia	Researcher	Industrial	6	6%	
Case study	Academia	Student	Toy example	2	2%	
Case study	Industry	Practitioner	Industrial	8	8%	B
Case study	Industry	Researcher	Down-scaled	2	2%	
Case study	Industry	Researcher	Industrial	6	6%	
Case study	Industry	NM	Industrial	1	1%	
Laboratory experiment (HS)	Academia	Practitioner	Toy example	1	1%	
Laboratory experiment (HS)	Academia	Researcher	Toy example	1	1%	
Laboratory experiment (HS)	Academia	Student	Toy example	9	9%	C
Laboratory experiment (HS)	Industry	Practitioner	Industrial	3	3%	
Laboratory experiment (SW)	Academia	Researcher	Toy example	1	1%	
Laboratory experiment (SW)	Industry	Researcher	Industrial	1	1%	
Interview	Industry	Practitioner	Toy example	1	1%	
Interview	Industry	Practitioner	Industrial	1	1%	
			Total	**97**	**100%**	

technology adds value when used: in essence, ROI and comparison with the best alternative investment.

Looking at the papers in this review, there does not seem to be a spread; rather, a clear majority of the evaluations was made by researchers (68%) and most made on the scale of downscaled or toy examples ($46 + 12 = 58\%$). This result is not encouraging from a technology transfer perspective, as it means that the decision support material offered to practitioners has its basis far removed from a real setting. This is especially surprising in an engineering field that per its very definition should be applied. At least, the final products of research in the form of technologies should have the goal of being applied.

6.3 RQ3: To what extent does the state of research support the actual adoption of technologies?

To analyze the support for technology transfer offered in the papers, the evidence in Table 11 is mapped to the technology transfer model in Sect. 2.1 and visualized in Fig. 3. The papers are classified as dynamic and static evaluations in industry and evaluations in academia. A distinction is also made between evaluations carried out with practitioners as subjects and without (shown as the top and bottom part of the circles in Fig. 3). Articles classified as dynamic evaluations in industry consist of either case studies, action research, or lessons learned. These should also be applied in industry on industrial scale applications. Other research methods applied in industry are classified as static evaluations. Studies performed in academia are all mapped to evaluations in academia.

Figure 3 shows that 21 evaluations are classified as dynamic evaluations in industry (top left part of Fig. 3), of which 11 were carried out using practitioners as subjects (labeled A in Fig. 3).

As previously noted, the clear majority of the papers presents evaluations performed in academia (shown in the bottom part of Fig. 3). A minority of the papers is found to be static evaluations in industry (shown as C and D in Fig. 3), which is somewhat surprising as it is the natural step before carrying out larger scale evaluations. For evaluations to provide support for technology transfer, they need to offer evidence valued by practitioners and be performed in a relevant environment [1, 2, 12, 13, 19]. This implies that the 11 papers labeled A in Fig. 3 potentially offer the greatest support for technology transfer. These papers are further detailed in Table 12, which shows the degree of description of context, design, and validity of the study.

Looking at Table 12, only 5 papers of the 11 offer somewhat strong support for technology transfer. That is, the context of the study is described to the point where practitioners can evaluate and compare it to their own environment; the study design is described to a degree at which one can understand how the results have been produced and the validity of the study has been discussed to unveil any circumstances that may affect the results. Four papers are classified as being strong in all of these three properties, and one further paper deviates slightly only in one property. On further investigation of the papers, one of the papers presenting RAM [45] was found to describe a static evaluation and was removed from further investigation. Table 13 provides brief descriptions of the papers identified to support technology transfer fully; for detailed information about the studies, the reader is referred to the original papers.

Of the papers in Table 13, one presents more than one case in the evaluation, namely [43], which presents two cases of industry evaluation. However, both [56] and [50] present some information that shows the techniques were used on a wider scale than that presented in the papers. Two of the papers present investigations in a market driven context [43, 44].

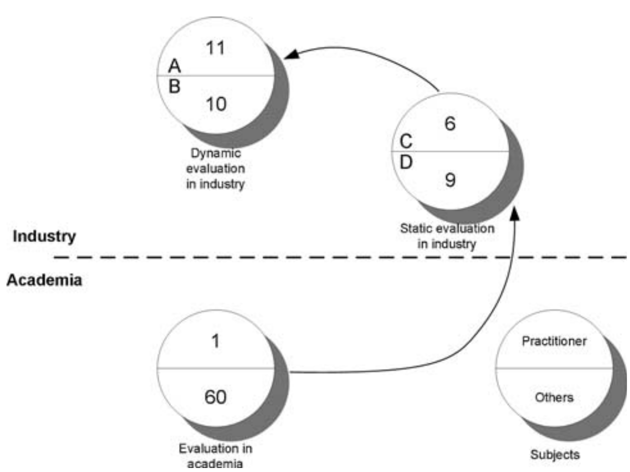

Fig. 3 Mapping of articles to technology transfer model

6.3.1 *RQ3.1: To what degree are the evaluations described?*

Table 14 shows the combination of research methods and the degree to which the context, design and validity of the evaluations are presented in all of the papers.

The distribution shown in Table 14 indicates that:

- Most of the papers utilizing conceptual analysis as a research method do not describe how the results are produced or their validity (labeled A in Table 14). This is expected, as there is no real reason to present these aspects in a basic evaluation. This also means that the evaluations offer slim decision support for technology transfer.
- Most of the case studies do not score strongly in all three properties in terms of the way in which the study

Table 12 Dynamic validation in industry (depicted A in Fig. 3)

Article	Technologies	Research method	Context described	Study design	Validity discussed
[50]	Scenarios, usability tests	Action research	Strong	Strong	Strong
[51]	Finite state machine	Action research	Strong	Medium	Weak
[45]	Requirements Abstraction Model (RAM)	Lessons learned	Strong	Strong	Medium
[52]	Participatory design	Case study	Medium	Weak	Weak
[44]	AHP, Disagreement charts, Distribution charts, Influence charts, Satisfaction charts	Case study	Strong	Strong	Strong
[36]	Language extended lexicon (LEL)	Case study	Strong	Medium	Weak
[53]	RM-Tool	Case study	Weak	Medium	Weak
[54]	MAGERIT, SIREN	Case study	Weak	Weak	Weak
[55]	Misuse oriented quality requirements engineering (MOQARE)	Case study	Weak	Medium	Weak
[43]	Requirements Abstraction Model (RAM)	Case study	Strong	Strong	Strong
[56]	Knowledge based Approach for the Selection of Requirements Engineering Techniques (KASRET)	Case study	Strong	Strong	Strong

Table 13 Brief description of papers found to support technology transfer

	Context	# of cases	Brief description of technologies	Results in short
[50]	A Danish manufacturer of sound and vibration measurement	1	Scenarios, including description of user tasks and usability tests with daily user tasks based on screen mock-ups	The number of usability problems per screen is reduced by 70% and the number of implementation defects per month is reduced by more than 20%. Sales figures showed that the product sells twice as many units and at twice the unit price
[44]	A Swedish CASE tool developer. Geographically distributed stakeholders that take part in prioritizing high-level requirements	1	Four different charts for visualizing and interpreting prioritization data	The distributed prioritization is useful and the visualization charts are a valuable decision support
[43]	A Swedish software and hardware developer of navigation, control, fleet management and service for automated guided vehicles. An international developer of power and automation technologies	2	Requirements Abstraction Model (RAM), a model for working with requirements on several levels of abstraction	Indications of that implementation of RAM has yielded substantial increases in both the accuracy of practices performed in requirements engineering and in requirements quality
[56]	A company developing real-time supply chain management solutions to optimize power networks	1	Knowledge-based Approach for the Selection of Requirements Engineering Techniques (KASRET), a decision support system for the selection of requirements engineering techniques for a specific software project	The majority of the projects members appreciated the high quality requirements gained from using the recommended techniques. The techniques selected were very helpful in reducing the delay of the project and improving the overall quality of the software specification

is described. Only four papers (labeled B in Table 14) of 34 presenting case studies also presented the context, study design and validity of the study.

- Most of the experiments carried out with human subjects score strongly in all three properties for the level on which the study is described (labeled D in Table 14).

Table 14 Research method combined with the description of the evaluation

Research method	Context described	Study design	Validity discussed	#	Percentage	
Action research	Strong	Strong	Strong	1	1%	
Action research	Strong	Medium	Weak	1	1%	
Action research	Medium	Strong	Weak	1	1%	
Action research	Weak	Strong	Weak	1	1%	
Action research	Weak	Weak	Medium	1	1%	
Conceptual analysis	Medium	Weak	Medium	1	1%	
Conceptual analysis	Medium	Weak	Weak	5	5%	
Conceptual analysis	Weak	Weak	Weak	32	33%	A
Lessons learned	Strong	Strong	Medium	1	1%	
Lessons learned	Medium	Weak	Weak	1	1%	
Case study	Strong	Strong	Strong	4	4%	B
Case study	Strong	Medium	Weak	1	1%	C
Case study	Strong	Weak	Medium	1	1%	C
Case study	Medium	Medium	Weak	2	2%	C
Case study	Medium	Weak	Medium	2	2%	C
Case study	Medium	Weak	Weak	10	10%	C
Case study	Weak	Medium	Medium	1	1%	C
Case study	Weak	Medium	Weak	3	3%	C
Case study	Weak	Weak	Weak	10	10%	C
Laboratory experiment (HS)	Strong	Strong	Strong	7	7%	D
Laboratory experiment (HS)	Strong	Strong	Medium	3	3%	
Laboratory experiment (HS)	Strong	Medium	Weak	1	1%	
Laboratory experiment (HS)	Medium	Strong	Weak	1	1%	
Laboratory experiment (HS)	Medium	Medium	Medium	1	1%	
Laboratory experiment (HS)	Weak	Medium	Weak	1	1%	
Laboratory experiment (SW)	Strong	Strong	Medium	1	1%	
Laboratory experiment (SW)	Medium	Strong	Weak	1	1%	
Interview	Strong	Strong	Strong	1	1%	
Interview	Weak	Weak	Weak	1	1%	
Total				**97**	**100%**	

Summary and Discussion of RQ3: To summarize the findings, the technologies identified in Sect. 6.1 as receiving most attention (Language extended lexicon (LEL) [34–37] and Use cases [38–42]), in terms of leading to several publications on the same technology, are not found to be among the four strongest papers in relation to supporting technology transfer. Thus, it cannot be assumed that more attention on the part of researchers automatically translates to stronger technology transfer support in terms of more and better evaluations, and a better description of evaluation design, context and validity discussions. This is surprising, as more papers dealing with the same technology could be assumed to be an evolution in terms of more and better technology transfer support. However, the papers presenting Language extended lexicon (LEL) do provide a gradual refinement in terms of evaluation, as all three type of evaluations (see Fig. 3) are covered by the four papers [34–37].

Only 4 of the 97 articles investigated in this study are found to offer strong evidence that supports technology transfer.

With respect to the way evaluations are presented, almost half do not describe how the evaluation was carried out, the context or the validity of the study. Only about one in ten studies has a clear description of all three aspects. This might not be remarkable in the case of general research papers, but it is important to remember that the papers investigated in this review are selected on the basis of the fact that they explicitly evaluate technologies. Furthermore, they are journal papers.

The research method that scores the highest in all three properties is laboratory experiments with human subjects. This may be because that the research field has some established guidelines for how to conduct and report experiments, e.g. [31] and [57].

6.4 The secondary studies

The classification process used facilitates an identification of all evaluations presented in the papers. The results presented so far have focused on the primary studies in the papers in order to make interpretation of the results easier for the reader. Studies are classified as primary if they are the focus of the paper, or if that cannot be determined, the study presenting the highest level of support for technology transfer is selected as the primary evaluation. Identifying all evaluations reported in the papers is important, as the purpose of the review is to identify support for technology transfer. All evaluations presented are candidates for providing technology transfer support. Thus, for completeness, this section presents the studies classified as not being primary in the review. Twenty-eight secondary evaluations were identified in the review and are presented in Table 15.

Table 15 shows the research method, context, subjects and scale of the secondary studies. The studies are reported in 21 different papers. Ten studies are candidates for providing technology transfer support and thus influence the results given in Sect. 6.3, as the research method used is either case study, action research or lessons learned. However, on closer investigation, only one study would influence the results presented for RQ3, as this was carried out in a real environment with real subjects and an application of industrial scale. This study is the second case study presented in [43], which is noted in Table 13.

7 Conclusions

This paper presents a systematic literature review that examines all of the papers published in Requirements Engineering journal that contain any type of technology evaluation. The aim is to gauge the support for technology transfer (see the definitions in Sect. 2.1), i.e., the degree to which industrial practitioners can use the reporting of technology evaluations as decision support for adopting the technologies in industrial practice. This is done by evaluating the type of research performed in terms of research methodology, who performs the evaluations, where the evaluations are performed, and to what extent the design, context, and validity are described.

The major findings of the review are:

- Only 4 of the 97 papers investigated provide strong technology transfer support (see Sect. 6.3). This was rather unexpected, as requirements engineering is, as the name indicates, an engineering field. In addition, the papers investigated are journal publications, which would indicate mature research with a focus on validations of some sort. On a possible positive note, all the papers that provide strong technology transfer support were written in the last 7 years, which might indicate a positive trend.
- The majority of evaluations were performed by the researcher him/herself (see Sects. 6.2.3, 6.2.5). This limits the realism provided in evaluations and hence the support offered for technology transfer.
- As concerns the technologies presented, few have a pre-project timeline focus (see Sect. 6.1.2). Most technologies presented are meant to be used when the understanding of the system under consideration has become rather mature. Technologies used to acquire

Table 15 Classification of secondary studies

Research method	Context	Subjects	Scale	Number
Conceptual analysis	Academia	Researcher	Toy example	8
Conceptual analysis	Academia	Student	Toy example	1
Conceptual analysis	Academia	Researcher	Down-scaled real example	1
Case study	Academia	Researcher	Industrial	2
Case study	Academia	Student	Toy example	3
Case study	Industry	Practitioner	Industrial	1
Case study	Academia	Researcher	Toy example	1
Case study	Industry	Student	Industrial	1
Lessons learned	Academia	Researcher	Toy example	1
Laboratory experiment (HS)	Industry	Practitioner	Industrial	2
Laboratory experiment (HS)	Academia	Student	Toy example	1
Laboratory experiment (HS)	Academia	Student	Industrial	1
Laboratory experiment (SW)	Academia	Researcher	Toy example	1
Interview	Industry	Practitioner	Industrial	2
Action research	Industry	Researcher	Industrial	1
Survey	Industry	Practitioner	Not mentioned	1
Total				28

this understanding seem to be somewhat overlooked, thus potentially limiting the value of the technologies presented. In addition, none of the technologies investigated is meant to be used after a project has concluded (see Sect. 6.1.2). This may have implications for the progression in research, as there are, e.g., few established ways to evaluate requirements engineering performance.

- With regard to the presentation of the studies, experiments with human subjects are on the whole well described (see Sect. 6.3.1). However, case studies that have the potential to provide more realism in evaluations are rather poorly described. Presenting the context, study design, and validity of evaluation is important in case studies, as this aids practitioners in deciding, whether the results are transferable to another setting, and also evaluate the overall validity of any claims and conclusions.
- The scale on which the evaluations were performed is also an important aspect. Forty-one percent were on an industrial scale, which would be positive from a technology transfer support perspective. However, this means that the rest (59%) were on a downscaled or toy example level. This can be explained to some extent in terms of an introduction of a new technology perhaps being best illustrated using a simple, small-scale example. Two factors might counter this argument, however. The first is that the papers investigated contained technology evaluations. Papers that purely presented a new technology were not included in the review at all. Second, as only journal publications were reviewed, one might argue that the maturity of the technology presented would have passed the "yet another technology" stage, where toy examples are the norm.

The aspects of technology transfer, and decision support for practitioners to adopt technologies that are described and defined in this review are naturally somewhat simplified. Many aspects not covered, such as attaining management support and initiation threshold, are also relevant. However, the perspectives covered in this paper aim to describe a minimum level of support factors. These factors (e.g. description of context, design, validity), and the way in which the evaluations were performed in terms of subjects and scalability, are critical from both an academic and industry perspective. Thus, the findings of this review do have a number of implications for research:

- The major potential for improvement as concerns technology transfer support seems to be the subjects used in the evaluations. Not using industrial practitioners as subjects excludes several important aspects of an evaluation. First, successfully convincing industry of participating in an evaluation speaks to the perceived relative advantage of a technology. Second, it also involves training practitioners, which highlights issues having to do with usability. There is an apparent risk of missing addressing these issues if researchers themselves make the evaluations.
- Regarding the time-line focus of the technologies, there is a tendency to focus on technologies to be used when projects have already started (in-project). The implication for research is whether this reflects the actual needs in industry. Looking at the studies identified in this review as providing the strongest technology transfer decision support, half have a more complete view, looking at pre-project requirements engineering also. The question is whether the industry commitment for validation and piloting was obtained as industry problems were better addresses by these studies.
- The results indicate that empirical evidence from realistic evaluations in industry are scarce, but even more significant is the level of description of study design, context and validity as this is within the control of the researcher and not dependent on hard to obtain external resources. This limits the value of the research from an adoption perspective. Using guidelines such as those presented by Kitchenham [57] could improve the decision support material offered to practitioners. This would also be beneficial from an academic perspective as it improves the credibility of the research, and the ability to gauge the value of the contribution.
- The scale of the evaluation is predominantly toy or down-scaled examples. The implication for research is that it is hard to judge whether the technology scales in relation to applicability in practice. Performing evaluations on a more realistic scale could enable better technology decision support material but also mean that unrealistic solutions are discarded earlier [6].

Looking at the evaluations, it is very difficult to attain good support for technology transfer by utilizing only one research method or, for that matter, making only one evaluation. A combination of methods, such as an experiment for control and a case study in industry for scalability and realism, would offer a combination of evaluation results superior to any single evaluation (see Fig. 3). This combinatory approach of incremental evaluation and technology transfer would not only increase the support for technology transfer but would also improve the evaluations performed, both from a research and industry adoption perspective.

Appendix

Technologies identified in the review

Technology	#	Technology	#
AHP	2	Problem frame	2
Cost-value approach	1	Alloy	1
Albert II	1	Promela	2
Analogy-based domain analysis (ABDA)	1	Rationale	1
Disagreement chart	1	Relationship analysis	1
Distribution chart	1	Release planning	N/A
Domain model	1	Release Planner Provotype (RPP)	1
Multi-Agent Domain Engineering Methodology (MADEM)	1	RENISYS	1
Domain reduction abstraction	1	Requirements Abstraction Model (RAM)	2
EPMcreate (EPM Creative Requirements Engineering TEchnique)	1	Requirements generation model	1
Entity relationship model	1	Requirements Similarity Analysis	1
Facilitator	1	Retrenchment	1
Co-operative Requirements Capture (CRC)	1	RM-Tool	1
Fault tree analysis	1	Role activity diagram (RAD)	1
Feature dependencies	1	SAP Goal Map	1
Finite state machine (FSM)	1	Satisfaction chart	1
Fitness relationship	1	Scenarios	2
Gap typology	1	Goal-Based Requirements Analysis Method	1
Generating Natural Language from UML (GeNLangUML)	1	Immersive scenario-based requirements engineering (ISRE) method	1
Goal-orientation	N/A	Language extended lexicon (LEL)	4
Assumption/commitment style reasoning	1	Requirements Analysis Support System (NDRASS)	1
Attributed goal oriented requirements analysis (AGORA)	1	scenario-based requirements analysis method	1
NFR framework	3	PC-RE	1
Grounded theory	1	System reliability analyser (SRA)	1
I* framework	1	Timed automata	1
Analysis of Web Application Requirements (AWARE)	1	User manual as requirements specification	1
Formal Tropos	1	Set theory	2
Requirements engineering framework (REF)	1	SIREN	1
RESCUE	1	Soft Systems Methodology (SSM)	N/A
Scenario advisor tool	1	SISTeM	1
TCD approach	1	SPIN	2
IDEFØ activity model	1	Arcade	1
Influence chart	1	State machine hazard analysis	1

Table continued

Technology	#	Technology	#
Inspection	N/A	State-charts	2
Defect-based reading	1	STATEMATE	1
N-fold inspection	1	Viewcharts	1
Issue-Based Information Systems (IBIS)	1	UML	1
KAOS	N/A	Usability tests	1
Obstacle analysis	1	Use cases	5
ReqMon	1	MisUse Case	N/A
Knowledge-based Approach for the Selection of Requirements Engineering Techniques (KASRET)	1	Misuse-oriented quality requirements engineering (MOQARE)	1
LOTOS	1	Requirements pattern language (RPL)	1
MAGERIT	1	SCORES	1
MemReq	1	Use case map	1
MetaEdit+	1	Use case statecharts	1
Mismatch Handling for COTS Selection	2	User interface and requirements prototyping (UIRPing)	1
Mismatch Handling for COTS Selection Sensitivity Analysis (MiHOS-SA)	1	Venn diagram	1
Multi level feature trees	1	View-point	N/A
Multimedia specification process model	1	MultiSpec	1
NFR-classifier	1	RECOCASE	1
Object oriented analysis	2	Viewpoint-Oriented Requirements Definition method (VORD)	1
Object-Z	N/A	WinWin negotiation	N/A
Formal Object-Oriented Method (FOOM)	1	EasyWinWin negotiation	N/A
Participatory design	1	Negotiation model checking	1
Problem dependency map	1	Z	1
		CADiZ	1

References

1. Pfleeger SL (1999) Understanding and improving technology transfer in software engineering. J Syst Softw 47(2–3):111–124. doi:10.1016/S0164-1212(99)00031-X
2. Pfleeger SL, Menezes W (2000) Marketing technology to software practitioners. IEEE Softw 17(1):27–33. doi:10.1109/52.819965
3. Sjoberg DIK, Anda B, Arisholm E et al (2002) Conducting realistic experiments in software engineering. In: Proceedings of the 18th international symposium on empirical software engineering, pp 17–26
4. Hsia P, Davis AM, Kung DC (1993) Status report: requirements engineering. IEEE Softw 10(6):75–79. doi:10.1109/52.241974
5. Neill CJ, Laplante PA (2003) Requirements engineering: the state of the practice. IEEE Softw 20(6):40–45. doi:10.1109/MS.2003.1241365
6. Potts C (1993) Software-engineering research revisited. IEEE Softw 10(5):19–28. doi:10.1109/52.232392
7. Glass RL (1994) The software-research crisis. IEEE Softw 11(6):42–47. doi:10.1109/52.329400
8. Zelkowitz MV, Wallace D (1997) Experimental validation in software engineering. Inf Softw Technol 39(11):735–743. doi:10.1016/S0950-5849(97)00025-6
9. Tichy WF, Lukowicz P, Prechelt L et al (1995) Experimental evaluation in computer science: a quantitative study. J Syst Softw 28(1):9–18. doi:10.1016/0164-1212(94)00111-Y
10. Maiden N (2005) What has requirements research ever done for us? IEEE Softw 22(4):104–105. doi:10.1109/MS.2005.113
11. Kitchenham B, Charters S (2007) Guidelines for performing systematic literature reviews in software engineering. Technical report EBSE 2007-001, Keele University and Durham University Joint Report
12. Guba EG, Lincoln YS (1989) Fourth generation evaluation. Sage, Newbury Park
13. Zelkowitz MV, Wallace DR, Binkley DW (1998) Culture Conflicts in Software Engineering Technology Transfer. Tech. report, University of Maryland, College Park

14. Gorschek T, Garre P, Larsson S et al (2006) A model for technology transfer in practice. IEEE Softw 23(6):88–95. doi:10.1109/MS.2006.147
15. Kaindl H, Brinkkemper S, Bubenko JA et al (2002) Requirements engineering and technology transfer: obstacles, incentives and improvement agenda. Requir Eng 7(3):113–123. doi:10.1007/s007660200008
16. Redwine ST, Riddle WE (1985) Software technology maturation. In: Proceedings of the 8th international conference on Software engineering, IEEE Computer Society Press, London, England, pp 189–200
17. Raghavan SA, Chand DR (1989) Diffusing software-engineering methods. IEEE Softw 6(4):81–90. doi:10.1109/52.31655
18. Zelkowitz MV (1996) Software engineering technology infusion within NASA. IEEE Trans Eng Manage 43(3):250–261. doi:10.1109/17.511836
19. Rogers EM (1995) Diffusion of Innovations. Free Press, New York
20. Wohlin C, Höst M, Mattsson C (1996) A framework for technology introduction in software organizations. In: Proceedings of the software process improvement conference, Brighton, UK, pp 167–176
21. Linkman S, Rombach HD (1997) Experimentation as a vehicle for software technology transfer—a family of software reading techniques. Inf Softw Technol 39(11):777–780. doi:10.1016/S0950-5849(97)00029-3
22. Kitchenham B, Pearl Brereton O, Budgen D et al (2009) Systematic literature reviews in software engineering—a systematic literature review. Inf Softw Technol 51(1):7–15. doi:10.1016/j.infsof.2008.09.009
23. Glass RL, Vessey I, Ramesh V (2002) Research in software engineering: an analysis of the literature. Inf Softw Technol 44(8):491–506. doi:10.1016/S0950-5849(02)00049-6
24. Jørgensen M (2007) Estimation of software development work effort: evidence on expert judgment and formal models. Int J Forecast 23(3):449–462. doi:10.1016/j.ijforecast.2007.05.008
25. Wieringa R, Maiden N, Mead N et al (2006) Requirements engineering paper classification and evaluation criteria: a proposal and a discussion. Requir Eng 11(1):102–107. doi:10.1007/s00766-005-0021-6
26. Wieringa R, Heerkens J (2006) The methodological soundness of requirements engineering papers: a conceptual framework and two case studies. Requir Eng 11(4):295–307. doi:10.1007/s00766-006-0037-6
27. Davis A, Dieste O, Hickey A, et al (2006) Effectiveness of requirements elicitation techniques: empirical results derived from a systematic review. In: Proceedings of the 14th IEEE international requirements engineering conference (RE'06), IEEE Computer Society, pp 176-185
28. Parviainen P, Tihinen M (2007) A survey of existing requirements engineering technologies and their coverage. Int J Softw Eng Knowl Eng 17(6):827–850. doi:10.1142/S0218194007003513
29. Dyba T, Dingsoyr T, Hanssen GK (2007) Applying systematic reviews to diverse study types: an experience report. In: Proceedings of the first international symposium on empirical software engineering and measurement (ESEM), pp 225–234
30. IEEE Computer Society (2004) SWEBOK. IEEE, Los Alamitos
31. Wohlin C, Runeson P, Höst M et al (2000) Experimentation in software engineering. Kluwer, Boston
32. Maiden N, Manning S, Jones S et al (2005) Generating requirements from systems models using patterns: a case study. Requir Eng 10(4):276–288. doi:10.1007/s00766-005-0010-9
33. Robson C (2002) Real world research. Blackwell, Cornwall
34. Leite J, Rossi G, Balaguer F et al (1997) Enhancing a requirements baseline with scenarios. Requir Eng 2(4):184–198. doi:10.1007/BF02745371
35. Leite J, Hadad GDS, Doorn JH et al (2000) A scenario construction process. Requir Eng 5(1):38–61. doi:10.1007/PL00010342
36. Cysneiros LM, Leite J, t Neto J (2001) A framework for integrating non-functional requirements into conceptual models. Requir Eng 6(2):97–115. doi:10.1007/s007660170008
37. Leite J, Doorn JH, Hadad GDS et al (2005) Scenario inspections. Requir Eng 10(1):1–21. doi:10.1007/s00766-003-0186-9
38. Juric R, Kuljis J (1999) Engineering requirements through use cases in complex business environment. Requir Eng 4(2):65–76. doi:10.1007/s007660050013
39. Dutoit AH, Paech B (2002) Rationale-based use case specification. Requir Eng 7(1):3–19. doi:10.1007/s007660200001
40. Richards D (2003) Merging individual conceptual models of requirements. Requir Eng 8(4):195–205. doi:10.1007/s00766-002-0158-5
41. Siau K, Lee L (2004) Are use case and class diagrams complementary in requirements analysis? An experimental study on use case and class diagrams in UML. Requir Eng 9(4):229–237. doi:10.1007/s00766-004-0203-7
42. Saiedian H, Kumarakulasingam P, Anan M (2005) Scenario-based requirements analysis techniques for real-time software systems: a comparative evaluation. Requir Eng 10(1):22–33. doi:10.1007/s00766-004-0192-6
43. Gorschek T, Garre P, Larsson S et al (2007) Industry evaluation of the requirements abstraction model. Requir Eng 12(3):163–190. doi:10.1007/s00766-007-0047-z
44. Regnell B, Höst M, och Dag JN et al (2001) An industrial case study on distributed prioritisation in market-driven requirements engineering for packaged software. Requir Eng 6(1):51–62. doi:10.1007/s007660170015
45. Gorschek T, Wohlin C (2006) Requirements abstraction model. Requir Eng 11(1):79–101
46. Aurum A, Wohlin C (2005) Engineering and managing software requirements. Springer, Heidelberg
47. Brinkkemper S (2004) Requirements engineering research the industry is (and is not) waiting for. In: Proceedings of the 10th anniversary international workshop on requirements engineering: foundation for software quality (REFSQ'04), Riga, Latvia, pp 41–54
48. Karlsson L, Dahlstedt AG, Regnell B et al (2007) Requirements engineering challenges in market-driven software development—An interview study with practitioners. Inf Softw Technol 49(6):588–604. doi:10.1016/j.infsof.2007.02.008
49. Parnas DL (2007) Stop the numbers game. Commun ACM 50(11):19–21. doi:10.1145/1297797.1297815
50. Lauesen S, Vinter O (2001) Preventing requirement defects: an experiment in process improvement. Requir Eng 6(1):37–50. doi:10.1007/PL00010355
51. Breen M (2005) Experience of using a lightweight formal specification method for a commercial embedded system product line. Requir Eng 10(2):161–172. doi:10.1007/s00766-004-0209-1
52. Cherry C, Macredie RD (1999) The importance of context in information system design: an assessment of participatory design. Requir Eng 4(2):103–114. doi:10.1007/s007660050017
53. Lang M, Duggan J (2001) A tool to support collaborative software requirements management. Requir Eng 6(3):161–172. doi:10.1007/s007660170002
54. Toval A, Nicolás J, Moros B et al (2002) Requirements reuse for improving information systems security: a practitioner's approach. Requir Eng 6(4):205–219. doi:10.1007/PL00010360
55. Herrmann A, Paech B (2008) MOQARE: misuse-oriented quality requirements engineering. Requir Eng 13(1):73–86. doi:10.1007/s00766-007-0058-9

56. Jiang L, Eberlein A, Far B (2008) A case study validation of a knowledge-based approach for the selection of requirements engineering techniques. Requir Eng 13(2):117–146. doi:10.1007/s00766-007-0060-2

57. Kitchenham BA, Pfleeger SL, Pickard LM et al (2002) Preliminary guidelines for empirical research in software engineering. IEEE Trans Softw Eng 28(8):721–734. doi:10.1109/TSE.2002.1027796