THESIS FOR THE DEGREE OF DOCTOR OF PHILOSOPHY

Wicked Problems in Engineering Education

Preparing Future Engineers to Work for Sustainability

JOHANNA LÖNNGREN

Department of Communication and Learning in Science

CHALMERS UNIVERSITY OF TECHNOLOGY

Gothenburg, Sweden 2017

Wicked Problems in Engineering Education
Preparing Future Engineers to Work for Sustainability
JOHANNA LÖNNGREN
ISBN 978-91-7597-602-0

Doktorsavhandlingar vid Chalmers tekniska högskola
Ny serie nr 4283
ISSN 0346-718X

Department of Communication and Learning in Science
Chalmers University of Technology
SE-412 96 Gothenburg
Sweden
Telephone + 46 (0)31-772 1000

Cover:
Addressing wicked problems requires an *integrative* approach. In such an approach, problem parts (dark red) and improvement measures (light green) are viewed as interconnected in complex ways. The problem is addressed through an iterative process that aims to improve an unsatisfactory situation, while all the time considering stakeholders' different values and interests and monitoring and anticipating changing circumstances and unintended consequences (see pp. 53, 55, and 57 for more information).

Reproservice Chalmers
Gothenburg, Sweden 2017

Wicked Problems in Engineering Education

Preparing Future Engineers to Work for Sustainability

JOHANNA LÖNNGREN

Department of Communication and Learning in Science

Chalmers University of Technology

ABSTRACT

Most engineering education today does not adequately prepare students to contribute to sustainability. For example, engineering students often do not learn how to address complex and ill-structured sustainability problems that involve different stakeholders, value conflicts, and uncertainty; such problems are also called *wicked problems*. Efforts to improve engineering education in this regard are hampered by a lack of research on how engineering education can prepare students to address wicked problems.

This thesis aims to address this gap in two parts. The research described in Part 1 aimed to explore what engineering students need to learn to be able to address wicked problems. For this purpose, a pre-study literature review and two empirical studies were conducted. For the empirical studies, engineering students were interviewed and the interviews were analyzed using qualitative content analysis (Study 1) and a phenomenographic approach (Study 2). The research in Part 2 aimed to link the theoretical results from Part 1 to engineering education practice by focusing on teaching and assessment. The research in Part 2 comprises two empirical studies in which pragmatic action research (Study 3) and design-based research (Study 4) was used.

The results of the research include (a) a description of engineering education-specific challenges in addressing wicked problems; (b) 3 descriptions of wicked problems and design principles for wicked problem descriptions; (c) description of four different approaches that engineering students have used in addressing a wicked problem; (d) 22 intended learning outcomes, 3 assessment approaches, an analytic assessment rubric, and a rubric-based intervention for students' ability to integratively address wicked problems; (e) validity, reliability, and utility evaluations of the assessment rubric; and (f) insights about students' performance, their approaches to wicked problems, and affordances for learning in differently scaffolded activities during the rubric-based intervention.

Conclusions from the research include that an *integrative* approach to wicked problems is most appropriate, that students are *able* to use such an approach, but that they may need *instructional support* to do so. Conclusions further include that strong cognitive scaffolding with a highly detailed assessment rubric can support students' understanding of the nature of wicked problems and students' performance in written responses to wicked problems, but possibly also limit affordances for deep and transferable learning.

Keywords:
wicked problems, ill-structured problems, problem-solving, engineering education, sustainability, assessment, rubric, phenomenography, action research, design-based research

Wicked problems i Ingenjörsutbildningen

Att förbereda blivande ingenjörer för att arbeta för hållbarhet

JOHANNA LÖNNGREN

Institutionen för Vetenskapens kommunikation och lärande

Chalmers tekniska högskola

SAMMANFATTNING

Många ingenjörsutbildningar brister idag i hur väl de förbereder sina studenter för att arbeta för hållbarhet. Exempelvis får många ingenjörsstudenter otillräcklig träning i att hantera komplexa hållbarhetsproblem som karakteriseras av värdekonflikter mellan intressenter och en stor grad av osäkerhet. Sådana problem kallas på engelska *wicked problems*. Det saknas idag forskning om hur ingenjörsutbildningar kan förbereda sina studenter för att hantera wicked problems, vilket gör det svårt för universitet och lärare att förbättra utbildningen i det avseendet.

Den här avhandlingen består av två delar. Forskningen som beskrivs i Del 1 syftade till att undersöka vad ingenjörsstudenter behöver lära sig för att kunna hantera wicked problems. Detta gjordes genom en litteraturbaserad förstudie och två empiriska studier. De empiriska studierna baserades på intervjuer med ingenjörsstudenter. Intervjuerna analyserades genom kvalitativ innehållsanalys (Studie 1) och fenomenografi (Studie 2). Forskningen i Del 2 syftade till att länka de teoretiska resultaten från Del 1 till utbildningspraktiken genom att fokusera på undervisning och bedömning. Forskningen i Del 2 består av två empiriska studier där pragmatisk aktionsforskning (Studie 3) och design-baserad forskning (Studie 4) användes.

Forskningsresultaten omfattar (a) en beskrivning av ingenjörsstudenters specifika utmaningar i att hantera wicked problems; (c) 3 beskrivningar av wicked problems och designprinciper för att beskriva wicked problems; (c) beskrivning av fyra sätt på vilka ingenjörsstudenter har hanterat ett wicked problem; (d) 22 lärandemål, 3 bedömningsmetoder, en bedömningsmatris och en utbildningsmodul för att utveckla ingenjörsstudenters förmåga att hantera wicked problems; (e) utvärdering av validitet, reliabilitet och användbarhet av bedömningsmatrisen; och (f) insikter om studenters prestationer, deras sätt att ta sig an wicked problems och lärandemöjligheter som uppstod när studenter fick olika typer av stöd i en utbildningsmodul.

En rad slutsatser kan dras från dessa resultat, bland annat att en *integrativ* ansats är lämplig för att hantera wicked problems, att studenter ofta *har förmågan* att använda en sådan ansats, men att de kan behöva *undervisningsstöd* för att faktiskt göra det. En annan viktig slutsats är att *starkt kognitivt stöd* i form av en detaljerad bedömningsmatris kan stödja studenternas förståelse av vad som karakteriserar wicked problems och därmed leda till bättre prestationer när studenter ombeds att hantera wicked problems i skriftliga prov, men att samma kognitiva stöd också kan minska möjligheter för djupinlärning.

Nyckelord:
Wicked problems, ostrukturerade problem, problemlösning, ingenjörsutbildning, hållbarhet, bedömning, bedömningsmatris, fenomenografi, aktionsforskning, design-baserad forskning

List of publications

This thesis is based on the work that is reported in the following papers:

Paper I: Lönngren, J., Svanström, M., Ingerman, Å., & Holmberg, J. (2016). Dealing with the multidimensionality of sustainability through the use of multiple perspectives – a theoretical framework. *European Journal of Engineering Education*, *41* (3), pp. 342-352.

Paper II: Lönngren, J., Ingerman, Å., & Svanström, M. (2017). Avoid, Control, Succumb, or Balance: Engineering Students' Approaches to a Wicked Sustainability Problem. *Research in Science Education*, *47*, pp. 805-831.

Paper III: Lönngren, J., & Svanström, M. (2015). Assessing "Wicked Sustainability Problem"-Literacy in Engineering Education. *Proceedings of the 122nd ASEE Annual Conference & Exhibition.* Seattle, USA, 14-17 June. American Society for Engineering Education.[1]

Paper IV: Lönngren, J., Adawi, T., & Svanström, M. (2017). Wicked problems and assessment in engineering education: Developing and evaluating an analytic rubric. *Proceedings of the 7th Research in Engineering Education Symposium.* Bogota, Colombia, 6-8 July.[2]

Paper V: Lönngren, J., Adawi, T., & Svanström, M. (n.d). Scaffolding strategies in a rubric-based intervention to promote engineering students' ability to address wicked problems. *Resubmitted to scientific journal after minor revisions.*

---

[1] Peer reviewed and awarded a prize for the best graduate student paper of the year in the Environmental Engineering Division of the American Society for Engineering Education.

[2] Peer reviewed.

Contribution report

Paper I: Lead author. Planned data collection together with co-authors. Independently conducted interviews and performed data analysis. Developed framework together with co-authors.

Paper II: Lead author. Planned data collection and analysis together with co-authors. Independently conducted interviews and performed most parts of the data analysis. Developed phenomenographic outcome space together with co-authors.

Paper III: Lead author. Planned data collection together with co-authors. Independently conducted interviews. Facilitated the workshops with support from co-author. Developed learning outcomes and descriptions of assessment activities together with co-author.

Paper IV: Lead author. Developed rubric and intervention together with co-authors. Planned data collection together with co-authors. Implemented intervention together with one co-author. Independently collected and analyzed most empirical material. Developed conclusions from results together with co-authors.

Paper V: Lead author. Developed rubric and intervention together with co-authors. Planned data collection together with co-authors. Implemented intervention together with one co-author. Independently collected and analyzed most empirical material. Developed conclusions from results together with co-authors.

Throughout the thesis, I use the first person singular pronoun, “I”, to indicate that the work described is primarily my own. I use the first person plural pronoun, “we”, to indicate that the work has been done in collaboration between the researchers who were involved in the specific study or paper (as indicated in the above list of publications). Consequently, “we” refers to different groups of researchers in different parts of the thesis.

Other relevant publications

In addition to the included papers, the following publications are relevant to the research reported in this thesis:

Paper A: Lönngren, J. (2014). *Engineering Students' Ways of Relating to Wicked Sustainability Problems.* Licentiate thesis. Chalmers University of Technology, Department of Applied IT. Gothenburg: Chalmers.

Paper B: Lönngren, J., Ahrens, A., Deppert, K., Hammarin, G., & Nilsson, E. (2010). Sustainable Development in Nano-Perspectives – An Innovative Student Initiative. *Proceedings of 5th Engineering Education in Sustainable Development Conference*. Gothenburg, Sweden, 19-22 September.

Paper C: Lönngren, J., Jacobsson, D., Mårsell, E., & Nilsson, E. (2012). Breaking Catch-22 of Engineering Education for Sustainable Development: An Example of Parallell Learning of Teachers and Students. *Improving Student Learning Through Research and Scholarship: 20 Years of ISL*. Lund, Sweden, 29-31 August, p. 28. Oxford Centre for Staff and Learning Development.

Paper D: Lönngren, J., Svanström, M. (2012). Navigating the maze of teaching and learning for sustainable development in engineering education–"Perspective shift" in relation to other key competences. *Proceedings of the XVth IOSTE International Symposium-Science & Technology Education for Development, Citizenship and Social Justice.* La Medine, Tunisia, 28 October – 3 November.

Paper E: Lönngren, J. & Svanström, M. (2016). Systems thinking for dealing with wicked sustainability problems: Beyond functionalist approaches. In W.L Filho & S. Nesbit (eds.) *New Developments in Engineering Education for Sustainable Development*. First edition. World Sustainability Series. Springer International Publishing.

Paper F: Zhang, Z., Fyn, D., Langelotz, L., Lönngren, J., McCorquodale, L., & Nehez, J. (2014). Our way(s) to action research: Doctoral students' international and interdisciplinary collective memory work. *Action Research*, *12* (3), pp. 293-314.

Abbreviations used in the thesis

AQ	additional question (part of the mid-term exam in Study 4)
Chalmers	Chalmers University of Technology
EER	engineering education research
ESE	environmental and sustainability education
ESER	environmental and sustainability education research
FQ	final exam question (part of the mid-term exam in Study 4)
ICC	intraclass correlation coefficient
IT	information technology
MQ	mid-term exam question (part of the mid-term exam in Study 4)
NTD	neglected tropical disease
PBL	problem-based learning
RQ	research question (the abbreviation is used only when referring to specific research questions, such as "RQ 2.3")
SD	sustainable development
UNESCO	United Nations Educational, Scientific and Cultural Organization
WHO	World Health Organization
WSP	wicked sustainability problem

Table of contents

Acknowledgements

"If we create a sustainability course for the students in engineering nanoscience, will you help us develop and teach that course?" These words are one of the main reasons why I have written this thesis. In 2007, we were a group of three undergraduate students in engineering nanoscience who felt strongly about the lack of sustainability-related discussions in our courses. We worked hard to lobby for a mandatory course on sustainability in our educational program. We were actually surprised when we realized that we had succeeded, but we were even more surprised when Knut Deppert, who at the time was the Program Director for engineering nanoscience, asked us to participate in designing and teaching the course.

What followed was an exciting and challenging journey of trying to learn how to teach engineering students to discuss complex sustainability problems from multiple perspectives. None of the members in the course development group had any experiencing in teaching anything related to sustainability. Yet, together, we envisioned the kinds of learning we wanted to achieve and possible ways in which our teaching could support that learning. In 2010, the course was given for the first time. That same year, I graduated from Lund University. I realized that there was so much more to learn about how to prepare engineering students to work for sustainability, and I decided to pursue a Ph.D. in engineering education research. I was able to participate in further developing and teaching our course during the first years of my Ph.D. studies. I will always be grateful to Knut Deppert and Elisabeth Nilsson for trusting us students to participate in designing and teaching this course. I will also always be grateful to Knut Deppert, Elisabeth Nilsson, Andreas Ahrens, Greger Hammarin, Erik Mårsell, Daniel Jacobsson, Olof Persson, Frida Lindberg, Sofia Yngman, and Carina Fast for sharing the excitement and challenges of working with the course throughout the years.

Pursuing a Ph.D. in engineering education research was equally exciting and challenging, and I would not have made it this far without the help of my supervisors, colleagues, friends, and family. I consider myself very lucky to have had four very competent and supportive supervisors. Throughout my Ph.D. studies, they provided constructive feedback that helped me challenge and clarify my ideas. They also helped me narrow my research focus rather than opening up ever new research avenues. Maybe most importantly, they helped me deal with my uncertainty about the worth of my own work. Many times during my Ph.D. studies, I felt an urge to discard everything I had done so far and start over from scratch. At those times, my supervisors helped me understand that there is no such thing as "perfect research" and they encouraged me to keep going. Magdalena Svanström, Åke Ingerman, John Holmberg, and Tom Adawi, thank you so much for all your support!

I also want to thank Jens Kabo, who acted as an informal tutor throughout my Ph.D. project. Thank you for all our discussions about the joys and challenges of academic work, and thank you for patiently listening to all my ideas, even those that definitely went beyond what I could do within the scope of my Ph.D. project.

In the research underlying this thesis, I attempted to integrate two different research fields, engineering education research (EER) and environmental and sustainability education research (ESER). Trying to navigate between these fields sometimes left me feeling that I did not belong anywhere. I want to thank all research groups that welcomed me into their communities and made me feel that I belonged – despite differences in research foci and/or approaches. To all members of those communities (for example the NNEER, REEN, ASEE, ESER, and EESD communities; the RESD, Variation Theory, and Action Research groups at Gothenburg University; and the Department of Engineering Education at Virginia Tech): Thank you for all the interesting discussions that I was allowed to participate in, for constructive feedback on my work, and for making me feel that my participation in, and contributions to, the communities were valued. Thank you also to my colleagues in our own research group, EER at Chalmers University of Technology, for interesting seminars and discussions in a friendly and supportive atmosphere.

I want to direct a special thanks to the leaders and members of the CUL-group for mathematics and science education at Gothenburg University. I am immensely grateful for the opportunity to participate in so many constructive and interesting seminars and productive writing retreats. Even more importantly, I am grateful for the much needed emotional support and friendship that was so generously given and that helped me persevere through many challenging periods during my Ph.D. studies.

During most of my Ph.D. studies (until our research group moved to the Department of Communication and Learning in Science in May 2017), the Department of Applied IT provided a safe and friendly working environment and the resources I needed for my research. I am grateful to my former colleagues at that department for helping me unwind from the stress of my work by making me laugh during lunch and fika breaks, departmental conferences, julbord dinners, relay races, and many hours in the climbing gym.

Nothing allows me to forget about work as well as flying and exercising. To my paragliding, hang gliding, Ultimate Frisbee, and climbing friends: Thank you for helping me remember that life is about more than a Ph.D. project – and thank you for putting up with me when I was stressed or tired from work. Kajsa and Sebastian, I would not have made it through these years without the two of you. Mama, Papa, Hendrik, Hans-Martin, and the rest of my wonderful family – thank you for always being there for me and believing in me, but also for making me feel that you love me irrespective of how well I succeed in my work. Who would I be without you?

I owe special thanks to the research participants who volunteered their time and personal experiences to contribute to my research. Studying students' approaches to wicked problems and working together with educators in trying to support students in their attempts to address wicked problems has taught me a lot, not only about wicked problems in engineering education, but also about how to deal with the lack of control, unpredictability, and ambiguity that is a necessary part of my research.

I also want to thank all those who devote their time and energy to improving engineering education and to preparing engineering students to work for sustainability. I feel honored to be a part of an engaged, creative, and optimistic community in which so many people contribute unique ideas and initiatives.

Last but not least, I am indebted to the Board of Undergraduate and Master's Studies at Chalmers University of Technology for financing my Ph.D. project. I am also grateful for travel grants from Adlerbertska Forskningsstiftelsen, Chalmerska forskningsfonden, Kungliga Vetenskaps- och Vitterhets-Samhället i Göteborg, and Stiftelsen för internationalisering av högre utbildning och forskning; these grants allowed me to participate in international conferences and thus develop my research network in both EER and ESER.

Johanna Lönngren, Gothenburg, 26 July 2017

WICKED PROBLEMS IN ENGINEERING EDUCATION

Preparing Future Engineers to Work for Sustainability

INTRODUCTION

Problem statement

The need to include environmental and sustainability education (ESE) in higher education in general, and engineering education in particular, is today widely recognized both internationally and in Sweden. Already in 1977, the United Nations Educational, Scientific and Cultural Organization (UNESCO) held its First Inter-governmental Conference on Environmental Education in Tbilisi, Georgia. Based on this conference, the Tbilisi Declaration (UNESCO, 1978) was published, which included "recommendations for the wider application of environmental education in formal and non-formal education" and which has been highly influential in shaping the development of ESE around the world (Palmer, 1998, p. 8). Two other important documents are the Rio Declaration (UNCED, 1992c) and the Agenda 21 (UNCED, 1992a). These documents were developed during the United Nations Conference on Environment and Development in Rio de Janeiro in 1992. The Rio Declaration is a statement of 27 principles of sustainability. These principles provide a basis for international cooperation on sustainability issues. The Agenda 21 is an action program with detailed descriptions of how participating nations can address a wide range of sustainability challenges such as poverty, toxic waste, and education. Chapter 36 in the Agenda 21 specifically addresses the question of how participating nations can promote and develop ESE (UNCED, 1992b). More recently, the UNESCO focused attention on ESE through the UN Decade of Education for Sustainable Development 2005-2014 (UNESCO, 2005) and the Global Action Plan on Education for Sustainable Development (UNESCO, 2014).

Other organizations have focused on ESE specifically in engineering education. One important document is the Barcelona Declaration, which was formulated during the Engineering Education for Sustainable Development Conference in 2004 (EESD Conference Scientific Committee, 2004). Similarly, the American Association of Engineering Societies and the World Engineering Partnership for Sustainable Development have published a document about the role of engineering in sustainable development (AAES, 1994). In Great Britain, the Royal Academy of Engineering published influential guiding principles for ESE in engineering education (Royal Academy of Engineering, 2005), and in the United States, the Engineering Accreditation Commission has included ESE-related learning outcomes in their Criteria for Accrediting Engineering Programs (ABET, 2016).

Also in Sweden, several policy documents stress the importance of ESE. For example, Sweden has closely collaborated with neighboring countries in the Baltic region; this collaboration resulted in the Haga Declaration (Ministers of Education, 2000) and a local

Agenda 21 action plan for the Baltic region (Baltic 21, 2002). In addition, the Swedish Ministry of Education and Research has formulated national requirements for ESE at higher education institutions in general (Ministry of Education and Research, Sweden, 1992) and at engineering education institutions in particular (Ministry of Education and Research, Sweden, 1993).

Despite this national and international consensus, most engineering education today does not adequately prepare students to contribute to sustainability (Fenner, Cruickshank, & Ainger, 2014; Lundqvist & Svanström, 2008; The EESD Observatory, 2009; Thompson, 2002). For example, engineering students often do not learn how to address complex and ill-structured sustainability problems such as those mentioned in the Agenda 21 (Seager, Selinger, & Wiek, 2012). These kinds of problems are commonly called *wicked problems* (p. 19). Addressing wicked problems is particularly challenging since these problems typically involve different stakeholders with conflicting interests and values. Wicked problems are also characterized by irreducible uncertainty, and it is not possible to develop definite problem descriptions or solutions (Rittel & Webber, 1973).

Rather than learning to address wicked problems, engineering students typically learn to solve well-structured *story problems*, which do not contain any uncertainty; all parameters of the problems are specified in the problem statement. Story problems have knowable, correct solutions and there are established methods for how to arrive at these solutions:

> When learning to solve story problems in engineering, students learn to translate relationships about unknowns into equations, solve the equations to find the value of the unknowns, and check the values found to see if they satisfy the original problem. This linear process implies that solving problems is a procedure to be memorized, practiced, and habituated, a process that emphasizes getting answers over making meaning. (Jonassen, Strobel, & Beng Lee, 2006, p. 139)

An excerpt from one of the interviews that I conducted for this thesis illustrates the lack of attention to wicked problems in engineering education from a student perspective:

> JL: Do you usually encounter problems like these [the wicked problem of water shortage in Jordan] in your educational program?
>
> Student: Not really.
>
> JL: What is the difference?
>
> Student: The difference is that [, in school,] you're just supposed to learn something. Sure, sometimes you get problems to solve, like when you solve math problems. But that problem, you can often solve it pretty quickly with some kind of calculation and then it's solved, you're done. But that's not really how it is, it doesn't work that way [in real life]. Because [t]here you might solve a problem in one country, and then you come to the next [country]. Or you might solve a problem in one country, but that might not turn out so well, so you have to continue. And that is what—, that's how it's going to be in real life. It's not like I'm just going to calculate this formula and then we have everything, then you'll all get water. In our educational program, it's been like that [merely calculating solutions with the help of mathematical formulas], and I think it's been like that in [high school] as well. The problems you get to work with there have a solution. And that's the one [solution] that is correct, and there aren't a lot of issues to discuss. Because it [the solution] just is the way it is and there isn't really anything else to think about.

Unfortunately, learning to solve well-structured problems, such as story problems, does not readily transfer to wicked problems (Schraw, Dunkle, & Bendixen, 1995; Shin, Jonassen, &

McGee, 2003). At the same time, learning to address wicked problems is challenging for engineering students: “Often [engineering] students want to learn an easy trick by which they can make their designs not only effective and efficient but also sustainable. It is often disappointing [for the students] to recognize that there is no easy trick that leads to sustainable technology” (Mulder, Segalàs-Coral, & Ferrer-Balas, 2010, p. 3). In fact, learning to address wicked problems may be *more* challenging for engineering students than for students in other majors. The reason is that engineering students are more likely than students in other majors to view knowledge as certain (Paulsen & Wells, 1998), which in turn has been shown to negatively affect students’ ability to address ill-structured problems, such as wicked problems (King & Kitchener, 1994; Schraw, Dunkle, & Bendixen, 1995).

Relatively little research has been done related to engineering students’ ability to address wicked problems. In particular, more research is needed on what engineering students need to learn to be able to address wicked problems, what pedagogical methods could be used to support that learning, and how learning outcomes could be assessed (Mulder, Segalàs-Coral, & Ferrer-Balas, 2010). Such research could support engineering educators in the challenging process of including wicked problems in their teaching, both by providing practically applicable research results and by providing opportunities for teacher professional development in the context of research activities. This support for educators is particularly important since many educators themselves lack adequate training in addressing wicked problems and in teaching students how to address wicked problems (Mulder, Segalàs-Coral, & Ferrer-Balas, 2010).

Research aim and questions

The overarching aim of the research underlying this thesis was to *contribute to improving engineering education practice such that it could better prepare engineering students to contribute to sustainability*. More specifically, the research addressed the overall question of how engineering education can prepare students to address wicked problems.

The research consisted of two parts. The first part of the research focused primarily on *student learning*, and the second part focused primarily on *teaching and assessment* related to wicked problems in engineering education. The research questions in the second part emerged from the results of the first part.

In total, four empirical studies were conducted, two in each part of the research. Each of the four studies was designed to answer the specific questions that emerged during the research. These questions differed in character; to answer the different kinds of questions, different theoretical perspectives and methodological approaches were used (see p. 99 for a discussion on the use of multiple theoretical perspectives in the thesis). The answers to the different kinds of questions resulted in different kinds of knowledge claims and different kinds of contributions to the overarching aim of the research (e.g. mainly theoretical contributions in Study 2 versus mainly practical contributions in Study 3).

Part 1: What do engineering students need to learn to be able to address wicked problems?

The aim of Part 1 was to develop theoretical understanding of what it means to learn to address wicked problems in the context of engineering education, for example through the use of multiple perspectives. The following seven research questions (RQs) guided the inquiry in this part:

- RQ 1.1 What characteristics of wicked problems are particularly challenging for engineering students?
- RQ 1.2 What could be relevant examples of wicked problems for undergraduate engineering students and how could these problems be described for the students?
- RQ 1.3 What are relevant characteristics of perspectives in the context of learning to address wicked problems?
- RQ 1.4 What are some ways in which students interact with their perspectives?
- RQ 1.5 In what ways do engineering students approach a wicked problem?
- RQ 1.6 Which of these approaches may be considered appropriate in the normative context of ESE?
- RQ 1.7 What are salient aspects of integratively addressing wicked problems for engineering students?

RQs 1.1 and 1.2 were addressed in a pre-study literature review (p. 31). RQs 1.3-1.5 were addressed in two studies that both relied on the same empirical material from ten in-depth interviews with undergraduate engineering students. Study 1 addressed RQs 1.3 and 1.4 with the help of qualitative content analysis (Cohen, Manion, & Morrison, 2011) of the material (p. 45). In Study 2, a phenomenographic approach (Marton & Booth, 1997) was used to address RQ 1.5 (p. 51). To address RQs 1.6 and 1.7, the results from Studies 1 and 2 were synthesized (p. 55).

Part 2: How can the ability to address wicked problems be taught and assessed in engineering education?

Part 2 aimed to develop theoretical and practical understanding of how engineering education can prepare students to address wicked problems through specific teaching and assessment approaches. The results from the research in Part 1 served as a basis for formulating the following eight research questions for Part 2:

- RQ 2.1 What do engineering educators view as important barriers for teaching their students to integratively address wicked problems?
- RQ 2.2 What learning outcomes do educators view as important for students’ ability to integratively address wicked problems?
- RQ 2.3 How could (some of) these learning outcomes be assessed?
- RQ 2.4 What kinds and levels of reliability, validity, and utility can be achieved with an analytic rubric for assessing students’ written responses to wicked problems?
- RQ 2.5 How do students perform when different scaffolding strategies are used?
- RQ 2.6 How do students approach wicked problems when different scaffolding strategies are used?
- RQ 2.7 What affordances for learning do different scaffolding strategies provide?

RQs 2.1-2.3 were addressed in Study 3, in which a pragmatic action research approach (Johansson & Lindhult, 2008) was used (p. 61). RQs 2.4-2.7 were addressed in Study 4, in which a design-based research approach (McKenney & Reeves, 2012) was used in an undergraduate engineering education context (p. 75).

Thesis overview

This thesis consists of eight chapters. After this *Introduction* chapter, I provide an overview of the *Research context* by describing the historical developments and disciplinary cultures of my primary research field, engineering education research (EER), and my secondary research field, environmental and sustainability education research (ESER). I also explain how I integrated these two fields in my research. In the next chapter, I review the literature on *Wicked problems*.

The main body of the thesis consists of two chapters in which I describe the research in *Part 1* and *Part 2* respectively. To facilitate an understanding of how the research questions in the second part emerged from the results in the first part, the structure of the two chapters mirrors the emergent design of the research: I address one or two research questions at a time rather than describing all methods and methodologies in one chapter and all results and conclusions in another chapter.

Figure 1 can be used as a map to navigate through the two chapters on Parts 1 and 2. To facilitate such a use of the figure, a duplicate of the figure is printed on a foldout page at the very end of the thesis. Opening this foldout page makes it possible to see the figure while reading the thesis. The figure provides an overview of the two parts of the research (gray boxes); studies and papers in those parts; and research questions (blue boxes). The figure also contains results and conclusions for each of the research questions (yellow boxes).

After the description of the research in Parts 1 and 2 follows a *Discussion* chapter. In that chapter, I first discuss the use of multiple theoretical perspectives within the context of the research underlying this thesis. Second, I discuss limitations in the individual studies. Finally, I describe contributions, implications, and questions for future research. I end the thesis with a *Conclusion*, in which I collate the most important findings, conclusions, and contributions of the research presented throughout the thesis.

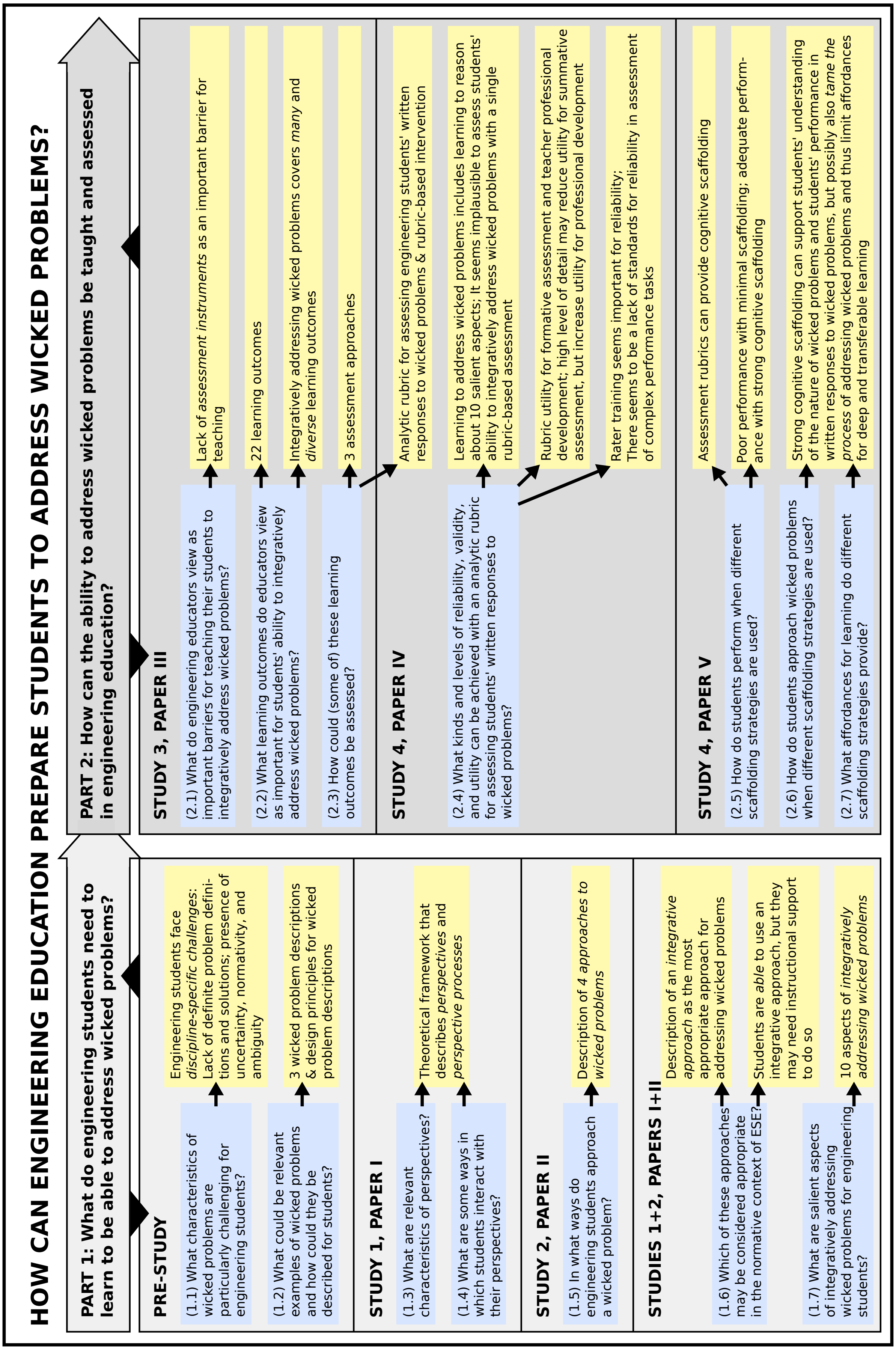

Figure 1. Overview of the two parts of the research (gray boxes); studies and papers; research questions (blue boxes); and results and conclusions (yellow boxes).

RESEARCH CONTEXT

The research underlying this thesis was financed by the Board of Undergraduate and Master's Studies at Chalmers University of Technology (Chalmers). The board's decision to finance a Ph.D. position in EER was based on the expectation that the research conducted during the Ph.D. project would contribute to informing and improving undergraduate and/or Master's education at Chalmers, with a particular focus on how education at the university could prepare students to contribute to sustainability. The board's expectations were closely aligned with my own ambitions to contribute to educational practice and to sustainability.

At Chalmers, I have been a part of the EER division. Thus, the primary research context for this thesis is the field of EER. However, I also draw on literature from the field of ESER, which serves as a secondary research context. Throughout my Ph.D. project, I participated in scholarly discussions in both EER and ESER to discuss and further develop my research. Engaging with both fields allowed me to include diverse theoretical perspectives and methodological approaches and develop critical awareness about the importance of theoretical perspectives in my research.

Primary research context: EER

Engineering education as a *research* field has recently emerged out of the needs of educators to understand and adapt to the changing demands on engineering education *practice* (Borrego & Bernhard, 2011; Seely, 2005). Therefore, an important purpose of EER is to address challenges that arise in engineering education practice (Bernhard & Baillie, 2016).

EER as a way to address challenges in engineering education practice

Engineering education practice has a long history around the world. Throughout this history, engineering education has undergone profound changes in response to socio-cultural, economical, environmental and technological developments. Originally, the purpose of engineering education was to serve the needs of local small-scale industries, and it was mainly carried out in the form of individual "hands-on apprenticeship" (Seely 2005, p. 115) between a master and an apprentice. With the beginning of the industrial revolution, the nature of technological innovation changed dramatically. In particular, the pace of innovation increased, which placed higher demands on engineers' scientific knowledge. Engineering education adjusted to these demands with a stronger focus on the natural sciences at the expense of learning and application in specific engineering contexts (Lohman 2008; Wankat, Felder, Smith, & Oreovicz 2002). Other driving factors for the development of engineering education towards more scientific approaches were military developments after World War II, and the space programs of the 1950s and 60s (Wankat, Felder, Smith, & Oreovicz 2002). More recently, the demands on engineering education have again changed as the social and environmental contexts of engineering are recognized as important (Jamison, Hyldgaard Christensen, & Botin, 2011; Wisnioski, 2012). Engineering students are expected to develop generic skills such as design skills and an ability to handle global complexity, diversity, sustainability issues, and stakeholder interests (Lohman 2008; Wankat, Felder, Smith, & Oreovicz 2002). While demands for understanding the social context of engineering have increased, demands for developing a strong scientific knowledge base have remained high. Thus, the demands on engineering students and educators have increased and at the same time become more diverse.

In this context of changing and increasing demands on engineering education, educators grapple with questions about what is considered high-quality engineering education and how it could be achieved in practice. The field of EER provides one avenue for educators to address these questions.

Dominance of post-positivism in EER

Engineering involves the use of specialized mathematical and scientific knowledge and technical skills to physically construct "highly technical and mechanized products" (Pawley, 2009, p. 310). Thus, engineering as a field has much in common with mathematics and the physical sciences (Biglan, 1973).

Drawing on Kuhn (1962/1970), Biglan argued that engineering, mathematics, and the physical sciences are "characterized by the existence of paradigms that specify the appropriate problems for study and the appropriate methods to be used" (1973, p. 195). According to Guba and Lincoln, "a paradigm may be viewed as a set of *basic beliefs*. These beliefs represent assumptions about the nature of the 'world', the individual's place in it, and the range of possible relationships to that world and its parts" (1994, p. 107, italics in original; see also Creswell, 2003):

- *Ontological* assumptions concern the form and nature of the world.
- *Epistemological* assumptions concern the nature of knowledge about the world, i.e. the relationship between that which can be known about the world and those who hold that knowledge.
- *Methodological* assumptions concern the nature of inquiry, i.e. they provide strategies or plans of action that influence what methods researchers choose to use in their inquiry.
- *Axiological* assumptions are personal or professional values that guide researchers' understanding of what may be important questions to focus on in their research and how they should conduct their research.

- Assumptions about appropriate *methods* concern beliefs about what specific methods may be appropriate in the context of a specific set of ontological, epistemological, methodological, and axiological assumptions.

For thousands of years, philosophers have discussed the validity of different sets of assumptions without reaching consensus. However, a relatively strong consensus exists within some academic fields. Biglan (1973) called such fields "hard" or "high-paradigmatic" fields, and contrasted them with "soft", "low-paradigmatic" fields. He suggested that engineering represents a high-paradigmatic field, where *post-positivism* is widely accepted and used as the dominant research paradigm (see also Gardner & Willey, 2016; Koro-Ljungberg & Douglas, 2008; Paulsen & Wells, 1998).

Researchers who work in a post-positivist tradition typically strive to develop generalizable, scientific explanations that enable control of processes and prediction of the outcomes of those processes. Post-positivist researchers typically acknowledge that objective truth is ultimately unattainable. However, they strive to *approach* truth by formulating, and attempting to falsify, hypotheses that could challenge theories that previously have been tentatively accepted as "probably true" (Guba & Lincoln, 1994, p. 110). Inherent in this focus on hypothesis testing is an assumption that there is an external "reality" that is independent of whether (and how) human or non-human beings experience and interact with the world. However, post-positivism also assumes that this external reality is not directly accessible to objective experience and description. Rather, descriptions of reality are viewed as *distorted* by subjective experiences and personal values; values are seen as "confounding variables" (Guba & Lincoln, 1994, p. 114) that researchers attempt to exclude from their research as they strive for high levels of objectivity. In line with this ambition to exclude values from research processes and results, post-positivist research is often reported in *passive* voice.

Because an important purpose of EER is to address challenges that arise in engineering education practice (Bernhard & Baillie, 2016), most researchers in engineering education have a background in engineering (Osorio, 2005) and are most familiar with research approaches and assumptions that are commonly accepted in *engineering research* (Koro-Ljungberg & Douglas, 2008). Therefore, similar to engineering research, EER can be described as a high-paradigmatic field in which post-positivism is the dominant research paradigm (Borrego, Douglas, & Amelink, 2009; Koro-Ljungberg & Douglas, 2008; Jawitz & Case, 2009).

Alternative paradigms in educational research

In contrast to EER, the broader field of educational research is generally described as "soft" or "low-paradigmatic" (Biglan, 1973; Gardner & Willey, 2016), i.e. as characterized by a higher level of diversity in research approaches. Jawitz and Case (2009) argued that the strong reliance on a single research paradigm in EER is problematic since it may restrict the scope of questions that researchers can ask and the kinds of results they can obtain. It is reasonable to assume that reliance on a single research paradigm in EER creates large areas of *blind spots* in the field – important questions that simply cannot be asked and addressed:

> All scientists operate in a world defined by what they think and know to be true. What they don't know well enough to even ask about or care about are their *blind spots*. What they know enough to question but not answer are their *blank spots.* (Wagner, 1993, p. 16, emphasis added)

More specifically, Douglas *et al*. (2012) argued that *epistemological pluralism* is necessary to allow investigation of complex, multi-layered phenomena in natural and social contexts (such as wicked problems). To broaden the scope of EER (and thus, in Wagner's language, to identify blind spots in the field), Jawitz and Case (2009) suggested that researchers in EER should consider two alternative paradigms that are commonly referred to in educational research, *interpretivism* and *critical inquiry*.

Table 1. Assumptions in three research paradigms that are commonly referred to in educational research (based on descriptions by Creswell, 2003, and Guba and Lincoln, 1994)

	Post-positivism	**Interpretivism**	**Critical inquiry**
What is the form and nature of reality? (Ontological assumptions)	There is an external world that can be described and understood	There is an external world, but we do not have objective access to it. We can only describe and understand different ways in which people experience the external world	There is an external world, and the structures of that world restrict humans' actions and experiences. We can understand and change these structures
What is the form and nature of knowledge? (Epistemological assumptions)	Objective and generalizable knowledge about the external world	Subjective and contextualized knowledge about human experience of the external world	Objective and contextualized knowledge that contributes to disrupting unequal power relations in the external world
How do researchers judge knowledge? (Epistemological assumptions)	In terms of how well the results correspond to what is true in the external world	In terms of how well the results describe and explain the specific situation, striving for consensus	In terms of how well the results contribute to disrupting unequal power relations, striving to disrupt constraining consensus
How do researchers try to gain knowledge? (Methodological assumptions)	Deductively testing hypotheses through experiments and measurement	Inductively interpreting what is expressed in specific contexts	Emancipation, participation in democratic dialog, and critical reflection
How are values understood in relation to research? (Axiological assumptions)	Values are seen as "confounding variables" that should be excluded as much as possible	Values are seen as inextricably linked to research; Explicit attention to values can help researchers to investigate non-dominant experiences	Values are seen as important drivers for research; Research is explicitly guided by values of democracy and social justice
What methods are considered appropriate? (Assumptions about methods)	Most often quantitative	Most often qualitative and interpretative	Most often qualitative and participative

While post-positivism aims to objectively describe an external "reality", *interpretivism* rests on the assumption that all knowledge is based on subjective experiences and that it, therefore, is impossible to gain objective knowledge about the world. In line with this assumption, interpretivism generally aims to describe and understand different ways in which people experience the external world. Interpretative research typically uses qualitative and inductive approaches to research (Guba & Lincoln, 1994). To stress the importance of personal perspectives and values, interpretative research is commonly reported in *active* voice.

Critical inquiry is explicitly guided by social values such as democracy and social justice. It rests on the assumption that there is an external "reality" in which unequal power relations constrain humans' experiences of, and actions in, the world. Thus, critical inquiry aims to disrupt (rather than merely describe and understand) unequal power relations. For this purpose, critical inquiry typically uses qualitative methodologies, most often emancipatory and participatory approaches (Guba & Lincoln, 1994). Table 1 provides a comparison of basic assumptions in post-positivism, interpretivism, and critical inquiry, based on descriptions by Creswell (2003) and Guba and Lincoln (1994).

Several EER authors have suggested that EER lacks *methodological diversity*; the dominant approach in engineering education is *quantitative* research, based on, for example, statistical comparison of treatment and control groups in educational interventions. For example, Koro-Ljungberg and Douglas (2008) found that the majority of articles published in the Journal of Engineering Education (one of the leading journals in the field) between 2005 and 2006 were quantitative research papers. Malmi *et al*. (2016) found that the proportion of qualitative research in the field has increased during the past decade, but that research published in the European Journal of Engineering Education (another leading journal in the field) in 2008, 2009 and 2013 still mostly employed "simple" quantitative approaches such as descriptive statistics, correlations or comparison of means. Malmi *et al*. argued that most of the qualitative research published in those years is research in which "qualitative data have been analyzed but no clear method and/or analysis process is reported" (2016, p. 5). Similarly, Koro-Ljungberg and Douglas (2008) found that some of the (few) qualitative research papers in their study did not fully embrace qualitative research perspectives. For example, authors sometimes applied quantitative analytic methods to qualitative data. Case and Light described a number of qualitative methodological approaches as "emerging" in EER. They argued that these approaches are not yet widely used in EER, but that they could be used to broaden the scope of researchable questions and develop an "[EER] field that can extend its domain of questions to those that are patently needed to be asked" (2011, p. 190).

While there appears to be a broad consensus in the field that paradigmatic, epistemological, and methodological pluralism is valuable (Borrego, Douglas, & Amelink, 2009; Case & Light, 2011; Koro-Ljungberg & Douglas, 2008; Moskal & Brown, 2014), there is still evidence that this pluralism has not yet been achieved in all parts of the engineering education research community. For example, The Steering Committee of the National EER Colloquies published a Special Report in which the authors outlined "the research agenda" for EER. The text is currently used as a guide for scholars who want to publish research articles in the Journal of Engineering Education. In that Special Report, the authors suggested that research in EER should be performed "similar to the way in which research is performed and used in the traditional engineering disciplines" (SCNEER, 2006, p. 259) – which indirectly (and possibly unintentionally) seems to suggest that research in EER should be based on post-positivist research perspectives and quantitative research methods. Another example is found in a guest editorial in the Journal of Engineering Education, in which Fortenberry (2006) compared EER to large-scale, biomedical research rather than educational research. In line with the norms of natural scientific research, Fortenberry called for coordination of EER under "an overarching research framework". Such a framework, Fortenberry argued, would be based on "clearly defined deliverables and specific end-points" and lead to an increased "return on investment" in EER (*ibid*., p. 4). In fact, readers of the Journal of Engineering Education raised concerns that publication of qualitative research is not valued or encouraged in the journal (Moskal & Brown, 2014). Similarly, Borrego, Douglas, and Amelink (2009) found that engineering education practitioners lament a lack of qualitative studies in EER and a bias towards quantitative approaches in the peer review process. However, they also found that the same practitioners still enacted a quantitative, post-positivist bias when they were the ones critiquing other researchers' work. To address this lack of paradigmatic, epistemological, and methodological pluralism in EER, I used the field of ESER to broaden my understanding of what can be considered adequate research questions, theoretical perspectives, and methodological approaches.

Secondary research context: ESER

ESER as a way to address sustainability challenges

ESER emerged in the context of rising concerns about the threat of environmental degradation in the 1960s, combined with a growing awareness that environmental problems are not purely scientific problems. Environmental problems were viewed as arising out of interactions between cultural and natural processes and therefore, ESE was considered "a necessary component of any solution to the environmental crisis" (Gough, 2013, p. 14).

Stevenson, Wals, Dillon, and Brody (2013, p. 2) identified five characteristics of ESER that may distinguish it from other kinds of educational research:

- Since "environmental issues are fundamentally normative or value-laden by nature", ESER "embraces normative questions", i.e. questions related to value judgments rather than facts.
- ESER is an interdisciplinary field that is rooted in the "interdisciplinary nature of people-society-environment relationships".
- Rather than predominantly focusing on developing students' knowledge, skills, attitudes, and values, ESER also focuses on developing students' agency and action competence to enable them to contribute to sustainability and "enact solutions".
- Rather than predominantly focusing on formal educational contexts, ESER acknowledges the importance of non-formal and informal contexts. And
- since "the scale of environmental issues ranges from the local to the global", ESER includes both local and global orientations.

Paradigmatic and methodological diversity in ESER

Similar to EER, ESER emerged in the context of a high-paradigmatic field with a dominant post-positivist paradigm. ESER was originally closely associated with the environmental sciences, in which a post-positivist paradigm was the norm. ESER continued to be dominated by a post-positivist paradigm through the 1980s. However, during the 1990s, researchers in the field increasingly adopted interpretative and critical paradigms (Gough, 2013; Hart & Nolan, 1999; Palmer, 1998). Recently, the field has been described as embracing "widely differing discourses" (Sauvé, 2005, p. 11) and an "extraordinary diversity of perspectives" (Ardoin, Clark, & Kelsey, 2013, p. 514).

As ESER transitioned from high- to low-paradigmaticity, methodological diversity also increased. In the late 1970s to 1980s, over 90% of the research in the field employed quantitative methodologies (Hart & Nolan, 1999). During the 1990s, however, the share of qualitative approaches such as case study methodology, action research, and narrative inquiry has increased considerably (Gough, 2013; Hart, 2006; Hart & Nolan, 1999; Palmer, 1998). Recently, Gough argued that today, "a wide range of research methodologies [are] being engaged in environmental education research" (2013, p. 18). Expressions such as "thousand flowers bloom" (Gough, 2013, p. 18) are used to describe the methodological diversity in the field. However, at the same time as methodological diversity increased, an antagonistic debate emerged in which researchers from different philosophical schools attempted to "establish which methodological research approach is the 'most desirable for environmental education'" rather than acknowledging the value of all methodological approaches (Connell, 1997, p. 118, citing Robottom & Hart, 1995, p. 5). Possibly as a result of this debate, quantitative approaches may today be underrepresented in ESER (Reid & Scott, 2006).

Despite the broadening of the field since the 1990s, some researchers have suggested that discussions about philosophical perspectives (Sund & Greve Lysgaard, 2013), acceptance of alternative paradigms (McKenzie, 2016), and explicit discussions of methodological questions (Dillon & Wals, 2006) remain insufficient. However, already in the 1990s, Connell (1997) argued:

> There has been extensive dialogue about the 'politics of method' in environmental education research during the past decade (…) [, which] is an indication of the growing acceptance of newly evolved and alternative methodological approaches to educational research. (…) This dialogue is evidence of healthy on-going debate; it promotes rigour and flexibility and enables the environmental education movement to change when new information and insights are able to challenge the current state of affairs. (Connell, 1997, p. 118; see also Gough, 2013)

Similarly, prominent journals in ESER invite researchers to submit manuscripts about research that is based on a variety of methodological approaches and schools of thought. For example, the editor of Environmental Education Research (one of the leading journals in the field) explicitly invited researchers to submit manuscripts about studies in which "novel ways of thinking about emerging issues" (Reid, 2014) have been used. This should be contrasted with the calls for post-positivist research perspectives and quantitative research methods in the Journal of Engineering Education (Fortenberry, 2006; SCNEER, 2006) as discussed above.

Integration of primary and secondary research contexts

EER and ESER share a background in high-paradigmatic, predominantly post-positivist fields in which quantitative methodologies used to be the norm. For both fields, it has been, and continues to be, important to strive for paradigmatic and methodological diversity. However, the fields have dealt with this challenge in different ways, which has led to the development of different research foci in terms of commonly used and accepted concepts, theories, and methodologies. In Wagner's terms, the fields are therefore characterized by different "configuration[s] of their blind spots", which are "areas in which existing theories, methods, and perceptions actually keep us from seeing phenomena as clearly as we might" (Wagner, 1993, p. 16). In the research underlying this thesis, I used concepts, theories, and methodologies from EER to illuminate blind spots in ESER, and I used concepts, theories, and methodologies from ESER to illuminate blind spots in EER.

ESER has a long history (Palmer, 1998) and can be considered a more mature field than EER in the sense that a higher degree of paradigmatic and methodological diversity has been achieved. Thus, ESER can provide concrete examples of how diverse paradigms and methodological approaches can be used in an originally post-positivist field and in educational research in general. At the same time, ESER may have marginalized post-positivist perspectives and quantitative methodologies. EER can provide an example of a more nuanced approach to its post-positivist origin: While researchers in EER consistently advocate methodological diversification through the use of alternative methodologies, they typically avoid suggesting that post-positivist research does not have any value for EER. Thus, EER can also provide concrete and recent examples of how quantitative approaches can be used in educational research. In the research underlying this thesis, I used both EER and ESER literature to inform my choices of theoretical perspectives and methodological approaches. For example, my decision to use an action research approach in Study 3 was mostly informed by the ESER literature, while my decision to include a quantitative component in Study 4 was mostly informed by the EER literature. I also used the "extensive dialogue about the 'politics of method'" (Connell, 1997, p. 117) in ESER to develop critical awareness about the importance of theoretical perspectives in my research; this awareness encouraged me to write this thesis predominantly in an *active* voice, even though many researchers in EER may be more familiar and comfortable with research reports written in *passive* voice.

EER and ESER also share an ambition to contribute to addressing practical challenges. However, research in the two fields has focused on addressing different kinds of challenges (challenges in engineering education practice versus sustainability challenges, as described above). These different foci can complement each other in research that combines insights from both fields. For example, in EER, there is a lack of focus on concepts such as "ethics", "stakeholders", and "social justice" (as illustrated for example by the low number of entries related to these concepts in a recently developed taxonomy of the field; University of Michigan, 2016). ESER, on the other hand, has a strong focus on "normative questions" related to "people-society-environment relationships" (Stevenson, Wals, Dillon, & Brody, 2013, p. 2). Thus, I could use ESER to develop a general normative base for my research, i.e. my ambition to ultimately contribute to sustainability. Similarly, it has been argued that much of the work in ESER *predominantly* focuses on environmental and sustainability concerns and not sufficiently on *educational* concerns (Sund & Greve Lysgaard, 2013). The strong connection between EER and engineering education practice, and the strong focus on educational concerns in EER, helped me to stay focused on educational questions that matter for engineering education practice. Table 2 provides an overview of characteristics of both fields.

Table 2. Overview of characteristics of EER and ESER

	Characteristics of EER	**Characteristics of ESER**
Maturity of research field	Recently described as a "new discipline" (SCNEER, 2006), a "maturing (…) research field" (Jesiek, Newswander, & Borrego, 2009), and an "emerging field of inquiry" (Borrego, Streveler, Miller, & Smith, 2008)	Described as a "young and evolving" field during the late 1970s to 1980s (Palmer, 1998), a "rapidly expanding" field in the late 1990s (Hart & Nolan, 1999), and recently as a "vibrant" field (Ardoin, Clark, & Kelsey, 2013)
Reason for emergence of the field	Changes in engineering education that created new educational challenges (Seely, 2005)	Concerns about the environment (Palmer, 1998)
Paradigmaticity	Origin in high-paradigmatic field (engineering); today still relatively high-paradigmatic, but in transition to lower paradigmaticity (as illustrated by calls for paradigmatic diversity, e.g. Jawitz & Case, 2009)	Origin in high-paradigmatic field (environmental sciences); during the 1990s in transition to lower paradigmaticity; today low-paradigmatic with "widely differing discourses" (Sauvé, 2005) and "extraordinary diversity of perspectives" (Ardoin, Clark, & Kelsey, 2013), but also continued critique of limited acceptance of non-dominant paradigms (McKenzie, 2016)
Dominant research paradigm	Post-positivist both historically and today (Borrego, Douglas, & Amelink, 2009; Koro-Ljungberg & Douglas, 2008; Jawitz & Case, 2009); lack of discussion about theoretical perspectives today (Jawitz & Case, 2009)	Post-positivist during the late 1970s to 1980s; during the 1990s, broadening to include interpretative and critical perspectives (Palmer, 1998); today, post-positivist perspectives may be underrepresented (Connell, 1997; Reid & Scott, 2006); lack of discussions about philosophical perspectives both in the 1990s and today (Sund & Greve Lysgaard, 2013)
Epistemology	Lack of discussions about epistemology today (Douglas *et.al.*, 2012)	Lack of discussions about epistemology during the late 1990s (Hart & Nolan, 1999; Palmer, 1998); "shift from an objectivist view of knowledge to views that regard knowledge as more provisional, problematic, and socially constructed" during the 1990s (Hart & Nolan, 1999)

Table 2 continued. Overview of characteristics of EER and ESER

	Characteristics of EER	**Characteristics of ESER**
Methodology	Lack of diversity today: dominance of quantitative approaches (Borrego, Douglas, & Amelink, 2009; Case & Light, 2011; Koro-Ljungberg & Douglas, 2008; Moskal & Brown, 2014); during the past decade, the share of qualitative research increased (Malmi, *et.al*., 2016); recent call for using case study methodology, action research, and narrative inquiry (Case & Light, 2011)	Lack of diversity in the late 1970s to 1980s (over 90% quantitative); during the 1990s, the share of qualitative approaches increased (Hart & Nolan, 1999; Hart, 2006; Palmer, 1998); call for using case study methodology, action research, and narrative inquiry during the late 1980s to early 1990s (Palmer, 1998); today high diversity ("thousand flowers bloom", Gough, 2013), but lack of explicit discussions of methodological questions (Dillon & Wals, 2006); recently calls for more broadening, e.g. by increasing the share of quantitative research (Reid & Scott, 2006); lack of diversity in research on some topics, but not others (Rickinson, 2001)
Aims and scope of journals	Journal of Engineering Education: EER should be similar to engineering (SCNEER, 2006) or large-scale biomedical research (Fortenberry, 2006), i.e. post-positivist and quantitative; reviewers may be biased towards post-positivist and quantitative perspectives (Borrego, Douglas, & Amelink, 2009; Moskal & Brown, 2014)	Environmental Education Research: explicit call for "novel ways of thinking about emerging issues", and diverse methodologies and schools of thought (Reid, 2014)
Relationship to practice	Ambition to improve engineering education: increase diversity, improve public image of engineering, prepare students to solve complex problems, meet industry's employee needs, better understand and improve learning (Borrego & Bernhard, 2011)	Ambition to improve sustainability; lack of focus on educational concerns (Sund & Greve Lysgaard, 2013)

WICKED PROBLEMS

Overview of the literature on wicked problems

In 1973, design theorists Horst Rittel and Melvyn Webber published a paper in which they argued that problems in social policy are fundamentally different from problems in the natural sciences. To describe this difference, they introduced the terms “wicked problems” for problems in social policy and “tame problems” for problems in the natural sciences. Rittel and Webber (1973, pp. 161–167) suggested that wicked problems have the following ten characteristics:

1. “There is no definitive formulation of a wicked problem.” It is not possible to definitely describe a wicked problem before already knowing what the solution to the problem will be.
2. “Wicked problems have no stopping rule.” There is no point in time at which work with a wicked problem is completed since it is always possible to further improve a solution to a wicked problem. Further, because of complex systems interactions, a solution to a wicked problem will have consequences that will reach far into the future and into distant parts of the system. To ultimately and definitely evaluate a solution to a wicked problem, one would need to wait until all these consequences have occurred – which is impossible.
3. “Solutions to wicked problems are not true-or-false, but good-or-bad.”
4. “There is no immediate and no ultimate test of a solution to a wicked problem.” This is related to characteristic 2. Since there is no point in time at which work with a wicked problem is completed, it is also impossible to evaluate the full consequences of a wicked problem. Further, there are no generally accepted criteria according to which a solution to a wicked problem should be evaluated.
5. “Every solution to a wicked problem is a ‘one-shot operation’; because there is no opportunity to learn by trial-and-error, every attempt counts significantly.” Since wicked problems are located in the real world, every attempt to address a wicked problem will also have consequences in that world that need to be defendable. Further, every attempt to address a wicked problem will change both the problem and the context in which it occurs. Thus, the problem essentially becomes a different one after

each intervention – which in turn may require different approaches for addressing the problem.

6. "Wicked problems do not have an enumerable (or an exhaustively describable) set of potential solutions, nor is there a well-described set of permissible operations that may be incorporated into the plan" for addressing a wicked problem.
7. "Every wicked problem is essentially unique. (...) There are no classes of wicked problems in the sense that principles of solution can be developed to fit all members of a class."
8. "Every wicked problem can be considered to be a symptom of another problem."
9. "The analyst's 'world view' is the strongest determining factor in explaining a discrepancy" between his/her descriptions of a wicked problem and other actors' descriptions. Descriptions of wicked problems, and preferences for how they should be addressed, are strongly influenced by questions related to norms and values.
10. "The planner has no right to be wrong. (...) Planners are liable for the consequences of the actions they generate".

Since the publication of Rittel and Webber's seminal paper on wicked problems, the concept has been applied to problems in a wide range of disciplines, the original description of the concept has been discussed and reformulated, and strategies for addressing wicked problems have been suggested. Other scholars have focused on identifying gaps in the wicked problems literature or critiquing the theory as a whole. In this chapter, I will review this literature and present the rationale for using the concept in this thesis.

Research focusing on applying the concept

The wicked problems concept has been applied "in a wide range of real-world problem-solving fields" (McCall & Burge, 2016, p. 201), including fields related to engineering, such as software engineering (DeGrade & Stahl, 1990) and systems engineering (Kovacic & Sousa-Poza, 2013), and fields related to environmental and sustainability sciences, such as environmental policy (Balint *et al*., 2011), and environmental management (DeFries & Nagendra, 2017). The environmental and sustainability literature provides ample examples of the application of the concept to sustainability problems such as health inequalities (Blackman, *et.al*., 2006), water resource management (Hearnshaw, Tompkins, & Cullen, 2011), fisheries and coastal governance (Jentoft & Chuenpagdee, 2009), disappearance of coral reefs (Hughes, Huang, & Young, 2012), global climate change (Levin, Cashore, Bernstein, & Auld, 2009), the effects of global warming on wildfire incidence (Chapin Ill, *et.al*., 2008), fire risk management (Carroll, Blatner, Cohn, & Morgan, 2007; DeFries & Nagendra, 2017), child abuse (Devaney & Spratt, 2009), changing environmental behavior (Krasny, 2013), logistic challenges of preparing for and responding to a disaster (Tatham & Houghton, 2011), input subsidy programs (Ricker-Gilbert, Jayne, & Shively, 2013), state fragility (Menkhaus, 2010), and biodiversity conservation (DeFries & Nagendra, 2017).

Research focusing on developing the theory

Besides describing specific problems as "wicked", scholars have also attempted to strengthen Rittel and Webber's theory by reformulating or reinterpreting the original description of the concept. One such approach aims to *condense* Rittel and Webber's characteristics of wicked problems to "eliminated repetition in the original" while "remain[ing] true to Rittel and

Webber's conceptual construct" (Duckett, Feliciano, Martin-Ortega, & Munoz-Rojas, 2016, p. 45). Examples of such approaches are found in Conklin (2005); Duckett, Feliciano, Martin-Ortega, and Munoz-Rojas (2016); Farrell and Hooker (2013); Norton (2012); and Seager, Selinger, and Wiek (2012).

In another approach, scholars have attempted to *identify the essence* in Rittel and Webber's description of wicked problems. For example, Norton suggested that "we can see all ten of the characteristics as symptoms of underlying value conflicts" (2012, p. 458), and McCall and Burge suggested that "difficulties in identifying cause-effect relationships are the central theme of wicked problems" (2016, p. 207).

A third approach aims to critique individual criteria and *suggest alternative wordings*. McCall and Burge (2016) took such an approach. They argued, for example, that the original descriptions should be reworded such that trial and error is recognized as a necessity (rather than an impossibility) in addressing wicked problems. They suggested that the presence of unforeseen consequences in wicked problems makes it impossible to implement a design in any other way than a "trial". They concluded that a commitment to *redesign* after implementation is central in addressing wicked problems. McCall and Burge also argued that planners cannot be held accountable for unforeseeable consequences (rather than stating that "the planner has no right to be wrong") since "the one thing more dangerous than designing is failing to design. While actions can have undesirable consequences, so can failing to act" (2016, p. 207).

Finally, a fourth approach aims to *connect wicked problems theory to other theories*, for example when Farrell and Hooker (2013) argued that Rittel and Webber's description of wicked problems is compatible with contemporary descriptions of science problems, or when Coyne (2004) argued that viewing design problems as wicked should be understood as the first step in applying postmodern philosophy to design. However, McCall and Burge argued that such approaches weaken rather than strengthen the original theory since they represent "attempts at hostile takeovers of wicked problems theory by rival theories" and since "none of these attempts is compatible with the statements of Rittel and Webber" (2016, p. 201).

Research focusing on suggesting strategies for addressing wicked problems

Another strand of wicked problems literature focuses on *suggesting strategies for addressing wicked problems*. Duckett, Feliciano, Martin-Ortega, and Munoz-Rojas (2016) described different kinds of strategies that have been suggested in the literature. These kinds of strategies include[3]

- theoretical approaches, such as suggesting that addressing wicked problems requires a post-normal approach to science (Batie, 2008);
- analytic approaches that can be used to cope with uncertainty, such as modeling tools (Batie, 2008) and multi-criteria analysis (Hearnshaw, Tompkins, & Cullen, 2011);

[3] Most references in the following list are taken directly from Duckett, Feliciano, Martin-Ortega, and Munoz-Rojas (2016) description, others provide examples from the literature that have not been mentioned by Duckett, Feliciano, Martin-Ortega, and Munoz-Rojas.

- atomistic approaches, such as breaking problems down into smaller parts (Shindler & Cramer, 1999) and locking down problem formulations (Conklin, 2010);
- interdisciplinary approaches (Batie, 2008) and transdisciplinary/transacademic approaches (Brown, Harris, & Russell, 2010; Brundiers & Wiek, 2011; Conklin, 2010; Krasny, 2013; Palmer, 2012);
- systems thinking approaches, such as graph analysis and mind mapping (Berkes, 2011; Hearnshaw, Tompkins, & Cullen, 2011; Lönngren & Svanström, 2016);
- envisioning, such as scenario building (Batie, 2008; Whyte & Thompson, 2012);
- participatory/deliberative approaches, such as Dialogue mapping (Conklin, 2005), Issue-Based Information Systems (Kunz & Rittel, 1970; Noble & Rittel, 1988), and public participation (Coyne, 2004; Sharman, 2009); and
- iterative approaches, such as redesign (McCall & Burge, 2016) and adaptive management (DeFries & Nagendra, 2017).

However, Duckett, Feliciano, Martin-Ortega, and Munoz-Rojas (2016) found that many of the suggested strategies did not address all characteristics of wicked problems; rather, the strategies focus on different *subsets* of the characteristics of wicked problems.

Scholars have also focused on *explicitly describing taming strategies*, i.e. strategies that are inherently inappropriate for dealing with wicked problems:

> Attempting to tame a wicked problem, while appealing in the short run, fails in the long run. The wicked problem simply reasserts itself, perhaps in a different guise, as if nothing had been done. Or, worse, sometimes the tame solution exacerbates the problem. (Conklin, 2005, p. 12)

Conklin (2005, p. 11) described six ways of taming wicked problems:

1. locking down the problem definition rather than allowing iterative redefinition,
2. asserting that the problem is solved when it clearly is not from all perspectives,
3. specifying objective parameters by which to measure the solution's success,
4. casting the problem as 'just like' a previous problem that has already been solved,
5. giving up on trying to get a good solution to the problem, and
6. declaring that there are just a few possible solutions, and that one just needs to choose to "best" one among those.

De Fries and Nagendra described taming strategies as *traps* that "can curtail incremental, partial improvements" in addressing wicked problems: Trap A involves "oversimplifying a problem and assuming that a technical solution will fix a wicked problem" (2017, p. 269). This description is similar to Conklin's first, second, third, fourth, and sixth ways of taming wicked problems. Trap B involves "making a problem overly complex" such that "managers trained in technical problem-solving can be ill-equipped to confront" the problem. "The result can be inaction from the inability to identify an incremental, partial solution" (2017, p. 269). In other words, there is a risk of inaction, in line with Conklin's fifth way of taming wicked problems. Roberts suggested that taming can occur if the "authority to define a problem and come up with a solution" (2000, p. 4) is given to a few stakeholders (such as a scientific committee, a CEO, or the supreme court) rather than addressing a problem collaboratively (for example in alliances, partnerships, or joint ventures).

Research focusing on identifying gaps

As described in the previous sections, numerous articles have been published in which the wicked problems concept has been applied and refined, and in which strategies for addressing wicked problems have been suggested and discussed. However, the number of peer reviewed, cited publications on wicked problems remains “modest” (Xiang, 2013, p. 2). One reason may be that much of the existing research on wicked problems focuses on “raising awareness, preaching for acceptance [of the concept], and advocating creative adaptation strategies and innovative approaches”. Thus, “as substantive as it may seem, [much of the research] remains largely a repetitive description of the social reality of wickedness, rather than well-grounded theoretical explorations or empirical investigations” (Xiang, 2013, p. 2).

Several scholars have also argued that there is a lack of convergence in the literature on wicked problems. For example, Duckett, Feliciano, Martin-Ortega, and Munoz-Rojas found that many of the strategies that have been proposed for addressing wicked problems “only allowed a partial mapping” to the characteristics of wicked problems because all characteristics “were not necessarily tackled [with each strategy]” (2016, p. 45). This indicates that the suggested strategies have not yet been consolidated to integrative strategies that could address *all* characteristics of wicked problems. As described above, there have also been multiple but independent attempts at reformulating or condensing Rittel and Webber’s original description of wicked problems. Consolidation in that literature is restricted to acknowledging that other authors have presented different sets of reformulated criteria (e.g. Duckett, Feliciano, Martin-Ortega, & Munoz-Rojas, 2016). Turnbull and Hoppe go as far as claiming that ”there has been no convergence whatsoever on which problems are wicked nor what we should do about them” (2017, p. 4). One reason for this lack of convergence may be that authors who publish about wicked problems are “geographically scattered (…) across the world” (Xiang, 2013, p. 2) rather than concentrated to specialized research groups. Another reason may be that many alternative concepts are used in the literature to describe similar ideas (see the next section for a description of some of these concepts). Xiang noted that “many people have been, and may well continue to be, working with wicked problems without knowing it or without calling it as such” (2013, p. 3). This lack of connectedness of research on similar concepts may further inhibit consolidation and convergence of research that, in principle, has much in common.

Research on related concepts

Ill-structured and ill-defined problems

The terms “ill-structured problems” and “ill-defined problems” have been used to describe problems that

- involve a high degree of uncertainty (Jonassen, 1997);
- lack definite right or wrong solutions (Cho & Jonassen, 2002; Kitchener, 1983; Simon, 1981; Voss, Greene, Post, & Penner, 1983);
- are highly contextualized (Jonassen, 1997);
- involve political considerations (Fernandes & Simon, 1999);
- are characterized by a high level of inherent ambiguity and conflicting norms and values (Jonassen, 1997; King & Kitchener, 1994); and

- involve unclear goals, unstated constraints, and multiple criteria for evaluating solutions (Jonassen, 1997; Voss, 1987).

These characteristics have much in common with Rittel and Webber's description of characteristics of wicked problems. According to a recent search in the Scopus database (Table 3), both terms are used more frequently in literature related to engineering education than the term "wicked problems". However, the term "wicked problems" is much more common in literature related to sustainability education.

Table 3. Number of hits in the Scopus database for different search term combinations (row headings + column headings); the search was performed on 29 June 2017.

Search terms	**"engineering education"**	**"sustainability" + "education"**
"wicked problem*"	17	33
"ill-structured problem*"	47	2
"ill-defined problem*"	24	0
"complex problem*"	289	46
"design problem*"	602	19
"grand challenge*"	116	39
"real-world problem*"	465	35
"big problem*"	289	46
"socio-scientific issue*"	1	17
"complex sustainability problem*"	1	5
"wicked sustainability problem*"	2	3

One reason for the limited use of the term ill-structured problems in the literature related to sustainability education may be that the term sometimes is used with positivist undertones. For example, Simon's description of the term (1981) is based on a "positivist or empiricist philosophy" (Buchanan, 1992, p. 19). Buchanan argued that "Simon's methods are still analytic, directed toward the discovery of solutions in some sense already known rather than the invention of solutions yet unknown" (1992, p. 19). Another argument against using the terms ill-structured and ill-defined problems is that both are more narrow than the term wicked problems since they focus on only *one* characteristic of problems (i.e. either problem structure or problem definition). Wicked problems can thus be seen as *a type of ill-structured and ill-defined problems* for which several other characteristics are important, for example the presence of value conflicts (Norton, 2012).

Complex problems

The term "complex problems" is frequently used in the literature related to both engineering education and sustainability education (Table 3). However, Andersson, Törnberg, and Törnberg (2014) argued that *complexity* is only one of two important dimensions of wickedness, the other being *complicatedness*. They further argued that commonly used strategies to address problem complexity are inappropriate for addressing problem complicatedness – and thus for addressing wicked problems.

Design problems

Of all the terms reviewed here, the term "design problems" is most frequently used in the literature related to engineering education (Table 3). Design problems are also often described as wicked problems (e.g. Buchanan, 1992; Burge & McCall, 2014; Cherry, 1999; Coyne, 2004; Cross, 1984; Dorst, 2004; Dorst, 2006; Nelson, 2003; Rith & Dubberly, 2007; Schön, 1987). However, Buchanan (1992) noted that there are many different understandings of "design" because design researchers have backgrounds in diverse professions and academic disciplines. Buchanan suggested that "Rittel argued that *most* [but not all] of the problems addressed by designers are wicked problems" (1992, p. 15, emphasis added). In fact, McCall and Burge (2016) identified mismatches between Rittel and Webber's description of wicked problems and certain kinds of design problems. For example, as described above, McCall and Burge argued that the descriptions of some of the characteristics in Rittel and Webber's theory are inappropriate for problems such as small-scale product design and problems in software design (see p. 21 in this thesis). McCall and Burge met this discrepancy with suggestions to reformulate the relevant characteristics of wicked problems. Another approach would be to conclude, in line with Buchanan's suggestion, that *not all design problems are wicked*.

Grand challenges

Another term that is frequently used in the literature related to engineering education and sustainability education is "grand challenges" (Table 3). The term refers to "a family of initiatives fostering innovation to solve key global health and development problems" (Grand Challenges, 2003-2016). The frequent use of the concept could indicate wide acceptance. However, in the engineering education research literature, the concept has been critiqued on ideological grounds; in a special issue in the International Journal of Engineering, Social Justice, and Peace, several authors argued that the concept is flawed because of "authorial particularism, double standards in engineering's contributions to these challenges, bracketing of the 'social' from 'technical' realms, and deterministic definitions of progress" (Cech, 2012).

Real-world problems and big problems

The term "real-world problems" is widely used in the literature related to engineering education and sustainability education (Table 3). Real-world engineering problems are also frequently described as wicked problems (e.g. Bozic, Escalas Tramullas, Čizmić, & Pavlović, 2014; Jonassen, Strobel, & Beng Lee, 2006; Nagel, Pierrakos, Zilberberg, & McVay, 2012). However, the term is very unspecific; theoretically, it could be used to describe the real-world problem of brushing one's teeth in the morning. The term "big problems" is equally broad and could, for example, refer to non-wicked problems such as computational problems.

Socio-scientific issues

While there is a large body of research on socio-scientific issues in the science education literature, publications related to engineering education and sustainability education seem to be limited (Table 3). The term also excludes problems that are not directly related to scientific issues and is therefore unnecessarily narrow.

In this thesis, I have chosen to use the concept of wicked problems because it allows me to focus my research on important characteristics such as the lack of problem structure and definition and the presence of value conflicts. In Papers I-III, we have used alternative terms: *complex sustainability problems* in Paper I, and *wicked sustainability problems* in Papers II and III. However, these terms are not commonly used in the literature (Table 3). The meaning of those terms in the papers is equivalent to the meaning of the term *wicked problems* as described by Rittel and Webber and as used in this thesis.

Research focusing on critiquing the theory

Rather than attempting to improve research that uses wicked problems theory (or any of the alternative concepts), Turnbull and Hoppe critiqued the wicked problems concept on theoretical and philosophical grounds. They argued that "the wicked/tame problem distinction is simply the old distinction between social and natural sciences, rewritten in the language of policy and planning" (2017, p. 5) and that this distinction itself is flawed; scientific problems are "far from tame" because "scientific practice is a social process, one conducted through intuitive, habituated actions, infused with socio-cultural norms and bound up in political conflicts" (*ibid.*, p. 7)[4]. On these grounds, Turnbull and Hoppe questioned the validity of the ontological distinction between wicked and tame problems:

> If not only 'social' problems can be 'wicked', but problems of the natural sciences as well, then the distinction between the 'ill-defined' problems of the social sciences and the 'well-defined' or 'tame' problems of science (i.e. logic, mathematics and the sciences of the non-human world, especially physics) and engineering is also obsolete. (*ibid.*, p. 8)

Turnbull and Hoppe further argued that "the ontological reading of wicked problems promoted by Rittel and Webber" entails the problematic assumption that it is possible to analyze wicked problems from a neutral and detached position:

> The presumption is that problems can be analyzed from above, as though ontologically distinctive and abstractable from social activity around those problems. That is, the problem itself is assumed to have an autonomous, unique nature of its own, much like a scientist observing bacteria through a microscope. This isolates the problem itself from the surrounding context, including from the observer. This analytical ontology combines reductionist thinking about problems with the decontextualisation of policy analysis via the 'view from nowhere'. (*ibid.*, p. 8)

Turnbull and Hoppe concluded that "as much as Rittel and Webber outlined an alternative to the systems [analytics] view, they were also constrained by it" (*ibid.*, p. 25); thus, using wicked problems theory uncritically can perpetuate the reductionist paradigm that it was designed to overcome. According to Turnbull and Hoppe, "generations of scholars" have used the concept in an uncritical way as they "followed in playing this 'scientific' language game of essentializing and ontologizing 'wicked' problems via attempts at classifying problems by looking for the (relative) presence of their properties" (*ibid.*, p. 5) for the purpose of "draw[ing] attention and resources to certain problems" (*ibid.*, p. 2). Rather than being a valuable tool for rigorous research, Turnbull and Hoppe argued that the term "wicked" has "*rhetorical* appeal" (*ibid.*, p. 2, emphasis added) since it "elicits our natural curiosity about 'the most difficult' policy problems" (*ibid.*, p. 4) and since it "resists precise definition" (*ibid.*, p. 4):

---

[4] Farrell and Hooker (2013) made a similar argument when they described science problems as wicked. But rather than arguing that the wicked problems concept therefore does not hold any value, Farrell and Hooker used the concept to argue for a different understanding of, and approach to, science problems.

> The term can be applied to many different policy problems and put to many different uses, by practitioners, as well as academics. The many articles on wicked problems – each with their own adaptation of Rittel and Webber's original definition, along with their own recommendations for solutions – demonstrate the fecundity of the term for generating new work, but also the problems with it, because there has been no convergence whatsoever on which problems are wicked nor what we should do about them (Head, 2008; Head, 2016; Nordegraaf, Douglas, Geuijen, & van der Steen, 2016). (Turnbull & Hoppe, 2017, p. 4)

Rationale for using wicked problems theory in this thesis

Turnbull and Hoppe's (2017) critique of the wicked problems theory is relevant and important. However, it is written from the perspective of social policy research, a field in which reductionistic, systems analytics approaches to problem-solving have been critiqued for a long time (Turnbull & Hoppe, 2017). In engineering and EER, on the other hand, the dominant research paradigm is still post-positivism and systems analytic approaches are seldom questioned (see p. 10 in this thesis). I therefore argue that the concept may do "similar work" in EER today as it did in the field of social policy in the 1970s; even if the concept to some degree is based on a reductionistic paradigm, it has so far been successful in raising awareness about limitations of systems analytics approaches: "there has clearly been a steadily rising awareness of wicked problems and an increasingly broad – yet often reluctant – acceptance of their intractability among scholars, practitioners, stakeholders, and the general public" (Xiang, 2013, p. 2). This awareness of wicked problems and acceptance of their intractability "have already been, and will continue to be, enabling people to conscientiously give up the unrealistic hope for scientific solutions to tame the untamable so that they shift focus to the public process of working with wicked problems" (Xiang, 2013, p. 3). For example, Xiang noted that "the tone in which people describe their relationship with wicked problems has become progressively softer, from 'tame,' 'deal with,' 'handle,' 'tackle,' to even 'work with' (Australian Public Service Commission, 2007, pp. 11, 17, 35), 'live with' (Norton, 2012, p. 460), and 'embracing' (Raisio, 2010), reflecting a greater degree of acceptance of wicked problems as a sustained social reality that human society has to live with" (Xiang, 2013, p. 2).

My personal experience is that the concept indeed "elicits our [engineering educators' and engineering education researchers'] natural curiosity about 'the most difficult' policy [engineering] problems" (Turnbull & Hoppe, 2017, p. 4), and that this curiosity is an important asset in slowly increasing engineering educators' and engineering education researchers' awareness of alternative paradigms. For example, the concept has been used to argue that traditional pedagogical approaches (typically described as relying on well-structured problem-solving) are inadequate for training engineering students to address wicked problems and that, therefore, alternative pedagogies need to be explored, such as problem-based and project-based learning (e.g. Guerra & Kolmos, 2012; Hess, Brownell, & Dale, 2014; Kanematsu & Barry, 2016).

Finally, counter to Turnbull and Hoppe's argument that the wicked problems concept "was a product of its time" (2017, p. 25), other scholars have argued that wicked problems are becoming more common and that the theory, therefore, becomes more and more relevant. For example, Norton (2012) argued that environmental problems become increasingly open-ended, complex, and situation-dependent. Similarly, Roberts (2000) argued that value differences become more visible and influential in an increasingly globally connected world; that the developments in information technology allow more people to express their own values and to actively engage in discussing and addressing wicked problems; and that increasing decentralization weakens traditional authority and thus limits the usability of authoritative approaches to addressing wicked problems.

Because of the power of the concept to elicit curiosity, raise awareness, and stimulate conversations among engineering educators and engineering education researchers, I have chosen to use the concept despite Turnbull and Hoppe's (2017) relevant criticism regarding the theoretical underpinnings of the concept. In the long run, however, I hope that engineering and EER will develop a larger degree of awareness of theoretical and methodological perspectives such that it is no longer necessary (or even fruitful) to rely on rhetorically engaging but theoretically imperfect concepts such as wicked problems.

PART 1

In this chapter, I describe the research in Part 1, which aimed to answer the question of *what engineering students need to learn to be able to address wicked problems*. The chapter consists of four sections that describe the pre-study literature review, Study 1, Study 2, and the synthesis of Studies 1 and 2 respectively. In the section on the pre-study literature review, I address RQs 1.1-1.2. In the section on Study 1, I address RQs 1.3-1.4, and in the section on Study 2, I address RQ 1.5. Finally, in the section on the synthesis of Studies 1 and 2, I address RQs 1.6-1.7.

PRE-STUDY

What characteristics of wicked problems are particularly challenging for engineering students? (RQ 1.1)

Addressing wicked problems requires recognizing and dealing with uncertainty, ambiguity, and conflicting values. It also requires integrating several topics, disciplines, and perspectives, and it requires iteratively shifting between different tasks (Jonassen, Strobel, & Beng Lee, 2006; Norton, 2012; Rittel & Webber, 1973; Seager, Selinger, & Wiek, 2012). Kahneman (2011) compiled convincing evidence from psychological research that these processes are challenging tasks for all humans, not only engineering students. However, to best support engineering students in developing their ability to integratively address wicked problems, it is valuable to develop a better understanding of what characteristics of wicked problems may pose significant challenges for engineering students. The characteristics of wicked problems are described in detail on p. 19. For the purpose of the discussion in this section, they can be summarized as follows:

1. There is no definitive formulation of wicked problems.
2. Wicked problems have no stopping rule.
3. Solutions to wicked problems are not true-or-false, but good-or-bad.
4. There is no immediate and no ultimate test of a solution to a wicked problem.
5. Every solution to a wicked problem is a 'one-shot operation'.
6. There are no exhaustible sets of solutions or methods from which to choose the most appropriate one.
7. Every wicked problem is essentially unique.
8. Every wicked problem can be considered to be a symptom of another problem.
9. Descriptions of wicked problems, and preferences for how they should be addressed, are strongly influenced by questions related to norms and values.
10. The planner has no right to be wrong.

Methods

To answer RQ 1.1, I reviewed the literature on general differences between academic fields and on specific aspects of engineering culture and identity, such as epistemic beliefs and assumptions about the relationship between engineering and other sectors of society. During this review, I focused on identifying *discipline-specific challenges* that engineering students may face when learning to address wicked problems.

Results and conclusions

Engineering as a high-paradigmatic field

> It can be argued, as within any community of practice, that engineering students as well as practitioners and educators live within a kind of 'common sense' that they have developed from their teachers and books and from the external social constructs of their society. (Baillie & Armstrong, 2013, p. 136)

As discussed above (p. 10), engineering research has been described as a "hard" or "high-paradigmatic" field, which is "characterized by the existence of paradigms that specify the appropriate problems for study and the appropriate methods to be used" (Biglan, 1973, p. 195; see also Gardner & Willey, 2016, and Paulsen & Wells, 1998). Studies of faculty behavior in different academic fields support the conclusion that diversity of opinions in general, and diversity of research approaches in particular, may not be highly valued in hard fields such as engineering (Braxton & Hargens, 1996, and Lattuca & Stark, 1995; both cited in Paulsen & Wells, 1998). In other words, the existence of a strong engineering paradigm seems to create a disciplinary culture in which diversity of perspectives and worldviews is not highly valued. Such a high-paradigmatic culture may create a number of challenges for those who attempt to address wicked problems. First, since the paradigm itself defines which problems should be studied and what methods should be used to study these problems, a high-paradigmatic culture does not provide the openness and flexibility that is required if Rittel and Webber's (1973) first and sixth criteria for wicked problems are to be taken seriously. Further, since high-paradigmaticity precludes alternative perspectives and worldviews, it precludes an understanding of different worldviews that could explain and address discrepancies in problem formulation and conflicting preferences for solution approaches. Thus, a high-paradigmatic culture cannot adequately address the ninth characteristic of wicked problems.

Students' view of knowledge as certain

Jehng (1993) proposed that disciplinary culture can influence students' epistemic beliefs, i.e. their beliefs about the nature of knowledge. In high-paradigmatic fields, where awareness of alternative paradigms often is low, students may develop an understanding of knowledge as "true" and "factual" rather than provisionary, uncertain, and subjective:

> The high degree of consensus among scholars [in high-paradigmatic fields] regarding the content and methods of the field (…) may communicate to students that knowledge is certain due to the various ways it is presented in classes and textbooks. (Paulsen & Wells, 1998, p. 377)

On the other hand, a stronger focus on ambiguity, diverse methodologies and viewpoints, and critical thinking in low-paradigmatic fields may support a view of knowledge as diverse, tentative, and open to change. Empirical studies of students' epistemic understandings have shown that students in high-paradigmatic fields, such as engineering, are more likely than students in low-paradigmatic fields to believe that knowledge is certain and unchanging (Jehng, 1993; Paulsen & Wells, 1998).

Viewing knowledge as certain again creates difficulties in dealing with several of the characteristics of wicked problems outlined by Rittel and Webber. First, such a view of knowledge is related to believing that it is possible to find unambiguous descriptions of problems and "correct", final solutions that can be evaluated with given criteria. This violates the first four characteristics of wicked problems. Second, similar to high-paradigmaticity, viewing knowledge as certain and excludes alternative perspectives and worldviews. Thus, such a view of knowledge renders discrepancies between different actors' descriptions of problems, as well as between different preferences for solution approaches and different evaluation criteria, unintelligible and thus impossible to address (characteristics four and nine). Finally, viewing knowledge as certain is only meaningful if knowledge is also viewed as generalizable to new contexts. Such a belief would be in direct contrast with an understanding of wicked problems as essentially unique (characteristic seven).

Difficulty in taking multiple perspectives, modifying one's own thinking, and recognizing complexity

A tendency to view knowledge as certain is related to other student difficulties that are relevant to addressing wicked problems. In separate empirical studies, Liu, Lin, and Tsai (2011) and Schommer-Aikins and Hutter (2002) found that students who tended to view knowledge as certain were less likely than those who believed in tentative knowledge to take multiple perspectives, modify their own thinking, and recognize complexity in decision-making processes. As described above, an ability to take multiple perspectives is needed to understand and address discrepancies in problem formulation, solution approaches, and solution evaluation (characteristics one, three, four, and nine). An ability to modify one's own thinking is needed to cope with the constantly changing nature and context of wicked problems (characteristics one, two, four, and eight). Finally, an ability to recognize complexity is needed to understand how wicked problems are related to other problems and to anticipate possible consequences of actions that aim to address wicked problems (characteristics eight and ten).

Difficulty in questioning authorities

Liu, Lin, and Tsai (2011) and Schommer-Aikins and Hutter (2002) also found that students who viewed knowledge as certain were less likely to question authorities than students who believed in tentative knowledge. Similarly, Jehng found that students in high-paradigmatic fields were more likely to believe in "omniscient authority" (1993, p. 33) than students in low-paradigmatic fields. Students who believed in omniscient authority regarded expert advice as "highly credible" and viewed knowledge as "handed down by teachers and other experts rather than formed by independent reasoning" (*ibid.*, p. 26). These results are in line with suggestions that engineering as a profession is characterized by a high degree of loyalty towards authorities and traditions (Riley, 2008). Gambetta and Hertog (2016) even suggested that engineers are over-represented among right-wing extremist groups and that this over-representation reflects an engineering mindset that is characterized by a desire for order and hierarchy.

Similarly, Paretti and McNair described how a narrow focus on technical problem-solving can be related to an understanding of engineering work as an activity that is "responsive rather than proactive" (2012, pp. 63-64). In a case study of how engineering identities were enacted in two different engineering contexts, Paretti and McNair described that in one of these contexts, a consumer goods manufacturing company, engineering work was enacted as "implementing a vision created and constrained by others". According to this vision, "engineering input focuses not on contributing to the design ideas, but on applying technical analysis to meet demands for performance, esthetics, and release dates specified by marketing, within the production constraints specified by manufacturing". While the engineers in the study anticipated and planned for decisions from marketing and manufacturing, they did so only "covertly" by building in flexibility into product features that would allow them to later adjust to changing requirements from marketing or manufacturing; the engineers, in this specific case, did not "actively or overtly" contribute to important decisions, nor challenge external requirements (*ibid*., p. 61).

Similar to an understanding of problems and methods as defined by a strong paradigm, an understanding of problems and methods as dictated by authorities defies engagement with iterative and perspective-dependent problem formulation (characteristics one and nine), subjective judgment of solution quality (characteristics three, four, and nine), and unlimited possibilities for ways of approaching a problem (characteristic six). Critical engagement with values and controversies becomes impossible, which again renders characteristic nine unintelligible and impossible to address.

Engineering as purely scientific and technical

In high-paradigmatic fields, the dominant paradigm can be expected to strongly influence work in the domain. As discussed above (p. 10), the dominant paradigm in engineering is typically described as (post-)positivist, where

> the traditional way of conducting research involves the use of the 'scientific method'. According to this method, testable hypotheses need to be generated, and these are either confirmed or refuted by means of objective investigations. This involves the researcher adopting a position as a neutral and objective observer. (Jawitz & Case, 2009, pp. 151-152)

It also involves heavy reliance on quantitative research approaches (see also Gardner & Willey, 2016; Riley, 2008).

Riley suggested that a positivist approach often leads to reductionism, i.e. "the notion that phenomena (or problems) can be broken down into smaller components for analysis and that analysis of the components can fully explain the system as a whole". In engineering, Riley argued, "a reductionist perspective is evident in the engineering problem-solving and engineering design processes" and in engineering education where students "learn to break problems down into small parts, solve the individual parts, and then work back up to a solution" (2008, p. 41; see also Downey & Lucena, 1997).

In fact, engineering is often conceived of as a purely scientific and technical discipline that values scientific facts and ignores social values. For example, Cech suggested that engineering is characterized by an "ideology of depoliticization", which Cech described as "the belief that engineering work can and should be disconnected from 'social' and 'political' concerns because such considerations may bias otherwise 'pure' engineering practice" (2014, p. 45). Riley made a similar point when she argued that engineering culture values "uncompromising objectivity" at the expense of making subjective judgments (2008, p. 38). This disconnect between engineering work on the one hand and social or political concerns on the other hand may lead to a dualistic understanding of technical and social domains according to which engineers would not be expected to have social competencies and should not engage with social issues such as public welfare (Cech, 2014).

In the context of engineering education, research has shown that students enter engineering with a genuine desire to help others (Bielefeldt & Canney, 2016) and to contribute to environmental and social causes (Haase, 2014). In fact, the percentage of incoming students who want to help others has increased in the past decade (Bielefeldt & Canney, 2016). However, Cech (2014) found that students' interest in public welfare concerns may decrease over the course of their education. Similarly, Bielefeld and Canney (2016) found that engineering students' confidence in their own ability to help others decreased over time. Cech concluded that "engineering education fosters a culture of disengagement that defines public welfare concerns as tangential to what it means to practice engineering" (2014, p. 45), i.e. a culture in which public welfare concerns are not considered to be at the core of engineering work. In such a culture, Cech further suggested that "while reactionary efforts such as whistle blowing are generally considered valid and important (Dorn, 2011), proactive contemplation of the broader social impacts of technology is often explicitly excluded from the purview of engineers' day-to-day design activities" (Cech, 2014, p. 49; see also the above discussion of difficulties to question authorities, p. 33 in this thesis). Bielefeldt and Canney also suggested that this culture of disengagement is not only found within the profession but also among the wider public: "Public perception of the engineering profession does not include the recognition that engineers are engaged with societal and community concerns" (2016, p. 1536).

Understanding engineering as purely technical creates several challenges for addressing wicked problems. First, without considering social aspects of wicked problems, it becomes impossible to evaluate a "good-or-bad" solution to a wicked problem (characteristic three). Second, understanding engineering as purely technical again renders conflicting values and perspectives unintelligible and impossible to address (characteristic nine). And third, not taking responsibility for social concerns violates characteristic ten, which ascribes social responsibility to those who address (or neglect to address) wicked problems.

Lack of training to address wicked problems

In a typology of problem types, Jonassen (2000) organized different kinds of problems on a continuum from well-structured to ill-structured problems. At the well-structured end of the continuum, he listed logical problems (abstract logical puzzles such as the Tower of Hanoi challenge or the Rubic's Cube™), algorithmic problems (mathematical problems such as long divisions or equation factoring), and story problems (algorithmic problems embedded in a "story", such as how long it takes for one car to overtake another car given certain traveling speeds). At the ill-structured end, Jonassen listed case analysis problems (analyzing specific cases in authentic contexts, such as legal cases or managerial problem-solving), design problems (designing artifacts or processes in response to constraints from multiple stakeholders, such as designing a bridge or developing an investment strategy), and dilemmas (social or ethical dilemmas that require balancing conflicting needs and values, such as resource management and pollution). In another study, Jonassen, Strobel, and Beng Lee (2006) found that the most common type of problems used in engineering education are well-structured *story problems*, despite the fact that engineering workplace problems have much more in common with ill-structured design problems. Addressing story problems

> [requires students to] translate relationships about unknowns into equations, solve the equations to find the value of the unknowns, and check the values found to see if they satisfy the original problem. This linear process implies that solving problems is a procedure to be memorized, practiced, and habituated, a process that emphasizes getting answers over making meaning. (Jonassen, Strobel, & Beng Lee, 2006, p. 139)

Engineering workplace problems, on the other hand, require students to manage and address "conflicting goals, multiple solution methods, non-engineering success standards, non-engineering constraints, unanticipated problems, distributed knowledge, and collaborative activity that rely on multiple forms of problem representation" (Jonassen, Strobel, & Beng Lee, 2006, p. 148).

Possibly due to these highly different processes in solving well-structured problems and addressing ill-structured problems, an ability to solve well-structured problems does not readily transfer to an ability to address ill-structured problems. Research on problem-solving skills and processes has found that addressing ill-structured problems such as wicked problems involves different skills and cognitive processes (An & Cao, 2014; Schraw, Dunkle, & Bendixen, 1995; Shin, Jonassen, & McGee, 2003) and requires more elaborate epistemic understandings (Bendixen, Dunkle, & Schraw, 1994; King & Kitchener, 1994; Kitchener, 1983) than addressing well-structured problems.

As described above (p. 32), engineering students often hold more naïve epistemic beliefs than students in low-paradigmatic fields. Thus, engineering students are less likely to be able to adequately address wicked problems without extensive training. However, since engineering students typically learn to solve story problems rather than to address wicked problems, they often do not receive sufficient training in addressing wicked problems. In addition, engineering students often do not receive adequate training in problem *scoping*. In a study of how students in an interdisciplinary course approach social justice, Kabo (2010) found that humanities students primarily tended to focus on problem definition and deconstruction, while engineering students tended to focus on problem-solving and construction. Other scholars have argued that iterative problem definition should receive greater attention in engineering education (see e.g. Atman *et.al*., 2008; Downey, 2005).

This lack of training to define and address wicked problems renders characteristics five and ten of wicked problems especially problematic for engineering students. With a lack of training, it is natural to make mistakes. However, according to characteristic ten, "the planner has no right to be wrong" and "planners are liable for the consequences of the actions they generate". Similarly, characteristic five states that it is not possible to use a trial-and-error approach for addressing wicked problems. Thus, it is natural that, as described above, engineering students' confidence in their own ability to address social concerns *diminishes* rather than increases during the course of their education (Bielefeldt & Canney, 2016).

Summary

Table 4 provides a summary of the results. Based on these results, it is reasonable to conclude that engineering students may need specific training and carefully designed instructional guidance to overcome discipline-specific challenges in addressing wicked problems. The following characteristics seem to pose the strongest discipline-specific challenges for engineering students:

- the lack of definite problem definitions (characteristic one),
- the lack of "right" solutions (characteristic three), and
- the need to act in the presence of uncertainty, normativity, and ambiguity (characteristics four, nine, and ten).

Table 4. Overview of what characteristics of wicked problems may pose significant challenges for engineering students (marked in gray), given certain characteristics of engineering culture and engineering education.

Characteristic of engineering culture and/or engineering education	**Characteristics of wicked problems that are difficult to address**									
	1	**2**	**3**	**4**	**5**	**6**	**7**	**8**	**9**	**10**
Engineering as a high-paradigmatic field	■					■			■	
Students' view of knowledge as certain	■	■	■	■			■		■	
Difficulty in taking multiple perspectives	■		■	■					■	
Difficulty in modifying one's own thinking	■	■		■				■		
Difficulty in recognizing complexity								■		■
Difficulty in questioning authority	■		■	■		■			■	
Engineering as purely scientific and technical			■						■	■
Lack of training to address wicked problems					■					■

What could be relevant examples of wicked problems for undergraduate engineering students and how could these problems be described for the students? (RQ 1.2)

My approach to this question was informed by my practical experiences during the undergraduate engineering course on wicked problems that I got involved in before starting my Ph.D. studies. In that course, students are asked to address a given wicked problem (Lönngren *et.al.*, 2010). Every year, the teaching team identifies a new problem for the course. To present this problem to the students, the teaching team develops a written *description* of the chosen wicked problem; this description is then used to frame discussions about the problem during the course. Due to the complexity and openness of wicked problems, wicked problem descriptions can never be complete and "true" representations of "actual" problems. Rather, the descriptions provide a limited selection of relevant (according to the judgment made by the teaching team) background information about the problem and a description of the task(s) that the students are expected to perform in relation to the problem. Throughout this thesis, I will use the term *wicked problem description* whenever I need to specifically refer to the textual artifact of a written description. I will use the term *wicked problem* in all other cases, i.e. even when students engage with wicked problems through the textual artifacts that represent (but not completely define) these problems.

During the first three years in which the course was given, the teaching team also engaged in a process of collaborative action research with the aim to identify important characteristics of wicked problem descriptions for engineering education (Lönngren, Jacobsson, Mårsell, & Nilsson, 2012). Empirical material for this action research project included personal reflections and focus group discussions in the teaching team, as well as students' course evaluations from three consecutive years in which the course was taught (2010-2012). We observed that it is challenging to design good problem descriptions. For us, particular challenges included designing problem descriptions that could help students to

- appreciate the relevance of the problem for their specific engineering discipline – and the relevance of their discipline for addressing the problem,
- engage with social and global aspects of the problem rather than predominantly focusing on technological aspects, and
- acknowledge and address conflicting values and interests rather than strive for consensus.

To address the first challenge, we tried to include explicit and implicit references to the students' discipline in the problem description. For our group of students in engineering nanoscience, we for example developed cases related to a research facility for nanotechnology research, nanomaterial coating of buildings, and climate engineering with nanoparticles. To address the second challenge, we simultaneously tried to ensure that the technological focus did not become too strong and that students still were required to engage with a broad set of non-technological perspectives. We did this primarily in the activities in which the problem description was used. For example, we used role-playing activities in which students were required to represent different stakeholder perspectives. However, students tended to redefine their roles such that the different stakeholder groups would reach an unproblematic consensus. Thus, to address the third challenge, we actively tried to adjust the power balance between the fictive stakeholder groups, for example by including information in the problem descriptions that suggested that certain stakeholders have more or less power than could be intuitively expected, but also by providing instructional guidance to groups that seemed to assume a low-power position in discussions about the problem.

Another important background for my approach to RQ 1.2 was a conference paper that introduced guidelines for designing problem-based learning (PBL) problems for ESE in engineering education. In that paper, Dobson and Tomkinson described "some of the important features to take into account in designing [ESE] scenarios for complex global issues" (2010, p. 9). They suggested that problems should

- match Rittel and Webber's characteristics of wicked problems;
- represent authentic, unsolved problems;
- address all three dimensions of sustainable development (ecological, social, and economical development);
- be framed as interdisciplinary challenges that cannot be reduced to technical design nor fragmented into parts that can be delegated to different disciplinary experts;
- allow students to train skills that they will need in their future professional lives;
- include both local and global perspectives, and bottom-up and top-down approaches;
- provide an adequate level of challenge for the specific student cohort;
- not advocate or devalue specific perspectives; and
- not include topics that students may experience as sensitive or discriminating.

Dobson and Tomkinson did not explicitly address the question of how such problems should be described for the students and how problem descriptions should be used in course activities.

Wicked problem descriptions used in this thesis

In the context of the research described in this thesis, I developed three descriptions of wicked problems that I used in different research and teaching contexts. At the beginning of my Ph.D. project, I developed a description of *water-shortage in Jordan*. I used this description in in-depth interviews with undergraduate engineering students who studied engineering nanoscience (Studies 1 and 2). For this problem, I also developed a set of idealized solution alternatives and question prompts which I used in the interviews to stimulate participants to consider the problem from different perspectives. Towards the end of my project, I developed descriptions of *literacy in Afghanistan* and *dengue fever in Sub-Saharan Africa*. I used these two descriptions in written exams in the context of a sustainable development course for students in information technology (IT) engineering (Study 4). These problems can serve as an inspiration for engineering educators who want to develop and use wicked problem descriptions in their teaching.

Water-shortage in Jordan

> Jordan is a country in the Middle East that is classified as an “upper middle income” country by the World Bank. The country has a stable and growing economy, and a number of free trade agreements with other countries all over the world. The developmental standard, and the standard of living, are high in a global comparison.
>
> But Jordan’s climate is dry, especially in the eastern parts of the country. It is unclear whether there will be enough water to support the 6.5 billion inhabitants in the future. Jordan is one of the world’s most vulnerable countries in terms of water shortage.
>
> In 2007, the annual water demand was estimated to be 1505 billion cubic meters. This number is expected to further increase and reach 1635 billion cubic meters in 2020. Today’s water resources are estimated to amount to 665 billion cubic meters annually. The difference between assets and demands is currently bridged by over-exploiting ground water resources.
>
> Apart from these natural limits, water supply is also restricted by an agreement among the countries that surround the Jordan valley: Jordan, Israel, Lebanon, and Syria. According to the “Jordan Unified Water Plan”, which was signed in 1955, these countries have specific water allocations that they are allowed to withdraw from the streams in the valley.
>
> Solution alternatives:
>
> 1. Dig deeper wells in order to get hold of more ground water.
> 2. Build desalinations plants in order to make use of the water in the Dead Sea.
> 3. Import water from, e.g., Sweden and Norway where there is a surplus.
> 4. Breach the Johnston agreement by withdrawing more water from the Jordan valley.
> 5. Liberalize the water market to achieve a water price that reflects the balance between assets and demands.
> 6. Nationalize the water market and limit the water usage to an entirely renewable amount (665 million m^3) through an equitable distribution plan.

Literacy in Afghanistan

> Assume the following professional role: You are employed at a large, international IT company. The company has a broad portfolio of services and products, and it wants to be perceived as a company that actively works to support all three dimensions of sustainable development.

The company's sustainability department continuously monitors developments in society to identify possible issues in which the company could get involved. The department has collected the following information about the educational situation in Afghanistan from the Internet:

On 25 September 2015, the UN international conference adopted 17 global Sustainable Development Goals. Goal 4 is to "ensure inclusive and quality education for all and promote lifelong learning" by the year 2030. This goal includes an ambition to ensure that all people should have access to inclusive and high-quality primary and secondary education. Current population statistics suggest that, worldwide, 250 million children in 4th grade and 774 million people overall lack basic reading and writing skills. In the latter group, two-thirds are women.

An example of a country that faces major problems in relation to the above-stated goal is Afghanistan. Current estimates suggest that only 38% of Afghanistan's population is able to read and write; for women, the percentage is only 24%. Reading and writing literacy is generally lower in remote areas of the country than in urban areas.

Afghanistan has a population of 32 million, and 76% of the population lives in rural areas. The capital, Kabul, has 3.3 million inhabitants. All other cities in the country have fewer than 500 000 inhabitants. According to the Human Development Index, Afghanistan was the 15th least developed country in the world in 2011. One of the factors that hinder development in the country is deficient energy supply, especially in rural areas. Another factor is many years of violent conflict. For example, Afghanistan is one of the world's most affected countries with regard to violence against schools. Between 2006 and 2009, 439 teachers and other school personnel have been killed. The country faces a severe lack of qualified teachers.

The sustainability department suggests that the above issue could be a suitable focus for your company. You are asked to investigate how the company could contribute to improving the educational situation in Afghanistan in a way that would support sustainable development.

Describe how you would approach the task of identifying suggestions for how your company could contribute to improving the situation. Describe also which factors of the situation you would take into account, and which improvement measures could be included in your suggestion. Motivate your answer.

Dengue fever in Sub-Saharan Africa

Dengue fever is a viral disease that is spread by mosquitoes in areas with tropical and subtropical climate. In recent decades, the number of dengue fever infections has increased dramatically. Recent research estimates that 390 million people become infected with dengue fever every year.

The World Health Organization (WHO) describes dengue fever as one of 17 neglected tropical diseases (NTDs) because it primarily affects people in poor parts of the world, and because it has not received as much attention as e.g. HIV, malaria, and tuberculosis. WHO has established a specific division that aims to combat NTDs. Also, the United Nations' Sustainable Development Goals contain a sub-target which includes an ambition to eradicate all NTDs by 2030.

Assume the following professional role: You are employed at WHO as an IT specialist. As one part of the organization's ambitions to eradicate all NTDs, WHO has created a number of interdisciplinary expert committees that are asked to focus on specific diseases and geographic areas. You have been assigned to work with a group that aims to combat dengue fever in sub-Saharan Africa. As a starting-point for the group's first meeting, you have been asked to investigate how IT could contribute to combating dengue fever in that area. Write a short summary of the problem situation and what you see as possible IT-related improvement measures. The text shall not exceed a maximum of 1000 words plus references.

Design principles for wicked problem descriptions

Despite several years of experience in working with wicked problem descriptions in engineering education, I experienced the development of the descriptions for my research, and the integration of these descriptions in my research and teaching, as extremely challenging. To support other educators and researchers in developing wicked problem descriptions and related course activities, I therefore decided to develop *design principles for wicked problem descriptions*. For this purpose, I combined Dobson and Tomkinson's descriptions of important features of PBL-problems for ESE in engineering education and my personal experiences of developing wicked problem descriptions for engineering students.

In developing the design principles, I assumed that wicked problem descriptions cannot be developed in isolation from the educational context in which they are to be used. In many cases, the description itself is not enough to, for example, ensure that students engage with social and global perspectives or to avoid that students gravitate towards an unproblematic consensus on fundamentally controversial questions. In these cases, the educational context needs to be designed to provide additional guidance for the students. The following design principles for wicked problem descriptions, therefore, also contain suggestions for how to design some aspects of the educational contexts in which the descriptions are to be used. To illustrate how the design principles can be used in practice, I discuss how each principle applies to one of the three descriptions of wicked problem that were used in this thesis (p. 39).

Principle 1. Ensure that the problem can be understood and discussed from many different perspectives.

It is important for students to learn to address wicked problems from multiple perspectives. One way to introduce diverse *disciplinary* perspectives could be to address wicked problems in interdisciplinary educational contexts. However, for many engineering educators, it is not feasible to create and teach interdisciplinary courses at their institutions. In these cases, problem designers need to pay particular attention to ensure that the problem description contains explicit and implicit references to multiple perspectives that together cover a broad range of thematic aspects related to the problem. Since engineering students seem to particularly struggle to consider social and global perspectives, designers should make an effort to stress these perspectives in the description. References to specific perspectives can be provided through background data. For example, social perspectives can be included by mentioning demographic information and information about for example distribution of resources among different societal actors.

The description of *literacy in Afghanistan* includes social and global perspectives in the form of demographic information for different population groups in Afghanistan, such as urban and rural populations, as well as women as a marginalized group in the Afghan society. The description includes different disciplinary perspectives by referring to both technical/infrastructural issues (energy supply) and social/political issues (conflicts, violence, poverty, lack of teachers). A global perspective is included by locating the problem in a culture that is very different from the culture that most of the students in the study were familiar with. To further support the potential of the description to help students engage with multiple perspectives in an educational context, it could for example be beneficial to design course activities in which the students are explicitly asked to discuss marginalized social perspectives that are mentioned in the problem description.

Principle 2. Ensure that conflicting values and interests cannot be ignored.

To avoid that students tame wicked problems by assuming, or artificially constructing, a consensus among stakeholders, problem designers need to explicitly describe some of the most important conflicts – or allocate time for course activities in which the students are guided to identify these conflicts themselves. These activities should be scheduled *before* students start to engage in any projects that aim to address the problem.

To collect ideas for suitable problems, it can be valuable to continuously monitor the public debate and take notes of contentious issues that seem to engage stakeholders with conflicting values and interests. When it is time to develop a problem description, problem designers can refer to their list and choose one or several of the most promising issues as a basis for their problem description.

In the description of *literacy in Afghanistan*, the presence of conflict is explicitly mentioned. While the conflicting values between those who support and those who oppose schools in Afghanistan are not explicitly stated, students are unlikely to ignore the information about the prevalence of violence against schools. To further support students in identifying and understanding the underlying value conflicts, it could be beneficial to ask students to work in groups to describe the main lines of conflict or to discuss how value conflicts influence efforts to address the problem.

Principle 3. Define an achievable "result" that does not allow definite answers or solutions to the problems.

Working with wicked problems necessarily includes uncertainty and ambiguity, and problem designers should make an effort not to tame the problem. Yet, students also need guidance to avoid frustration and maintain motivation. For this purpose, problem descriptions should include a description of the nature of the outcome that is expected from the students at the end of the course. This could for example be a multi-stakeholder-agreement on how to address a wicked problem or a policy draft for managing a wicked problem. Problem designers should ensure that it is impossible to achieve this outcome with a binary yes-or-no decision or a description of how a limited set of well-defined variables can be optimized. During course activities, teachers could provide additional support by explicitly discussing the type of result that is expected, for example by contrasting important features of policy drafts with important features of solutions to story problems.

In the description of *dengue fever in sub-Saharan Africa*, this principle is illustrated by the explicit description of a concrete professional role (IT specialist at WHO, part of an expert committee on dengue fever in sub-Saharan Africa), the expected output (a short summary of the problem situation and possible IT-related improvement measures), the recipient of the output (the other members of the expert committee), and the purpose of the output (to serve as a starting point for the first meeting of the expert committee).

Principle 4. Ensure that students can make a connection between the problem and their educational program.

Without proper guidance, not all students will appreciate the connection between multidimensional, complex sustainability problems and their disciplinary educational programs. To help students appreciate this connection, problem designers should make an effort to identify problems for which the connection is relatively obvious. For example, designers can focus on problems that are discussed (but have not yet been satisfactorily addressed) in the students' future professional context. Designers can also explicitly mention these discussions in the problem description or they can design a course activity in which the students actively and collaboratively explore the connection between the problem and the students' educational program.

The description of *dengue fever in sub-Saharan Africa* illustrates this principle by defining a professional role related to the students' field of study. While the description of the problem does not contain explicit references to IT, the problem has been discussed extensively in the professional IT community, and numerous IT-based solution approaches have already been suggested and/or implemented. I have used this problem description in the context of a take home exam, which allowed students to relatively easily find information about how IT could be (and already is) used to address the problem.

Principle 5. Design problem-specific support for students.

As mentioned above, it is important to carefully design not only wicked problem descriptions but also the educational contexts in which these descriptions are to be used. Since engineering students face discipline-specific challenges (p. 32) and since they have limited prior experience addressing wicked problems, it is not possible to simply provide students with a problem description and then ask them to address the problem without further guidance. A good problem description can highlight multiple perspectives, conflicting values, and interests, define an achievable "result", and highlight the connection between the problem and the disciplinary context. However, as students engage with the problem, they will also need continuous guidance to keep them on track towards the described "result" and to avoid that they tame the problem. Most importantly, students will likely need continuous support to avoid an overly technological focus and premature consensus. This support needs to be designed specifically for each problem; it could for example consist of helping students to engage with specific stakeholder perspectives or to deal with frustration about the lack of definite solutions.

The description of *water shortage in Jordan* illustrates this principle. The description includes explicit descriptions of alternative solution approaches. Two of these approaches are clearly technical (digging deeper wells and building desalination plants), and one approach can be interpreted as a technical approach (importing water). However, the remaining three approaches are mainly political and/or economical approaches. Explicitly stating these approaches provided support for students to realize that the problem could be approached from non-technological perspectives.

I used the description of water shortage in Jordan in in-depth interviews with students. During the interviews, I continuously asked the students to reconsider their choices of solution approaches. Thus, the discussion during the interviews provided additional support for the students to engage with different perspectives and avoid premature closure.

STUDY 1

What are relevant characteristics of perspectives in the context of learning to address wicked problems? What are some ways in which students interact with their perspectives? (RQs 1.3 and 1.4)

We addressed these two questions in Study 1. The aim of that study was to develop a theoretical framework that could facilitate discussions about engineering students' approaches to wicked problems. Our work with the framework was theoretically grounded in the phenomenological principle of intentionality (Brentano, 1874/2009), empirically grounded in the results of a qualitative content analysis of student interviews about a wicked problem, and practically grounded in our own experiences of trying to teach engineering students to use multiple perspectives when addressing wicked problems.

Research approach

Theoretical background for the framework

In the ESER literature, the concept of "perspectives" has frequently been used to describe the need to approach sustainability problems in many different contexts and from many different angles (e.g. Kates, Parris, & Leiserowitz, 2005; Svanström, Lozano-García, & Rowe, 2008; Wals & Blaze Corcoran, 2006). In a short review of existing descriptions of perspectives in the ESER literature, we observed that these descriptions mainly focus on *thematic* aspects of perspectives, such as different dimensions of a wicked problem that are considered. For example, Wals and Blaze Corcoran (2006) argued that an individual needs to explore four dimensions of perspectives when working for sustainability: cultural, generational, disciplinary, and national. Similarly, Seager, Selinger, and Wiek (2012) introduced a macro-ethical tool to guide individuals to consider six dimensions of sustainability: strong versus weak sustainability, local versus global concerns, humans versus nature, longevity versus adaptability, present versus future, and allocation versus distribution. While these descriptions are useful tools for highlighting the diversity of perspectives that should be considered in addressing wicked problems, they rest on the problematic assumption that it is possible to include all relevant thematic dimensions of perspectives (Bengtsson & Kronlid, 2016). We concluded that it was necessary to further develop the concept of perspectives. Specifically, we suggested that a more abstract description of *non-thematic* characteristics of perspectives and possible ways in which students could use and interact with their perspectives was needed.

Since the term *perspectives* can be used in many different ways, and since it can carry a large variety of meanings and connotations, we first needed to develop a general description of what we meant by perspectives. To develop such a general description, we used Franz Brentano's concept of "intentionality", according to which all mental processes are directed towards concrete objects. In other words, all thinking is thinking about physical objects (such as a house or a person) or mental objects (such as a memory or an idea) (Brentano, 1874/2009). For our framework, we defined perspectives as a kind of mental process that is directed towards the object of the perspective. In Paper I, we described perspectives as "a relationship between a person (an actor) and an object (…). In this relationship, the object is experienced and interpreted by the actor" (Lönngren, Svanström, Ingerman, & Holmberg, 2016, p. 344).

Foundations of qualitative content analysis

Qualitative content analysis has been developed from early work on, for example, bible interpretations, newspaper analysis, and dream analysis. In its current form, qualitative content analysis has been used since the middle of the 20th century (Mayring, 2000). According to Weber, the purpose of qualitative content analysis is to "classify large amounts of text into an efficient number of categories that represent similar meanings" (1990, in Hsieh & Shannon, 2005, p. 1279).

Cohen, Manion, and Morrison (2011) described three types of qualitative content analysis. They described *conventional* qualitative content analysis as an inductive approach in which coding categories are derived directly from text data. They further described *directed* qualitative content analysis as a deductive approach in which an existing theory is used to develop initial codes that are applied to the data. Finally, they described *summative* qualitative content analysis as an approach that combines deductive and inductive elements to obtain frequency counts of keywords or content mentioned in a text. In Study 1, I have used conventional qualitative content analysis, which Hsieh and Shannon suggested is useful for "concept development or model building" when limited prior theory exists (2005, p. 1281).

Methods

Empirical material

To obtain empirical material for this study, I conducted ten in-depth interviews with undergraduate engineering students. The purpose of these interviews was to obtain data that could illustrate how students use multiple perspectives on a wicked problem. For this purpose, we selected interview participants among undergraduate engineering students who were enrolled in a course on sustainable development in which the students were asked to discuss a wicked problem from different societal stakeholders' perspectives. This selection ensured that the interview participants had at least some experience in using multiple perspectives on wicked problems. However, since engineering students generally do not receive extensive training in using multiple perspectives or addressing wicked problems (and these students were no exception), using multiple perspectives on wicked problems still was a challenging task for the participants. To be able to use multiple perspectives on the problem, the participants needed active and continuous support from the interviewer throughout the interview. Therefore, we needed to depart from commonly used approaches in semi-structured qualitative interviews according to which the interviewer should try to avoid strongly influencing participants' accounts (see e.g. Brinkmann & Kvale, 2015). Instead, we followed Francis (1993), who suggested that, in phenomenographic interviews,

> even though the interviewer does not wish to prompt in such a way as to reduce fidelity, nevertheless, because the study is phenomenographic in aim, there must be serious attempts to "lead" the interviewee to comment on experience of learning particular content in particular contexts. That is to say that some pre-determined "leading experiences" and "leading prompts" are required to focus the interview appropriately for the aims of the study in question. (Francis, 1993, p. 70)

To "lead" the participants in our study to use multiple perspectives on a wicked problem, we used an explicitly interactive approach; I took a deliberately active part in the interviews as I repeatedly tried to stimulate the students to use different perspectives on the problem and to describe how they viewed the problem from those perspectives.

We organized the interviews around a written description of the wicked problem of *water-shortage in Jordan* and six idealized solution alternatives that different actors could propose to address the problem (p. 39). The purpose of the problem description was to establish a common focus for all interviews and to provide students with basic information about the problem. This basic information was intended to allow students to discuss concrete aspects of the problem. The purpose of the solution alternatives was to stimulate the use of diverse perspectives.

To develop the six solution alternatives, we used existing descriptions of thematic dimensions of perspectives in the literature (Seager, Selinger, & Wiek, 2012; Wals & Blaze Corcoran, 2006), which we complemented with brainstorming activities to identify additional dimensions of perspectives. In total, we identified over 100 dimensions. Using this collection of dimensions of perspectives, we attempted to incorporate as many and diverse perspectives as possible in the solution alternatives.

Each of the interviews proceeded through five phases. In the *first phase*, I asked the students to reflect on how they would want to address the problem. In this phase, I did not attempt to actively influence the students' reflections. We expected that at least some of the participants would directly dismiss several of the provided solution alternatives and quickly settle on one preferred approach (we were correct in this assumption). To avoid this kind of premature closure and support students in exploring multiple perspectives on the problem, we had developed a set of "prompt trails" (Francis, 1993, p. 71) that I could use in the *second phase* of the interviews to highlight alternative perspectives and challenge participants to develop a more complex understanding of the problem and the provided solution alternatives. For example, when students seemed to uncritically focus on one solution alternative, I asked "Do you see any problems with this solution alternative?", "If we choose this solution alternative, do you think there will be enough water?", or "Do you think that the local community would find that solution acceptable?" In some cases, I also used prompt trails in the form of photographic images and diagrams, such as an image of people protesting against the construction of a desalination plant and a diagram illustrating the effects of digging deeper wells. In the *third phase* of the interviews, I asked the students to disengage with the provided solution alternatives and describe any alternative approaches they would want to use to address the problem. In the *fourth phase*, I asked the students to assume two given professional roles on the problem, and in the *fifth phase*, I asked how they had experienced the problem itself and the discussion about the problem during the interview.

Before I conducted the interviews, I tested the problem description, solution alternatives and interview guide in two pilot interviews and made necessary adjustments. I also obtained written, informed consent from all participants, which allowed me to audio- and video-record all interviews. Each interview lasted approximately one hour. I later transcribed all interviews verbatim.

Data analysis and development of the theoretical framework

Table 5. Overview of the four stages of framework development in Study 1.

	Methods for analysis	**Foci of analyses**	**Versions of the framework**
Stage 1 (June-July 2012)	Inductive content analysis of interview transcripts 1, 2, 4, 10	Identify and categorize perspectives in terms of (a) different kinds of perspectives; (b) different "ways of taking perspectives" (analysis annotation on 17 June 2012); and (c) different ways of talking about perspectives (meta-reflection about perspectives)	*Versions 1-3* included (a) *thematic aspects*; (b) *non-thematic characteristics of perspectives*; and (c) a first attempt to include *perspective shifting* into the framework as an important perspective-related process
Stage 2 (August-September 2012)	Inductive content analysis of interview transcripts 3, 5, 6	As in Stage 1, but also challenge version 3 of the framework using three specific foci: (a) identify deviant cases, such as perspectives that cannot be described with the current version of the framework; (b) describe an "engineering perspective" (analysis annotation on 3 August 2012) to test the applicability of the framework to describe specific perspectives; and (c) focus on perspective shifting processes to identify different ways in which perspective shifting can occur and different reasons the students might have had to shift perspectives	*Version 4* included the concept of *meta-perspectives*
Stage 3 (October 2012 – May 2013)	Identify indicators for categorization and empirical examples for all categories in interview transcripts 1-6, 10	Discuss the fit of indicators and empirical examples with version 4 of the framework and the utility of the framework	*Version 5* focused on the most important characteristics of perspectives rather than trying to include all possible characteristics; *version 5* also included separate descriptions of *characteristics of perspectives* and *perspective processes*; The description of characteristics of perspectives in *versions 6 and 7* focused on non-thematic characteristics only
Stage 4 (May-July 2014)	Deductive content analysis of interview transcripts 4, 6, 9	Test the fit and utility of the framework in describing perspectives and perspective processes in the empirical material	Fit and utility of *version 7* confirmed

In Paper I, we described the process of developing the framework as "an analytical dialogue between a theoretical framework and empirical material consisting of transcripts of interviews with engineering students, but also drawing on the author group's considerable experience of [ESE], primarily in an engineering context" (Lönngren, Svanström, Ingerman, & Holmberg, 2016, p. 344). This analytic dialogue was an iterative process that included four kinds of activities: inductive content analysis of interview transcripts, discussion among authors, framework development and refinement, and deductive content analysis to test the final framework. The development proceeded through four stages, which are outlined in detail in Table 5.

Results and conclusions

Theoretical framework that describes perspectives and perspective processes

In response to RQs 1.3 and 1.4, we developed a theoretical *framework for perspectives*. In that framework, we described three non-thematic *characteristics of perspectives* that are relevant in the context of ESE. First, we described *perspective depth* as the level of detail and concretion in a discussion about a wicked problem. Second, we described *perspective coverage* as "the number of different thematic aspects that are considered simultaneously" and "the degree to which these aspects differ from each other" (Lönngren, Svanström, Ingerman, & Holmberg, 2016, p. 348). In this definition, thematic aspects are understood as specific facts of a wicked problem that are considered. These could for example be specific cultural aspects of a wicked problem or specific short- or long-term considerations in relation to the wicked problem. Third, we described *perspective complexity* as "the level of integration between different thematic aspects" (*ibid*., p. 348), i.e. the degree to which a perspective relates different thematic aspects of a wicked problem to each other rather than considering them in isolation.

In our framework, we further described six *perspective processes*, which we defined as "ways in which students adopt, change, reflect on, or otherwise interact with their perspectives" (*ibid*., p. 346). We described *perspective formation* as the process of establishing a first perspective on the wicked problem. Once students have established a perspective, they can change it either through perspective shifting processes or perspective integration processes. In *perspective shifting* processes, students simply shift their focus from one set of thematic aspects of the wicked problem to another set, for example from short- to long-term considerations in relation to the wicked problem. In *perspective integration* processes, students combine different perspectives on the wicked problem. For example, students could compare short- and long-term effects of a specific improvement measure and thus integrate short- and long-term considerations rather than shifting between them.

We used the term *meta-perspective* to describe students' focus on their own perspectives on the wicked problem rather than on the wicked problem itself. When students used such a meta-perspective, they could engage in three additional perspective processes: perspective perception, perspective reflection, and perspective evaluation. We described *perspective perception* as the process of becoming aware of one's own perspective. Such a process could for example be expressed by describing one's perspective in terms of which thematic aspects one holds in focal awareness. *Perspective reflection* processes involve explicit reflection on the nature of one's perspective, for example by describing one's perspective as "distanced and insentient" as one participant in Study 1 did. Finally, in *perspective evaluation* processes, students relate the nature of their own perspectives to external norms and values. For example, students could reflect on how other stakeholders would perceive their perspective in a specific context and then actively evaluate how appropriate their perspective is in that context.

Relating our framework to the literature on ESE competencies, we concluded that perspective integration, reflection, and evaluation may be particularly important processes in ESE.

STUDY 2

In what ways do engineering students approach a wicked problem? (RQ 1.5)

Research approach

We addressed RQ 1.5 in Study 2, in which we used a phenomenographic approach (Marton, 2015; Marton & Booth, 1997) to empirically study engineering students' approaches to a wicked problem. Phenomenography provides an inductive, empirical approach that is well-suited to study educational phenomena for which a solid theoretical basis is not yet established (Kinnunen & Simon, 2012; Svensson, 1997), which was the case in our research. Further, the phenomenographic research approach has been developed with the explicit aim to support educational development (Bowden, 2000; Collier-Reed & Ingerman, 2013; Marton, 2015; Marton & Booth, 1997), which suggests that it should be possible to use theoretical results from phenomenographic research to improve engineering educational practice in line with the aim of this thesis. In fact, Micari, Light, Calkins, and Streitwieser described phenomenography as "above all, a practical tool for improving education: The perspective it offers on differences in learning experience can enable educators to more deeply understand why and how their learners struggle, and how this struggle might be overcome" (2007, p. 463). Similarly, Case and Light argued that "the identification of different conceptions (…) makes phenomenography particularly well suited for the design of educational learning objectives, pedagogical strategies, assessments, and evaluations" (2011, p. 199).

Foundations of phenomenography

The phenomenographic research approach was first developed in the context of empirical studies that aimed to describe students' conceptions of, and approaches to, learning (see e.g. Marton, 1979; Marton & Säljö, 1976; Svensson, 1976). Since then, phenomenographic researchers have attempted to clarify the theoretical and methodological nature of phenomenographic research (e.g. Åkerlind, 2005; Harris, 2011; Marton & Pong, 2005; Pang, 2003; Richardson, 1999; Sin, 2010; Svensson, 1997). Today, phenomenography is widely used in educational research in some parts of the world, and it has also been identified as a promising and "emerging" research approach in EER (Case & Light, 2011).

Philosophically, phenomenography has some similarities with phenomenology (Svensson, 1997). For example, phenomenography focuses on subjective experiences rather than attempting to describe an external reality. However, these subjective experiences are viewed *in relation to* the external reality, since, according to Brentano's concept of intentionality, it is impossible to simply "experience"; it is only possible to experience specific (physical or mental) objects (Brentano, 1874/2009; see also p. 46). In phenomenography, the experienced objects are called "phenomena" and the relationships between individuals and phenomena are called "conceptions", "ways of experiencing", or "approaches" (Marton & Booth, 1997; Svensson, 2016).

An important characteristic of phenomenographic research is its focus on *variation*. In contrast to phenomenology, phenomenography does not aim to describe individual experiences as fully as possible. Rather, the aim of phenomenography is to map the variation in ways in which individuals can experience a phenomenon. This variation is described in terms of *educationally critical differences* between different ways of experiencing the structure and meaning of a specific phenomenon (Pang, 2003). The descriptions of educationally critical differences provide a link to educational practice since educators can use the descriptions to design educational interventions that help students develop more complex and complete ways of understanding the phenomenon.

Methodologically, phenomenography provides an inductive, empirical approach to data analysis. Since the aim is to identify different approaches towards a phenomenon, phenomenographic analysis is performed on the collective level rather than through an analysis of individual participants' experiences. For this purpose, all data is combined in a *pool of meaning* (Marton & Booth, 1997), typically in the form of interview transcripts that have been divided into shorter excerpts in which distinct conceptions of the phenomenon are expressed (although some researchers use complete interviews as a unit of analysis rather than shorter excerpts (Åkerlind, 2005)). In phenomenographic analysis, each of these excerpts is examined in relation to the entire set of excerpts in the pool of meaning rather than only in relation to other excerpts from the same interview (Åkerlind, 2005; Marton & Booth, 1997).

Phenomenographic analysis is typically a highly iterative process of developing and refining a set of categories. Each of these categories describes a distinct conception of the phenomenon in terms of its structure and meaning. Finally, the categories are related to each other in a phenomenographic *outcome space*. In the outcome space, the categories are organized hierarchically in terms of increasing levels of complexity of the conceptions that are described in the categories (Åkerlind, 2005; Marton & Booth, 1997).

According to Marton and Booth (1997), a phenomenographic outcome space must fulfill four quality criteria. First, the categories that form the outcome space need to be valid, i.e. they need to adequately represent the conceptions that were identified in the empirical material. Second, the categories need to be parsimonious, i.e. there may not be any redundant categories. Third, the categories need to be mutually exclusive, i.e. there may not be any overlap among categories so that no data excerpt could be ascribed to several categories. Finally, logical (often hierarchical) relationships among the categories need to be identified and the outcome space needs to be structured accordingly.

Empirical material and data analysis

As empirical data for our phenomenographic analysis, we used the interview data that was also used in Study 1 (as described on p. 46), which we analyzed in nine iterative rounds of categorization and description. In each round, I first created a pool of meaning from the interview transcripts. For this purpose, I printed the transcripts and cut them into smaller pieces so that distinct conceptions of the phenomenon were expressed in each piece. I then sorted the excerpts from the pool of meaning into collections of similar approaches to the phenomenon, developed descriptions of each of these collections, and developed descriptions of how these collections are related to each other. In discussions among all authors, we evaluated and challenged the validity, parsimony, mutual exclusiveness, and logical relatedness of the thus established categories (Marton & Booth, 1997). In addition to discussions among the authors, we engaged a local phenomenography research group in an evaluation of a preliminary outcome space.

In the iterative process of developing and evaluating the phenomenographic categories, we also developed more and more concrete descriptions of the phenomenon. This led us to reduce the scope of the pool of meaning to ensure that the analysis focused on a single phenomenon rather than several different phenomena. For example, we decided to exclude excerpts from two phases of the interviews. First, we excluded the initial discussion of the problem since we found that the participants had not yet established a clear relationship to the phenomenon. Second, we excluded a part of the interview in which the participants were asked to consider the problem from two given professional perspectives rather than describe their own understanding of the phenomenon. We also excluded excerpts from the remaining phases of the interviews if they were not clearly and directly related to the phenomenon (e.g. statements about secondary problems that arose when solution alternatives were applied to the original problem). I reprinted and recut the entire material twice during the analysis process to ensure that the pool of meaning matched our developing understanding of the phenomenon.

Eventually, additional iteration did not significantly change the descriptions of the categories or the outcome space and all authors were convinced that the outcome space fulfilled the criteria of validity, parsimony, mutual exclusiveness, and logical relatedness as described above. Finally, we confirmed the potential of the results to contribute to educational practice (Collier-Reed & Ingerman, 2013) in discussions with engineering educators (Study 3), and in a design-based research project (Study 4).

Results and conclusions

Phenomenographic outcome space for engineering student's approaches to a wicked problem

In response to RQ 1.5, we identified four qualitatively different approaches to the wicked problem of water-shortage in Jordan: *simplify and avoid*, *divide and control*, *isolate and succumb*, and *integrate and balance*. When using the *simplify and avoid* approach, students viewed the problem as something vaguely problematic and a solution as something that somehow would solve the problem completely and satisfy all stakeholders. The *divide and control* approach was more elaborate in that it acknowledged that the wicked problem had multiple parts that needed to be addressed with multiple solution approaches. However, the problem parts were considered to be largely independent of each other so that they could be treated in isolation. Multiple solution approaches were thus simply added together without integration. When using an *isolate and succumb* approach, students appreciated even more of the complexity of the problem. In this approach, they viewed the parts of the problem as interconnected in complex ways. However, they still attempted to address each problem part in isolation. As they realized that such an approach is inappropriate for wicked problems, students using this approach concluded that the problem cannot be solved and they gave up trying to find other ways to improve the situation. Both *divide and control* and *isolate and succumb* can be described as *taming* approaches (Conklin, 2005; DeFries & Nagendra, 2017; Roberts, 2000; see p. 22 in this thesis). *Integrate and balance* was the most complex approach that students in the study used. In this approach, both problem parts and solution approaches were viewed as a complex system in which all parts were connected in complex ways. A “solution” to a wicked problem was in this approach viewed as an iterative attempt to improve an unsatisfactory situation while all the time being vigilant of changing circumstances and unintended consequences. In an *integrate and balance* approach, taming is avoided. Figure 2 provides a graphic representation of the phenomenographic outcome space.

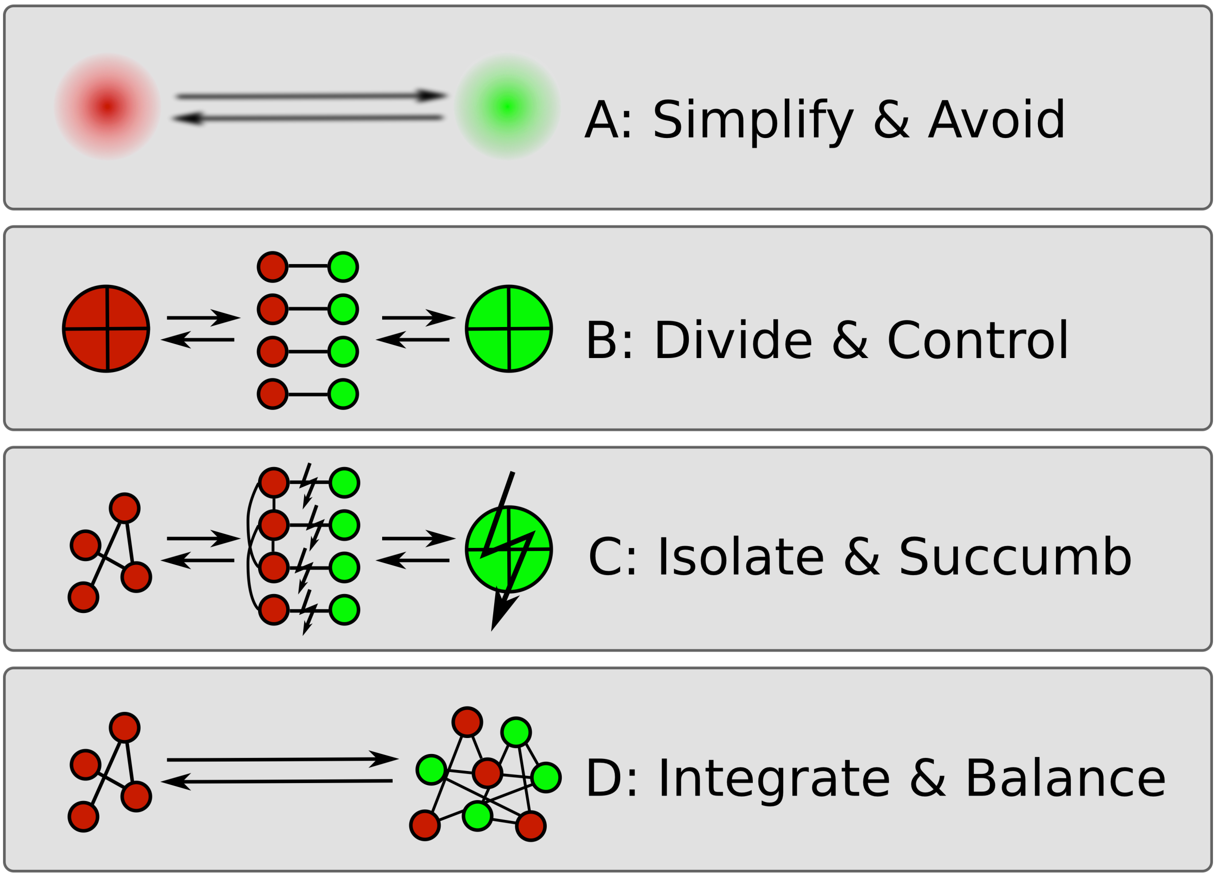

Figure 2. Graphic representation of four approaches to a wicked problem. Dark red items illustrate either the problem (approaches A and B) or parts of the problem (approaches B, C, and D). Correspondingly, light green items illustrate either what is perceived as a complete solution (approaches A and B) or improvement measures (approaches B, C, and D). Double-headed arrows are used to indicate that participants do not necessarily approach the wicked problem in a linear way by first attempting to define the problem and then identifying improvement measures.

SYNTHESIS OF STUDIES 1 & 2

Which of these approaches may be considered appropriate in the normative context of ESE? (RQ 1.6)

Methods

To answer RQ 1.6, we considered the results and conclusions from studies 1 and 2 in relation to each other. This was possible because the studies shared a common theoretical basis (the principle of intentionality) and utilized the same empirical material.

We also considered the results in relation to the normative context of ESE. In particular, we used descriptions of "key competencies" in ESE, since they can indicate what may be considered desirable goals of ESE. Based on a review of the literature on ESE competencies, Wiek, Withycombe, and Redman suggested that "the most critical check for the adequacy of the [ESE] competencies is the degree to which graduates can improve sustainability in the world" (2011, p. 214). This description is in line with the overall aim of this thesis, i.e. *to contribute to improving engineering education practice such that it could better prepare engineering students to contribute to sustainability*.

Results and conclusions

Integrative approach as most appropriate

In both studies, we identified an *integrative* approach as most appropriate for ESE. In Paper I, we argued that *perspective integration* is "crucial for designing resilient and sustainable solutions that take into account possible unintended consequences as well as different stakeholders' needs and interests" (Lönngren, Svanström, Ingerman, & Holmberg, 2016, p. 350). We further argued that learning to use perspectives with large coverage, depth, and complexity may support the use of perspective integration (*ibid*.; Lönngren, 2014). In Paper II, we identified *integrate and balance* as the most appropriate approach for addressing wicked problems since it "is most in line with the characteristics of [wicked problems]" and "is based on the most complex understanding of the problem, while also enabling students to see constructive ways of dealing with the problem and thus to 'improve sustainability in the world' (Wiek, Withycombe, & Redman, 2011, p. 214)" (Lönngren, Ingerman, & Svanström, 2017, p. 823).

We concluded that an integrative approach to wicked problems (either described as *perspective integration* or as an *integrate and balance* approach to wicked problems) allows students to avoid dividing wicked problems into isolated parts. Rather than merely shifting between different perspectives and addressing problem parts in isolation, perspectives are integrated and relationships among problem parts and improvement measures are actively and iteratively constructed and reconstructed as the understanding of the specific wicked problem develops in the process of addressing it.

Observation that students in the study are able to use an integrative approach, but that they may need instructional support to do so

Marton and Pong suggested that it is common in phenomenographic research to find that "the same individual can express different conceptions of the same phenomenon when asked different questions, or when in different situations" (2005, p. 346). In line with this suggestion, we found that the same students could use an *integrate and balance* approach in some parts of the interview and *divide and control* or *isolate and succumb* approaches in other parts of the interview. Similarly, in Paper I, we observed that one of the interviewees first performed a series of perspective shifts and then integrated several of these perspectives. From these results, we concluded that some of the students in the interviews for Studies 1 and 2 were *able* to use integrative approaches to address the wicked problem, but that they only did so under certain circumstances. This conclusion raises questions about what kinds of instructional support engineering students may need in order to be able to use integrative approaches to wicked problems, and what kinds of instructional support they may need to learn to *consistently* use integrative approaches when addressing wicked problems. An suggested that "there is no question about the importance of scaffolding [a type of instructional support, see p. 78 in this thesis] on ill-structured problem solving", but that "there is a need for additional research on the effectiveness of different types of scaffolds in supporting ill-structured problem solving" since previous research has found that not all kinds of scaffolds "have a positive effect on problem solving" (2010, pp. 724f). Study 4 aimed to explore some of these questions.

What are salient aspects of integratively addressing wicked problems for engineering students? (RQ 1.7)

Methods

To identify salient aspects of integratively addressing wicked problems, I used three sources of information. First, and most importantly, I used the phenomenographic description of the *integrate and balance* approach developed in Study 2 (p. 53). Treating the description itself as empirical material, I summarized and consolidated the content of the description until I had identified a set of salient aspects of integratively addressing wicked problems. Second, I used the literature on wicked problems (p. 19) to ensure that the identified aspects are salient *specifically for wicked problems*. Third, I used literature that indicates what aspects of wicked problems and what processes of addressing wicked problems may be particularly challenging for engineering students (p. 32) to ensure that the identified aspects are salient *specifically for engineering students*. During Study 4, we further refined the set of aspects in response to challenges that emerged when we used the aspects in an educational intervention, and in response to student and expert feedback (p. 77).

Results and conclusions

I identified ten salient aspects that need to be identified and reflected upon when integratively addressing wicked problems (shorthand descriptions in parentheses):

I. different problem parts that together make up the overall problem (problem parts);

II. different improvement measures that could be used to address the overall problem and/or individual problem parts (improvement measures);

III. interaction between problem parts, such as feedback mechanisms, symbiosis, or conflict, through which one problem part could alleviate or aggravate other problem parts (problem part interaction);

IV. interaction between improvement measures, such as feedback mechanisms, symbiosis or conflict, through which one improvement measure could facilitate or impede the successful implementation of other improvement measures (improvement measure interaction);

V. unintended, secondary problems that could be caused by improvement measures (secondary problems);

VI. stakeholders and their interests in relation to the overall problem, individual problem parts, and/or improvement measures (stakeholders);

VII. spheres of influence of different actors who could be involved in addressing the overall problem and/or individual problem parts (spheres of influence);

VIII. lack of accessible information, i.e. information that is not currently available but could be obtained through directed research and information-gathering activities (lack of information);

IX. the importance of incomplete control and predictability for the outcome of improvement measures (uncertainty);

X. the importance and influence of the local problem context for the outcome of different improvement measures (local context).

In the group of authors for Papers IV and V, we concluded that these ten aspects could provide a theoretical *framework for integratively addressing wicked problems*. We further concluded that this framework could be used as a basis for developing teaching and assessment activities for developing engineering students' ability to integratively address wicked problems.

PART 2

In this chapter, I describe the research in Part 2, which aimed to answer the question of *how the ability to address wicked problems can be taught and assessed in engineering education*. The chapter consists of two sections that describe Study 3 and Study 4 respectively. In the section on Study 3, I address RQs 2.1-2.3. In the section on Study 4, I address RQs 2.4-2.7.

STUDY 3

> The findings of traditional social science are of little or no use to members of organizations or practitioners (…); there is a division between academic research and the everyday practice that action research seeks to address. (Reason, 2006, p. 188)

The aim of Study 3 was to identify ways in which the theoretical results from Part 1 could be rendered more useful for engineering education practice. When we had developed the phenomenographic outcome space for engineering students' approaches to wicked problems in Study 2, we felt that the outcome space could be very valuable for practice. However, we also felt that there was a missing link between the theoretical results and practical application – the results were highly abstract and not readily applicable to educational practice. We decided to try to develop this missing link in Part 2 of the research underlying this thesis (Studies 3 and 4). In Study 3, we used a *pragmatic action research* approach in which we engaged educators in discussions and workshops to, first, identify problems in engineering education practice that could be addressed with the help of the theoretical results from Part 1, and second, on the basis of those theoretical results, develop practical tools to address some of the problems that had been identified.

Methodology for Study 3

Foundations of pragmatic action research

Action research has been described as an emerging methodology in EER (Case & Light, 2011) that explicitly focuses on practitioner engagement and acknowledges contribution to practice as one of the core outcomes of research. Action research has also been described as "a large family" of research approaches "rather than a particular research methodology" (Noffke, 1997, p. 306). Noffke described three dimensions of action research: The *professional* dimension of action research aims to "create stronger connections between university researchers and educational practitioners so as to bridge the 'theory-practice gap'" (*ibid.*, p. 324); it does not, however, aim to transform the "fundamental relationships" (*ibid.*, p. 306) between research and practice communities. The *personal* dimension of action research aims to support practitioners' individual development, for example by helping practitioners to develop a greater awareness of their own practices and helping them to explore how those practices align with their personal values. Finally, the *political* dimension of action research "seeks not additions to a knowledge base for teaching but a transformation of educational theory and practice toward emancipatory ends", for example by engaging practitioners in critical discussions about power relations (*ibid.*, p. 324). Noffke asserted that "all versions of action research have professional, personal, and political dimensions, but in the various efforts some aspects are more fully articulated, emphasizing different things in different ways" (*ibid.*, p. 322). Similarly, Reason (2006, p. 198) suggested that

> action research is full of choices: It is not possible, either theoretically or practically, to engage in an inquiry that addresses all dimensions [of action research[5]] fully and completely; rather, there will always be choices about what is important to attend to at any particular moment.

Reason further argued that the kind of action research that is used in a specific project depends on the specific aims of the project. These aims can vary widely and can include

- developing "immediate practical outcome[s]";
- helping to "articulate voices that are not being heard";
- "finding ways to open ourselves to different sorts of realities, or finding different ways of telling stories";
- "in-depth exploration into values, into what purposes are worthwhile pursuing, and into what issues most deserve our attention"; and/or
- "creating tentative beginnings of inquiry under very difficult circumstances" (2006, pp. 189-199).

The aim of Study 3 was most closely aligned with the first of the above-listed purposes: We wanted to develop "immediate practical outcomes" (Reason 2006, p. 189) in the form of practical tools that could help educators to address problems in engineering education practice as experienced by educators themselves. This aim was also most closely aligned with the *professional* dimension of action research in Noffke's (1997) description: We wanted to connect the theoretical results from Part 1 of the research with engineering education practice in order to "bridge the 'theory-practice gap'" (*ibid.*, p. 324). While the personal and the political dimensions of action research were present in the form of the overarching aim of the research to contribute to sustainability, they were not the main focus of Study 3.

Thus, the approach taken in Study 3 can be described as *pragmatic* action research, which aims to develop *useful practical knowledge* through "cooperation between all concerned parties", which in turn requires "finding and constructing a common ground" and thus "is characterized by consensus and conflict avoidance, and works hand in hand with existing structures of dominance in society" (Johansson & Lindhult, 2008, p. 100).

Johansson and Lindhult contrast pragmatic action research with *critical* action research, which "is focused upon the emancipation of underprivileged groups, and carries with it a strong ideological base in opposition to existing structures of society" (*ibid.*, p .99). Thus, critical action research aims to "de-establish established truths" (*ibid.*, p. 104) and "destabilize

---

[5] Note that Reason in this quote did not refer to Noffke's three dimensions of action research. Rather, Reason referred to what he described as "four characteristic dimensions" of action research "that present a broad range of criteria (…) against which quality [action] research might be judged". His criteria include that (1) "worthwhile practical purposes" are pursued, (2) the research process is democratic and participatory, (3) research and practice are integrated, and (4) the research design is iteratively and flexibly adapted to changing circumstances and new questions that emerge during the research process (2006, p. 187).

the dominant views" rather than "to contribute to their propagation" as may be the case in pragmatic action research (*ibid.*, p .108). The critical dimension of action research was not a primary focus of Study 3.

Study overview

While there appears to be a large variation in action research approaches and no consensus on what should be considered the "ideal form" of action research (Greenwood, 2007), there appears to be general agreement that action research should be conducted in a democratic and participative manner (Reason & Bradbury, 2001). Thus, participation is an important concern in action research methodology:

> At a methodological level participation is important because self-evidently one cannot study and improve practice without deep involvement of those engaged in that practice, for the necessary perspective and information [on practice] is simply not available [otherwise]. (Reason, 2006, p. 189)

In Study 3, we strove to engage engineering educators in order to be able to, first, identify barriers for engineering educators to teach their students to address wicked problems, and second, try to address some of these barriers in ways that educators would find useful, acceptable, and realistic. Study 3 proceeded through three stages in which we used different strategies to ensure participation. The stages built on each other in the sense that the results from one stage provided input for the next stage. Each of the stages addressed one research question (RQs 2.1-2.3 in Figure 1)[6]. The stages were as follows:

Stage 1. Individual interviews with engineering educators to identify barriers for teaching students to integratively address wicked problems, and to identify ways in which the results from Studies 1 and 2 could be used to address some of the identified barriers (RQ 2.1);

Stage 2. Workshop with educators to formulate intended learning outcomes for students' ability to integratively address wicked problems (RQ 2.2);

Stage 3. Workshops with engineering educators to develop assessment methods for the intended learning outcomes that were developed in Stage 2 (RQ 2.3).

To engage participants who could contribute a broad range of perspective, the research in the three stages was conducted in different contexts. The research in Stage 1 was conducted at a large technical University in the United States. The research in Stage 2 was conducted in Gothenburg, Sweden, involving participants from both Chalmers and Gothenburg University. The research in Stage 3 was conducted partly in Gothenburg and partly in New Orleans, United States.

---

[6] In Paper III, we described four stages of the study because we included Study 2 as one of the stages of Study 3. The three stages described here are the same as Stages 2-4 in Paper III.

What do engineering educators view as important barriers for teaching their students to integratively address wicked problems? (RQ 2.1)

Methods

Empirical material

We addressed RQ 2.1 in the first stage of Study 3. The aim of this stage was to identify problems in engineering education practice that could be addressed with the help of the theoretical results from Part 1 of the research in this thesis. For this purpose, I conducted two rounds of semi-structured interviews (Brinkmann & Kvale, 2015) with engineering educators at a large technical university in the United States. Between the two rounds of interviews, I also observed some of these educators' teaching practice related to ESE in engineering education. To establish contact with potential participants, I asked a local researcher for help. The local researcher identified six educators who were engaged in ESE at the university and who had a great interest in educational questions. All six educators agreed to participate in the study, and I obtained written, informed consent from all participants.

The purpose of the *first round of interviews* was to establish a first relationship with the research participants, to get an idea of their understanding, experiences, and current practices of engineering ESE, and to get an idea of their understanding and experience of EER in relation to their own teaching practice. During the interviews, I first asked questions about the participants' disciplinary background. I then asked them to describe their current ESE teaching in terms of topics covered, student ages, class sizes, pedagogical approaches used, learning outcomes, assessment approaches, and any deliberate strategies they might be using to continually improve their teaching practice. I also asked them about their current research and whether they had any experience of, or particular interests in, EER. I took written notes during these interviews.

After the first round of interviews, I invited three of the educators to participate in the remainder of this stage of the study. I chose these educators, first, because I expected that the theoretical results from Part 1 of the research underlying this thesis could be relevant to their practice (based on their descriptions of their ESE practice), and second, because they had expressed an interest in further participation.

I conducted *classroom observations* (Wragg, 2013) of some of the three selected participants' ESE teaching. For one of the participants, I observed a one-day field trip. For the other two participants, I observed a two-hour lecture or seminar, respectively. During these observations, I took written notes related to what the participants did during the lecture/seminar, how they approached the ESE content, how they interacted with the students, and what the students did. Through these observations, I was able to directly experience some of the participants' teaching, and thus develop a deeper understanding of the participants' educational context and their understanding of ESE.

Finally, I conducted a *second round of interviews* with the same three participants. As a preparation for this round of interviews, I had asked the participants to read, and reflect on, a four-page synopsis of the theoretical results from Study 2. The synopsis contained a short introduction to wicked problems in engineering education and a relatively detailed description of the phenomenographic outcome space developed in Study 2. The synopsis also contained the following description of the purpose of the interviews and some questions that I was hoping to discuss during the second round of interviews:

> During the interview, I would like to discuss with you what these results might mean to you as an engineering educator. How do you understand the implications of the research in relation to your own teaching? Do the results trigger any specific reflections related to your educational practices? Do they inspire you to make changes to your teaching? Or do you have any ideas about how the results could be further elaborated on to be more useful for your work as an engineering educator?

During the interviews, I asked the participants to describe how they had understood the research results presented in the synopsis and I addressed any questions they had about the content of the synopsis. I then asked them to describe whether they thought that the research was relevant to their ESE practice and whether any aspects of the results triggered reflections about, or ideas for, their own teaching. I also included specific questions about how the results related to concrete aspects of their own ESE teaching (based on their descriptions during the first round of interviews and/or my observations of their teaching). Finally, I asked the participants whether they had any ideas about how the results could be used to address problems in engineering education practice. I audio-recorded these interviews.

Data analysis

Since the focus of this stage of Study 3 was on *engineering educators' experiences* of barriers for teaching students to integratively address wicked problems, I primarily focused on analyzing the data from the *interviews*. I did not analyze the data from the classroom observations in detail; rather, I used that data to better understand the themes that emerged from the analysis of the interview data.

To analyze the data from the first round of interviews, I first summarized the written notes from each of the interviews. I did this directly after the interviews while I still remembered the interviews clearly. In these summaries, I organized the information from each interview in terms of information about the participants' teaching and information about the EER experience. Second, I produced an overall summary of all interviews in the first round, focusing on similarities and differences between the participants' descriptions of their ESE teaching and any challenges in their teaching practice that they had mentioned during the interviews. Once I had conducted the classroom observation (a few days to weeks after the first round of interviews), I compared the individual and overall summaries from the first round of interviews with the written notes from the classroom observations to better understand the challenges in practice that had been mentioned. Finally, I used the individual and overall summaries from the first round of interviews to formulate questions for the second round of interviews that could provide deeper insights about specific challenges in each of the participants' teaching practice.

To analyze the data from the second round of interviews, I first summarized the audio recordings from each of the interviews. In these summaries, I organized the information from each interview in terms of information about challenges in teaching practice and information about participants' thoughts on how the theoretical results from Part 1 of the research in this thesis could be useful for their teaching practice. Second, I produced an overall summary of all interviews in the second round, focusing on similarities and differences between challenges in teaching practice and ideas for using the theoretical results from Part 1 of this thesis for addressing them. Third, I compared the challenges that emerged from the summary of the interviews in the first round with those that emerged from the summary of the interviews in the second round. I also compared the descriptions of those challenges with the ideas for addressing them that had been mentioned, and I attempted to interpret the ideas for addressing challenges in light of the participants' descriptions of their EER experience. Finally, I tried to identify one or a few challenges that the participants experienced as barriers for teaching their students to integratively address wicked problems and which they indicated could potentially be addressed with the help of the theoretical results from Part 1.

Results and conclusions

In the analysis of the first round of interviews, it became clear that the participants had varying levels of experience with EER. However, they all expressed that EER should address challenges in engineering education practice. One of the participants described this very clearly: "EER needs to relate to realistic contexts and conditions" rather than assume ideal circumstances such as unlimited resources or ideal teaching facilities; EER should "address the *concrete challenges of everyday educational practice*". As one such concrete challenge, the educators identified *assessment of complex ESE competencies*.

In the second round of interviews, the participants reported that they experienced a lack of robust assessment tools for assessing students' approaches to wicked problems and that this lack of assessment tools created an important barrier for them to teach their students how to integratively address wicked problems. Indeed, as we later reviewed the literature on assessment of complex ESE competencies for Stages 2 and 3 of Study 3, we did not find any reports of carefully developed and evaluated assessment instruments that could be used to assess engineering students' ability to address wicked problems. We concluded that it would be a valuable contribution to practice to develop such assessment tools and a valuable contribution to theory to understand how the tools could function in concrete engineering education contexts.

What learning outcomes do educators view as important for students' ability to integratively address wicked problems? (RQ 2.2)

Theoretical background

The aim of the second stage of Study 3 was to identify *learning outcomes* that educators viewed as important for engineering students' ability to integratively address wicked problems. The research in this stage was conducted in Gothenburg, Sweden, and most participants were employed at either Chalmers or Gothenburg University. To facilitate practitioner participation during this second stage of the study (as well as the third stage), we decided to frame the research activities with the help of theories and concepts that would be familiar to the participants. In particular, we decided to use the framework of *constructive alignment*, which is used as a framework for curriculum development at Chalmers and widely communicated in seminars and pedagogical development courses at the university.

Constructive alignment has been described as

> an outcomes-based approach to teaching in which the learning outcomes that students are intended to achieve are defined before teaching takes place. Teaching and assessment methods are then designed to best achieve those outcomes and to assess the standard at which they have been achieved. (Biggs, 2014, p. 5)

Constructive alignment combines a constructivist view of knowledge with a behaviorist approach to curriculum development that is known as “instructional alignment”. In describing the meaning of *constructivism* in constructive alignment, Biggs referred to Tyler (1949) and Shuell (1986) and suggested that those authors described knowledge as “constructed through the activities of the learner” (Biggs, 2014, p. 9). In describing the meaning of *instructional alignment*, Biggs’ referred to Cohen who had described instructional alignment as “the extent to which stimulus conditions match among three instructional components: intended outcomes, instructional processes, and instructional assessment” (Cohen, 1987, p. 16). In combining constructivism and instructional alignment, Biggs argued that “the key to good teaching then is to get the learner to engage those activities (*sic*) that are most appropriate” to the specific learning outcomes that the learners are expected to reach (2014, p. 9).

Constructive alignment has been criticized for its use of constructivist theories of learning (especially in the context of science education), its use of behaviorist approaches to assessment, and the combination of constructivism and behaviorism (Jervis & Jervis, 2005). It has also been criticized for its narrow focus on educational outcomes, which may favor a reductionist approach to teaching with an undue focus on easily measurable, lower level competencies. However, constructive alignment has also been reported to provide an effective tool for achieving ”high quality learning outcomes and student satisfaction” (Biggs, 2014). Biggs (2014) also suggested that constructive alignment can be used to design teaching activities for lower level and complex competencies alike. Constructive alignment is today widely used in higher education curriculum development (Biggs, 2014; Jervis & Jervis, 2005), not only at Chalmers.

Methods

Empirical material

In response to RQ 2.2, my co-author and I organized a workshop with educators to develop intended learning outcomes for students’ ability to integratively address wicked problems. The workshop was part of a seminar at the Center for Environment and Sustainability in Gothenburg, Sweden. The seminar was organized in collaboration between Chalmers and Gothenburg University as part of a seminar series with recurring, bi-annual seminars on ESE. We chose this seminar as a context for our workshop because the seminar series typically attracts educators with a strong interest and engagement in ESE in higher education. 27 educators from Chalmers (n=10) and Gothenburg University (n=17) participated in our workshop; most of them had extensive experience in working with ESE in their own teaching practice. Participants from Chalmers had experience in working with ESE in engineering education; participants from Gothenburg University had experience in working with ESE in other disciplines in higher education, such as economy, environmental sciences, cultural sciences, and law.

The workshop lasted approximately 90 minutes and included group activities and plenum discussions as described in detail in Table 6. My role during the workshop was to introduce relevant theory, facilitate discussions, and observe group work; my co-author participated in the work of one of the groups. At the beginning of the workshop, I obtained oral, informed consent from all workshop participants, which allowed me to audio-record all plenum activities during the workshop. Throughout the workshop, the participants appeared to be actively and positively engaged in the activities.

Table 6. Overview of workshop activities in Stage 2 of Study 3.

Activity	**Duration**	**My role**
Theoretical introduction about • results from Study 2 and Stage 1 of Study 3, and • constructive alignment as a curriculum planning approach (Biggs, 1996; Biggs, 2014)	15 minutes	Present
Group work to identify concrete learning outcomes for students' ability to integratively address wicked problems (based on the results from Study 2) and summarize them on A2 paper sheets	50 minutes	Observe group work and answer potential questions
Plenum activity to present and discuss the groups' suggestions	25 minutes	Facilitate discussion
Summary and invitation for further collaboration for developing and evaluating assessment approaches in practice	2 minutes	Present

Data analysis

After the workshop, I summarized and consolidated the groups' suggestions for learning outcomes. For this purpose, I first developed a list of all suggested learning outcomes, irrespective of which group had suggested them. I then identified and removed duplicates. To organize the learning outcomes, I then developed a matrix of different types of learning outcomes. Along one of the axes, I categorized learning outcomes in terms of whether they focused on (a) a general understanding of sustainability, (b) an understanding of the nature of wicked problems, or (c) possible courses of action for addressing wicked problems. Along the second axis, I categorized the learning outcomes according to what kinds of learning outcomes they represent, based on existing categorizations of learning outcomes by the Swedish National Agency for Higher Education (Högskoleverket, 2010; Högskoleverket, 2012; these descriptions are commonly used in Swedish course and program descriptions in higher education), the UNESCO (UNESCO, 2010; these descriptions focus specifically on types of outcomes that are relevant in ESE), and Bloom's taxonomy (Bloom *et.al*., 1956; Krathwohl, 2002; this taxonomy is widely known and used by educators in both Sweden and the United States). The resulting category system for the second axis included five items: (1) knowledge and comprehension, (2) skills (application of knowledge), (3) advanced thinking processes (analysis, synthesis, evaluation), (4) attitudes and values (approaches, valuation), and (5) action competence. Where necessary, I combined or separated learning outcomes that were developed during the workshop to fit the structure of the matrix. Once I had organized the learning outcomes in this matrix, I homogenized the descriptions of the learning outcomes using active verbs that describe observable student actions (Biggs, 2014). Finally, my co-author and I refined the categorization and formulations in the matrix together.

Results and conclusions

The analysis and consolidation of the groups' suggestions resulted in a matrix of 22 intended learning outcomes (Table 7). We concluded that the ability to integratively address wicked problems covers a large number of diverse learning outcomes.

Table 7. Matrix of 22 intended learning outcomes for students' ability to integratively address wicked problems (here called wicked sustainability problems, WSPs)

Type of intended learning outcome	**(a) Sustainable Development (SD)**	**(b) WSPs – Problem situations**	**(c) WSPs – Courses of action**
(1) Knowledge and comprehension	*Describe* different perspectives on what could be seen as SD and what a sustainable society could be. *Describe* how different societal actors may use the SD concept in various contexts. *Describe* what it means that SD is a political concept.	*Describe* the general characteristics of WSPs, particularly in contrast to tame problems.	*Explain* why it is not possible to find "absolutely correct" solutions to WSPs.
(2) Skills (application of knowledge)	*Utilize* the SD concept in discussions about a WSP in accordance with how it is commonly used in the political and scientific context of SD (as opposed to how it is used in e.g. marketing).	Independently *identify* a WSP in the context of one's future profession and *describe* why it is a WSP. *Identify* relevant aspects of a WSP and *describe* how they are interrelated.	With reference to the general characteristics of WSPs, *describe* how different societal actors attempt to deal with a current WSP.
(3) Advanced thinking processes (analysis, synthesis, evaluation)	Critically *examine* and *assess* alternative descriptions of what could be seen as SD and what a sustainable society could be.	Critically *examine* and *assess* alternative descriptions of a WSP. Independently *identify* relevant knowledge that would contribute to a holistic understanding of a WSP, especially when one does not have substantial prior knowledge about the situation.	*Use* duly substantiated social, ecological, economical, cultural, political, and technical perspectives to *suggest*, *discuss*, and *assess* alternative courses of action for a WSP.

Table 7 continued. Matrix of 22 intended learning outcomes for students' ability to integratively address wicked problems (here called wicked sustainability problems, WSPs)

Type of intended learning outcome	(a) Sustainable Development (SD)	(b) WSPs – Problem situations	(c) WSPs – Courses of action
(4) Attitudes and values (approaches, valuation)	*Explain* how different sets of values among societal actors contribute to the diversity of descriptions of what could be seen as SD and what a sustainable society could be. *Identify* relevant personal values and describe their influence on one's own understanding of what could be seen as SD and what a sustainable society could be. Deliberately and transparently *apply* different sets of values to *develop* different descriptions of what could be seen as SD and what a sustainable society could be.	*Explain* how different sets of values among societal actors contribute to the diversity of descriptions of, and preference for certain courses of action for, a WSP. *Identify* relevant personal values and describe their influence on one's own understanding of, and preferences for certain courses of action for, a WSP. Deliberately and transparently *apply* different sets of values to *develop* different descriptions of, and assess possible courses of action for, a WSP. *Demonstrate* an open attitude towards, and *elicit support* from, different knowledge domains that may be relevant to describing and addressing a WSP in a holistic manner.	
(5) Action competence	*Demonstrate* initiative, perseverance, and a sense of responsibility for SD.	*Demonstrate* initiative, perseverance, and a sense of responsibility for addressing a WSP, despite high levels of uncertainty, lack of information and knowledge about the situation, the ambiguous and contested nature of the SD concept, and the need to work across e.g. disciplinary and national borders.	

How could some of these learning outcomes be assessed? (RQ 2.3)

Theoretical background

Assessment in higher education

The aim of the third stage of Study 3 was to develop a few ideas for how (some of) the learning outcomes developed in the second stage of the study could be assessed in engineering education. Thus, the focus of this stage was on *assessment*. Educational assessment has been defined as

> (1) The planned process of gathering and synthesizing information relevant to the purposes of (a) discovering and documenting students' strengths and weaknesses, (b) planning and enhancing instruction, or (c) evaluating progress and making decisions about students; (2) the process, instrument or method used to gather the information. (Cizek, 1997, p. 10)

A wide variety of assessment approaches is used in the context of higher education. An important distinction is between formative and summative assessment. The main purpose of *formative* assessment is to provide feedback that can guide students' learning (Panadero & Jonsson, 2013). *Summative* assessment, on the other hand, is primarily used to evaluate students' learning at the end of a course or study program, for example by assigning grades. Previous research has found that well-designed, timely feedback in formative assessment can support students' learning, motivation, and awareness of their learning (Gibbs & Simpson, 2005; Hattie, 1987; Hattie & Timperley, 2007; Pereira, Flores, & Niklasson, 2016).

The literature on assessment in higher education also often distinguishes between traditional and alternative assessment approaches. *Traditional* assessment approaches include, for example, multiple-choice exams and written exams with sets of mathematical or story problems. The term *alternative* assessment is often used to describe any type of assessment that provides "alternatives to traditional testing" (Baartman, Bastiaens, Kirschner, & van der Vleuten, 2007, p. 117). Therefore, alternative assessment can refer to, for example, performance assessment (as opposed to assessment of factual knowledge, see p. 78 in this thesis), self- and peer-assessment (as opposed to teacher assessment), and/or assessment using open-ended essay questions (as opposed to e.g. multiple-choice exams).

In the literature on educational assessment, there seems to be a strong consensus that assessment can influence student learning (Baartman, Bastiaens, Kirschner, & van der Vleuten, 2007; Gibbs & Simpson, 2005; Scouller, 1998; Tian, 2007): "[A]ssessment has been found to shape how much, how (their approach), and what (the content) students learn. It seems that most students will learn the forms of knowledge and develop the cognitive abilities that they are asked to demonstrate" (Scouller, 1998, p. 454). For example, multiple-choice exams have been found to promote surface approaches to learning, while essay exams may promote deep learning (Pereira, Flores, & Niklasson, 2016; Scouller, 1998).

In the context of engineering education, Gibson suggested, "there is a real and widespread concern about assessment amongst engineering educators across Europe" (2002, p. 465). This concern is partly due to the general move towards more learner-centered and competence-based teaching approaches, both in higher education in general (Baartman, Bastiaens, Kirschner, & van der Vleuten, 2007), and in engineering education in particular (Gibson, 2002). As teaching approaches change, assessment approaches need to be aligned with those teaching approaches (Baartman, Bastiaens, Kirschner, & van der Vleuten, 2007; Biggs, 2014; Gibson, 2002). Adequately assessing complex competencies may require combining several different assessment approaches, including both traditional and alternative approaches (Baartman, Bastiaens, Kirschner, & van der Vleuten, 2007; Pereira, Flores, & Niklasson, 2016).

Assessment related to wicked problems

Stage 3 of Study 3 focused on assessing engineering students' ability to integratively address wicked problems. King and Kitchener (1994) described four general requirements for assessing students' ability to address ill-structured problems, such as wicked problems. They suggested that, first, assessment of students' ability to address ill-structured problems requires actually using ill-structured problems in the assessment. The reason for this is that addressing ill-structured problems requires different skills than solving well-structured problems (Jonassen, 2000; Kitchener, 1983; Schraw, Dunkle, & Bendixen, 1995; see also p. 36 in this thesis). Second, assessment approaches that require students to provide definite "solutions" to problems are inappropriate since such approaches would be in conflict with the nature of ill-structured problems. In fact, providing such assessment would require students to treat ill-structured problems as if they were well-structured – which counteracts the purpose of supporting students to develop their ability to address *ill*-structured problems in a holistic and integrative manner. King and Kitchener suggested the use of interview or essay examinations. Third, it is important to assess the process of how students arrived at suggesting a certain course of action for addressing an ill-structured problem. Rather than merely asking students to describe how they would address the problem, the educator should also include questions about how the students arrived at their conclusions. Finally, assessment needs to be tailored to the specific group of students that the educator is working with. For example, engineering students may not have had extensive training in writing "social science-type essays" (Gibbs & Simpson, 2005, p. 23). Thus, they may need discipline-specific instructional support in essay-based assessment.

Methods

Empirical material

To address RQ 2.3, we held two workshops that focused on developing assessment activities for (some of) the learning outcomes identified in response to RQ 2.2 (Table 7). We held the first of these workshops in January 2015 during the annual pedagogical conference for teaching and learning at Chalmers. The annual conference typically attracts engineering educators from all departments at the university who are interested in learning about EER results and/or improving their educational practice. Apart from my co-author and me, 5 engineering educators participated in the workshop. One of those five participants had also attended the workshop in Stage 2 of Study 3.

The workshop lasted approximately 60 minutes. Similar to the workshop in Stage 2, it included group activities and plenum discussions, as described in Table 8. Also similar to the workshop in Stage 2, my role during the workshop was to introduce relevant theory, facilitate discussions, and observe group work, while my co-author participated in the work of one of the groups. At the beginning of the workshop, I obtained oral, informed consent from all workshop participants and started the audio-recording device. I also took written notes throughout the workshop. All participants appeared to be actively and positively engaged in the workshop activities.

Table 8. Overview of workshop activities in the first workshop in Stage 3 of Study 3.

Activity	Duration	My role
Theoretical introduction about • results from Study 2 and Stages 1 and 2 of Study 3, and • constructive alignment as a curriculum planning approach (Biggs, 1996; Biggs, 2014)	10 minutes	Present
Group work to develop assessment activities for one or several of the learning outcomes that the participants thought could be applicable to their own teaching practice	30 minutes	Observe group work and answer potential questions
Plenum activity to present and discuss the groups' suggestions	25 minutes	Facilitate discussion

I held the second workshop in Stage 3 of Study 3 in June 2016 during the American Society for Engineering Education Annual Conference and Exhibition in New Orleans. This conference typically attracts engineering educators from all fields of engineering, mostly from the United States. This workshop lasted approximately two hours. One engineering educator from an engineering college in United States participated. The participant had read Paper III prior to the workshop. The activities during this workshop are summarized in Table 9; I obtained oral, informed consent from the participant to audio-record the discussion. I also took written notes throughout the workshop.

Table 9. Overview of workshop activities in the second workshop in Stage 3 of Study 3.

Activity	Duration	My role
Theoretical introduction about • results from Study 2, Stages 1 and 2 of Study 3, and the first workshop in Stage 3 of Study 3, and • constructive alignment as a curriculum planning approach (Biggs, 1996; Biggs, 2014)	15 minutes	Present
Discuss the participant's ideas for assessment and try to develop assessment ideas that could be directly applicable to the participant's own ESE practice; this included the following activities: 1. the participant described his ESE practice, 2. we discussed any questions or comments he had about the material presented in the workshop introduction and Paper III, 3. we discussed the matrix of learning outcomes (Table 7) and which of them could be applicable to the participant's practice, 4. we chose one learning outcome, 5. we discussed how this learning outcome could be assessed in the context of the participant's teaching, and 6. we discussed further steps, i.e. how the participant can proceed after the workshop to develop his practice.	100 minutes	Stimulate reflection and scaffold development of assessment activity

Data analysis

The purpose of this third stage of Study 3 was to develop a few ideas for how (some of) the learning outcomes developed in the second stage of the study could be assessed in engineering education. Therefore, my main focus in the analysis of the empirical material was to develop a rich and detailed description of the ideas for assessment that emerged during the two workshops. To develop such rich descriptions, I primarily relied on my written notes from the workshops. I then used the audio-recordings from the workshops to verify and complement my notes to ensure that my descriptions would match the participants' suggestions as closely as possible.

In addition to describing each of the assessment ideas, I also strove to identify and describe similarities and differences between these ideas. On the basis of the description of these similarities and differences, I then compared the participants' assessment ideas with theoretical descriptions of what could be considered adequate assessment in education aimed at developing students' ability to address ill-structured problems (King & Kitchener, 1994; p. 71 in this thesis). Finally, my co-author and I suggested ways in which the groups' assessment approaches could be improved (Lönngren & Svanström, 2015).

Results and conclusions

Based on suggestions from the two groups of participants in the first workshop, we have described two very different approaches to assessing engineering students' ability to integratively address wicked problems (Lönngren & Svanström, 2015). The first group focused on one specific learning outcome, the ability to "independently identify a [wicked problem] in the context of one's future profession and describe why it is a [wicked problem]" (*ibid*., p. 9). The group suggested that to assess this learning outcome, the teacher could ask students to work in pairs to identify a wicked problem, to produce video-recorded argumentation for why the chosen problem should be described as a wicked problem, and to provide written feedback to other groups' video-recorded argumentations. The second group argued that it is not possible to separate the individual learning outcomes in practice. Therefore, they suggested that all 22 learning outcomes in the matrix should be assessed in a single essay task: "The teacher should 'simply ask students to address such a big, complex problem'" (*ibid*., p. 9). We concluded that both "approaches have advantages and disadvantages", but that the high level of guidance in the first approach can provide a means for teachers to carefully and gradually scaffold students' learning. We further suggested that such a guided approach "may be useful for students who are not familiar with working with sustainable development and/or complex problems" and that "many engineering students are likely to benefit from such an approach since complex problems are seldom used in engineering education" (*ibid*., p. 11).

In the second workshop, I worked with one engineering educator to develop assessment activities for a different learning outcome in the matrix, the ability to "use duly substantiated social, ecological, economical, cultural, political, and technical perspectives to suggest, discuss, and assess alternative courses of action for a [wicked problem]". The participant in this workshop used a similar approach to the first group in the previous workshop, i.e. a highly guided approach in which students would alternate between working in groups and presenting results from their group work in a whole-class setting, as well as producing individual reflective essays. The workshop participant further suggested that it would be necessary to develop an assessment rubric (see p. 78 in this thesis for a description of assessment rubrics) for each assessment activity (e.g. one rubric for the reflective essays, one for assessing general presentation skills in the group presentations, and one for assessing the content of the group presentations). Since Paper III was published before the second workshop, the results from that workshop are not included in the paper. The results from the workshop are included here because they provide an additional example of assessment activities for assessing engineering students' ability to address wicked problems, and because they provide a link between the focus on assessment activities in Study 3 and the development and evaluation of an assessment rubric in Study 4.

STUDY 4

Methodology for Study 4

In Study 4, we used a design-based research (DBR) approach to iteratively develop and evaluate an analytic assessment rubric for assessing engineering students' written responses to wicked problems. We also developed and evaluated an educational intervention in which the rubric was used as a tool for both teaching and assessment. The intervention was implemented in an undergraduate engineering course on sustainable development. It consisted of three questions on the mid-term exam of the course, two collaborative workshops, and one question on the final exam of the course. We hoped that the intervention would provide the students with an understanding of the nature of wicked problems and with a structured way of addressing wicked problems, and we hoped that the rubric would support that learning and serve as a valid, reliable, and useful tool for assessment of student learning.

Foundations of design-based research

> For decades, policies for educational research worldwide have swung back and forth between demanding rigor above all other concerns, and increasing emphasis on impact. These two qualities need not be mutually exclusive. (…) Educational design research (…) can yield both theoretical understanding and solutions to urgent educational problems. (McKenney & Reeves, 2012)

DBR (also called educational design research) has emerged as a framework for linking educational research and practice, and for making educational research more relevant to practice (Barab, 2014; van den Akker, Gravemejer, McKenney, & Nieveen, 2006). McKenney and Reeves (2012, p. 7) defined DBR as "a genre of research in which the iterative development of solutions to practical and complex educational problems also provides the context for empirical investigation, which yields theoretical understanding that can inform the work of others". Van den Akker, Gravemejer, McKenney, & Nieveen, (2006) further stressed that the design process in DBR not only aims to *generate* practical solutions and theoretical insights, but that it is also *based on* previous research and practical experiences. An important characteristic of DBR is that it is typically conducted in authentic, real world contexts. In contrast to research in laboratory environments, in DBR it is not possible (or even desirable) to carefully control the research contexts. Therefore, DBR requires a flexible approach to research, in which the research design can be continuously adjusted in response to changes in the research context, or in response to insights that emerge from ongoing data analysis.

> The emerging theoretical and practical insights and in some cases, even the research design, adjust course based on the empirical data, which are collected in real world settings. Educational design research is structured to explore, rather than mute, the complex realities of teaching and learning contexts, and respond accordingly. (McKenney & Reeves, 2012, p. 15)

To illustrate the process of DBR, McKenney and Reeves (2012) described three main types of project phases: analysis & exploration, design & construction, and evaluation & reflection. A typical DBR-project cycles through multiple iterations of these phases until a reasonably stable intervention has been developed and/or the researchers are able to formulate design principles that can help others in designing similar interventions.

Analysis & exploration phases aim to identify and explore a problem in educational practice, typically in collaboration with educational stakeholders. During these phases, researchers ask questions about the *problem* (e.g. 'What are the discrepancies between the current and the desired situation?' and 'What is already known about the reasons for these discrepancies?'), the *context* in which the problem exists (e.g. 'Who are the stakeholders and target groups?' and 'What are the physical, organizational, and educational contexts?'), and the needs and wishes of important stakeholders. To answer these questions, researchers perform literature reviews to understand how others have experienced and addressed similar problems. Researchers also engage in empirical research in the form of "field-based investigations" (*ibid.*, p. 94), as well as informal strategies such as site visits and professional conferences. These investigations aim to provide a better understanding of the *specific* problem, problem context, and stakeholder needs. At the end of each analysis & exploration phase, researchers revise the definition of the problem that is to be addressed and develop tentative design propositions for interventions that could address the problem.

Design & construction phases are creative and generative phases in which the problem is addressed by systematically designing and implementing interventions in the actual educational context in which the problem has been identified. During these phases, researchers openly explore ideas for addressing the problem. Ideas are evaluated in terms of their feasibility, as well al potential risks and benefits. This process requires "trade-off decisions" to be made in which "practical and theoretical perspectives, derived from the field and from literature, are urgently needed to help weigh potential options" (*ibid.*, p. 119). Once a potential solution idea has been chosen, the design propositions from previous analysis & exploration phases are refined. Based on these propositions, a "skeleton design" (*ibid.*, p. 122) is developed that is successively complemented with more and more detailed specifications. Finally, a prototype intervention is designed and/or revised.

Finally, *evaluation & reflection* phases aim to evaluate the intervention/design prototype that has been developed and implemented in design & construction phases. They also aim to develop theoretical insights that can later be used to refine the design or to inform the design of similar interventions in other contexts. For these purpose, researchers use a large variety of empirical data, often a combination of qualitative and quantitative data. Researchers also typically use a combination of deductive, inductive, and/or abductive strategies to interpret the data and to develop theoretical understanding and recommendations for future application of the results. In parallel with the three phases, researchers also plan for dissemination and future use of the designed intervention to increase the potential of the intervention to contribute to addressing problems in practice.

In DBR-projects, researchers typically have multiple roles simultaneously (e.g. researcher, developer, facilitator, and/or evaluator). To mitigate risks for evaluator bias, McKenney and

Reeves suggested that researchers should actively and intentionally switch between two perspectives in the evaluation of the intervention. First, the *advocate's* perspective allows researchers to utilize unique insights gained from the first-hand, live observations and experiences of implementing an intervention. Second, the *critic's* perspective allows researchers to critically re-examine design intentions and how well they have been met in a specific intervention.

Study phases

Study 4 proceeded through nine phases, which included one *analysis & exploration* phase and four iterations of *design & construction* and *evaluation & reflection* phases in the following order:

- *Analysis & exploration phase*: perform literature reviews and revisions of previous research on wicked problems in engineering education to define the aim of the study; perform field-based investigations to identify a suitable context in which the study could be conducted and to develop a thorough understanding of that context;
- *Design & construction phase 1*: develop the first prototype of the rubric (R1) and the mid-term exam question (MQ prototype);
- *Evaluation and reflection phase 1*: implement and evaluate R1 and the MQ prototype in a small pilot study in which students answered the MQ prototype and engineering educators assessed two of these answers with R1;
- *Design & construction phase 2*: develop a revised version of the rubric (R2) and the mid-term exam question for the intervention (MQ);
- *Evaluation and reflection phase 2*: implement and evaluate R2 and the MQ in the context of the intervention, and in inter-rater assessment comparisons;
- *Design & construction phase 3*: develop a revised version of the rubric (R3) and the workshop activities for the intervention;
- *Evaluation and reflection phase 3*: implement and evaluate R3 and the workshop activities in the context of the intervention, in a professional language review of R3, and in the context of a conference presentation;
- *Design & construction phase 4*: develop a revised version of the rubric (R4) and the final exam question (FQ) for the intervention; and
- *Evaluation and reflection phase 4*: implement and evaluate R4 and the FQ in the context of the intervention, in focus group interviews with students, in inter-rater assessment comparisons, and in seminars with engineering educators and engineering education researchers; evaluate the intervention as a whole.

Theoretical background

During the analysis & exploration phase, we performed a literature review in which we identified two theoretical concepts that could guide our study and frame our data analysis: *assessment rubrics* and *scaffolding*. We used the literature on assessment rubrics to guide the development and evaluation of the rubric, and we used the literature on scaffolding to guide the development and evaluation of the intervention (including the functions of the rubric in the context of the intervention).

Assessment rubrics

Assessing the ability to reason about ten important aspects of wicked problems can be described as *performance assessment* (Lane & Tierney, 2008), i.e. assessment of students' performance on complex tasks rather than assessment of students' declarative knowledge. If properly designed, performance assessment can be used to assess students' abilities to, for example, analyze and synthesize information, apply knowledge and skills that were developed during a course to new situations, and solve real-world problems (Brookhart & Chen, 2015).

There are many different types of performance assessment. The use of *assessment rubrics* is increasingly advocated as a tool for both formative and summative assessment of students' performance on complex tasks such as essay writing (Jonsson & Svingby, 2007; Moskal, 2000). In the United States, the use of assessment rubrics is particularly wide-spread. Thus, it is not surprising that the participant in the last workshop in Study 3 (which was held in New Orleans, United States) mentioned a need to develop assessment rubrics to be able to assess student learning related to wicked problems.

Jonsson and Svingby (2007) distinguished between two kinds of rubrics. *Analytic rubrics* are used to assign independent scores for all dimensions of performance that are assessed in a task. *Holistic rubrics*, on the other hand, provide an overall judgment of the quality of performance (see also Moskal, 2000). Jonsson and Svingby suggested that analytic rubrics may be particularly useful for educational purposes since such rubrics provide detailed information about students' performance. This information can help students and teachers identify students' strengths and learning needs in formative assessment. Well-designed, analytic rubrics can also increase reliability in grading performance tasks (Jonsson & Svingby, 2007) since the pre-defined criteria for assessment can reduce subjectivity in assessment (Moskal, 2000). The research literature on assessment rubrics further suggests that analytic rubrics can provide focused learning goals and performance criteria that can guide student learning and thus improve students' performance on complex tasks (Brookhart & Chen, 2015; Jonsson & Svingby, 2007; Moskal, 2000; Reddy & Andrade, 2010). In this way, assessment rubrics may have the potential to not only facilitate assessment but also provide *scaffolding*.

Scaffolding

In the synthesis of Studies 1 and 2, we concluded that engineering students may need instructional guidance to support them in learning to adopt integrative approaches to wicked problems (p. 56). It is reasonable to assume that the need for guidance is high also when engineering students learn to reason about the ten aspects in our *framework for integratively addressing wicked problems* (p. 57). Students' need for instructional guidance has been the focus of an intense debate in the educational research literature. This debate has mainly focused on what should be considered a proper *amount* of guidance (Hmelo-Silver, Duncan, & Chinn, 2007; Kirschner, Sweller, & Clark, 2006). Recently, Wise and O'Neill (2009) suggested a more nuanced understanding of instructional guidance. They suggested that guidance could vary along three dimensions: the *amount* of guidance, the *timing* of guidance, and the *context* in which guidance is provided.

A widely used concept to describe and analyze instructional guidance is *scaffolding*. Scaffolding is often associated with Vygotsky's socio-cultural theory of learning, which describes learning as originating in social rather than individual activity. In the socio-cultural theory of learning it is assumed that social interactions create opportunities for individual development, i.e. learning occurs first on the social level and then on the individual level (Stone, 1998; Vygotsky, 1987). Van de Pol, Volman, and Beishuizen defined scaffolding as "temporary support provided [by a teacher or more knowledgeable peer] for the completion of a task that learners otherwise might not be able to complete" (van de Pol, Volman, & Beishuizen, 2010, p. 271).

Based on a review of the use of scaffolding in the educational research literature, van de Pol, Volman, and Beishuizen (2010) described three key characteristics of scaffolding: *contingency*, i.e. adapting the support to the current level of students' performance; *fading*, i.e. reducing support over time; and *transfer of responsibility*, i.e. increasing students' responsibility for their performance and learning over time. Van de Pol, Volman, and Beishuizen further developed an analytic framework in which different scaffolding *strategies* can be characterized along two dimensions: scaffolding *intentions* (meta-cognitive, cognitive, or affective) and scaffolding *means* (feeding back, hints, instructing, explaining, modeling, and questioning). In Study 4, we used van de Pol, Volman, and Beishuizens framework to design and analyze scaffolding intentions in our intervention.

Scaffolding theory has been criticized for focusing on task completion rather than learning (Stone, 1998). However, Stone (1998, p. 345) argued that in Vygotsky's description of scaffolding theory,

> much more was at stake than merely completing the task. Instead, successful scaffolding was assumed to result in a better understanding on the part of the child [or student] of what was involved in successful completion of the task. That is, a genuine change in understanding had been accomplished, not merely some end state (e.g., a completed block tower). (…) [W]hat was being scaffolded was not the completion of a specific task but, rather, the child's [or student's] understanding of how to conceptualize the task and of the proper sequence of steps toward its accomplishment.

Thus, scaffolding theory assumes that, as students develop a better understanding of a task and of the processes needed to complete the task, less and less scaffolding is needed for students to be able to perform the task (Stone, 1998; Wood, Bruner & Ross, 1976). In our intervention, we designed scaffolding with the aim to provide students with opportunities to experience the process of integratively addressing wicked problems. We expected that these experiences would help the students to develop a better understanding of what it means to integratively address wicked problems and thus provide *affordances for learning* (see Norman, 2013 for a description of the term "affordances"). In particular, we expected that, after the intervention, the students would be able to independently identify and discuss important aspects of wicked problems in the context of their future professional roles.

Analytic rubric for assessing engineering students' written responses to wicked problems

According to Moskal, "the first step in developing a scoring rubric is to clearly identify the qualities that need to be displayed in a student's work to demonstrate proficient performance" (2000, p. 3). To identify these qualities for our rubric, we primarily used our *framework for integratively addressing wicked problems* (p. 57). As described above, this framework was based on our phenomenographic description of the *integrate and balance* approach to wicked problems from Study 2 (p. 53), theoretical descriptions of wicked problems (p. 19), and

Table 10. Rubric for assessing engineering students' written responses to wicked problems

The student demonstrates an ability to consider	Grade 3	Grade 4	Grade 5
I. problem parts	Indirectly indicate at least two problem parts and indirectly indicate how they contribute to the overall problem, e.g. by indicating something that can be improved with specific improvement measures	In general terms, describe at least two problem parts, and in general terms describe how they contribute to the overall problem	In concrete terms, describe at least two problem parts, and in concrete terms describe how they contribute to the overall problem
II. improvement measures	In general terms, describe at least two improvement measures, and indicate how they could contribute to addressing the overall problem and/or one/several specific problem part(s) that contribute to the overall problem	In concrete terms, describe at least two improvement measures, and in general terms describe how they could contribute to addressing the overall problem and/or one/several specific problem part(s) that contribute to the overall problem	In concrete terms, describe at least two improvement measures, and in concrete terms describe how they could contribute to addressing the overall problem and/or one/several specific problem part(s) that contribute to the overall problem
III. interaction between problem parts	Indirectly indicate at least one example of possible interactions between problem parts, e.g. by indicating that one problem part cannot be addressed without also addressing another problem part	In general terms, describe at least one example of possible interactions between problem parts, e.g. by describing, in general terms, how one problem part can increase or decrease the impact of another problem part on the overall problem	In concrete terms, describe at least one example of possible interactions between problem parts, e.g. by describing, in concrete terms, how one problem part can increase or decrease the impact of another problem part on the overall problem, and, on the basis of the described interaction(s), prioritize which problem part(s) should be addressed primarily
IV. interaction between improvement measures	Indirectly indicate at least one example of possible interactions between improvement measures, e.g. by indicating that one improvement measure can create the conditions required for implementing another improvement measure, or that one improvement measure can increase or decrease the impact of another improvement measure	In general terms, describe at least one example of possible interactions between improvement measures, e.g. by describing, in general terms, that one improvement measure can create the conditions required for implementing another improvement measure, or that one improvement measure can increase or decrease the impact of another improvement measure	In concrete terms, describe at least one example of possible interactions between improvement measures, e.g. by describing, in concrete terms, that one improvement measure can create the conditions required for implementing another improvement measure, or that one improvement measure can increase or decrease the impact of another improvement measure, and, on the basis of the described interaction(s), evaluate the appropriateness of the described improvement measures
V. secondary problems caused by improvement measures	Indirectly indicate at least one example of secondary problems that may be caused by a specific improvement measure, where the secondary problem(s) is/are relevant to the overall problem	In general terms, describe at least one example of secondary problems that may be caused by a specific improvement measure, where the secondary problem(s) is/are relevant to the overall problem	In concrete terms, describe at least one example of secondary problems that may be caused by a specific improvement measure, where the secondary problem(s) is/are relevant to the overall problem, and, on the basis of the described secondary problem(s), evaluate the appropriateness of the described improvement measure(s)

Table 10 continued. Rubric for assessing engineering students' written responses to wicked problems

The student demonstrates an ability to consider	Grade 3	Grade 4	Grade 5
VI. stakeholders and their interests in relation to the problem and/or improvement measures	Indirectly indicate at least two relevant stakeholders and indirectly indicate their interests in relation to the problem and/or possible improvement measures	Identify at least two relevant stakeholders, and describe, in general terms, their interests in relation to the problem and/or possible improvement measures	Identify at least two relevant stakeholders, and describe, in concrete terms, their interests in relation to the problem and/or possible improvement measures, and describe, in concrete terms, at least one example of how differences between their interests could be addressed to improve the outcome of at least one specific improvement measure
VII. actors' spheres of influence	For an actor in the given professional role, delineate, in general terms, a reasonable sphere of influence in relation to the problem and/or possible improvement measures, e.g. by describing, in general terms, s.th. within and/or s.th. beyond the actor's sphere of influence	For an actor in the given professional role, delineate, in concrete terms, a reasonable sphere of influence in relation to the problem and/or possible improvement measures, e.g. by describing, in concrete terms, both s.th. within and s.th. beyond the actor's sphere of influence	For an actor in the given professional role, delineate, in concrete terms, a reasonable sphere of influence in relation to the problem and/or possible improvement measures, e.g. by describing, in concrete terms, both s.th. within and s.th. beyond the actor's sphere of influence, and describe, in concrete terms, how an actor in the given professional role could cooperate with another actor who has a complementary sphere of influence
VIII. lack of accessible information	Identify at least one piece of information that is not currently available for an actor in the given professional role, but that could be gathered and then provide valuable input for addressing the problem	In concrete terms, describe at least one piece of information that is not currently available for an actor in the given professional role, but that could be gathered and then provide valuable input for addressing the problem	In concrete terms, describe at least one piece of information that is not currently available for an actor in the given professional role, but that could be gathered and then provide valuable input for addressing the problem, and evaluate the appropriateness of realizing at least one specific improvement measure as long as this piece of information is not available
IX. the importance of incomplete control and predictability for the outcome of improvement measures	Identify at least one example of incomplete control over, or predictability of, the outcome of at least one specific improvement measure, e.g. by identifying one example of random processes that can influence the outcome of an improvement measure	In concrete terms, describe at least one example of incomplete control over, or predictability of, the outcome of at least one specific improvement measure, e.g. by describing, in concrete terms, one example of random processes that can influence the outcome of an improvement measure	In concrete terms, describe at least one example of incomplete control over, or predictability of, the outcome of at least one specific improvement measure, e.g. by describing, in concrete terms, one example of random processes that can influence the outcome of an improvement measure, and, on the basis of the described example of incomplete control and/or predictability, evaluate the appropriateness of the improvement measure(s)
X. the importance of the local context for the outcome of improvement measures	Identify at least one example of how s.th. in the local problem context can influence the outcome of at least one specific improvement measure	In concrete terms, describe at least one example of how s.th. in the local problem context can influence the outcome of at least one specific improvement measure	In concrete terms, describe at least one example of how s.th. in the local problem context can influence the outcome of at least one specific improvement measure, and, on the basis of the described influence, evaluate the appropriateness of the mentioned improvement measure(s)

research that indicates what aspects of wicked problems and what processes of addressing wicked problems may be particularly challenging for engineering students (p. 32).

We iteratively revised the rubric throughout the course of the study as indicated on p. 77. The present[7] version of the rubric (R4) is reproduced in Table 10. It is structured into ten analytic assessment criteria (corresponding to the ten aspects in the *framework for integratively addressing wicked problems*). For each criterion, three levels of performance are described in line with the grading system at Chalmers. In this grading system, grade 5 is the highest achievable grade, and grade 3 is the minimum grade required to pass. For each criterion, the performance described for higher grades includes and extends the performance described for lower grades. "Fail" is assigned for a criterion if a text does not meet the requirements for grade 3 for that criterion. The overall grade for an assignment or exam is calculated as the average of the grades for all ten criteria.

For the Swedish version of the rubric, which was used in the intervention, we also developed a set of examples of how each criterion and performance level could be achieved in a written response to a wicked problem. For each of these examples, we developed explanations for why they are considered to fulfill each criterion for each indicated performance level. The Swedish version of the rubric, the examples and the explanations for those examples are currently unpublished; they are available upon request.

Rubric-based intervention

The intervention consisted of three integrated parts: three questions on the mid-term exam of the course, two collaborative workshops, and one question on the final exam of the course.

Mid-term exam

The mid-term exam was an in-class exam without access to literature and/or Internet resources. It consisted of a mandatory part, which contributed to overall course grades, and a voluntary part, which was not graded. Study 4 focused only on the voluntary part. This part included three questions, the first of which was the main question and which I will refer to as the *mid-term exam question* (MQ). I will refer to the second and third questions as the *additional questions* (AQ1 and AQ2).

In the MQ, students were asked to address the wicked problem of *literacy in Afghanistan* (see p. 39). In designing the MQ, I used the first four design principles for wicked problem descriptions described above (p. 41). To identify a suitable topic for the wicked problem, I consulted experts in IT-engineering as well as teachers and researchers who are familiar with the wicked problem concept; I also consulted a broad variety of online resources. We deliberately chose not to follow the fifth design principle since we were interested in what aspects of the task the students would be able to perform with minimal support. The problem description provided basic demographic and developmental background information, information about the level of literacy, and information about school-related violence in Afghanistan. A short role description contextualized the task of addressing the wicked problem as part of the sustainability engagement of a large, international IT company. The

[7] July 2017

MQ did not provide detailed directions for how the students were expected to answer the question.

To get a better understanding of potential challenges the students faced when they answered the MQ, we added two additional questions, AQ1 and AQ2, to the mid-term exam. AQ1 asked the students to describe how they experienced the problem in comparison with other problems that they encounter in their educational program. This question aimed to probe the students' epistemic understanding related to the problem given in the MQ. AQ2 asked the students to describe how satisfied they were with their answers and why. This question aimed to probe students' ability to evaluate their own approach to the wicked problem given in the MQ. Based on our previous teaching experience, we expected that the students would find the minimally scaffolded task of responding to the MQ very challenging.

Collaborative workshops

Prior to the workshops, we performed a preliminary analysis of the results from the mid-term exam. The results indicated that the responses to the MQ did not fulfill most criteria in the rubric. Based on these initial results (which were later confirmed, see p. 88, we expected that the students would find the workshop activities challenging; we therefore devoted a lot of effort to designing contingent scaffolding.

Two workshops, A and B, were offered in parallel. One of my co-authors on Paper IV (MS) offered workshop A and I offered workshop B. Both workshops were repeated two and a half weeks later to allow all students to participate in both workshops. Participation in the workshops was voluntary, but we stressed that the workshops provided preparation for the final exam.

Each of the workshops targeted five criteria in the rubric; workshop A targeted criteria I-V, and workshop B targeted criteria VI-X. The workshops were structured similarly; during the first hour, the respective criteria were introduced in collaborative and explorative exercises that referred back to the wicked problem that was addressed in the MQ. During the second hour, students worked first individually and then in groups to apply the relevant parts of the rubric (criteria I-V or VI-X) to an example response from the MQ. Throughout the workshops, we used cognitive, meta-cognitive, and affective scaffolding to support students in understanding the criteria in the rubric and applying them to an example text, understanding the relevance of the tasks, adhering to the tasks with proper time management and reasonable levels of frustration, and reflecting on their own performance and learning processes.

Final exam

The final exam was a take-home exam at the end of the course and students were given nine days to complete the exam. The exam consisted of four questions that contributed equally to the final exam grade. Only one of those questions was developed specifically for Study 4. I will refer to this question as the *final exam question* (FQ).

In designing the FQ, I used all five design principles for wicked problem descriptions outlined above (p. 41). To be able to probe for transfer of learning, I chose a different wicked problem: combating *dengue fever in sub-Saharan Africa* (see p. 40). Since the final exam was a take-home exam, the students were able to independently search for information about the problem. To avoid priming the students' responses with given information about the problem, we decided to keep the amount and detail of information in the problem description to a minimum and require students to search for information themselves.

Based on our experiences from the MQ and the workshops, we expected that fulfilling the requirements of contingency, fading, and transfer of responsibility in the FQ would require more scaffolding than was provided in the MQ, but less than was provided during the workshops. We used primarily cognitive scaffolding to support the students in understanding task requirements and constructing their responses to the FQ in line with those requirements. Most importantly, the students had access to the rubric, as well as a set of short examples that illustrated and explained how different performance levels could be achieved for each criterion in the rubric (p. 79). Students' grades for the FQ contributed 10% to their overall course grades.

What kinds and levels of reliability, validity, and utility can be achieved with an analytic rubric for assessing students' written responses to wicked problems? (RQ 2.4)

Methods

Reliability analysis

To evaluate the reliability of the rubric, we engaged seven raters who each independently scored three to five student texts. We used these scores to calculate intraclass correlation coefficients (ICCs) (Brookhart & Chen, 2015; Hallgren, 2012). For this purpose, we first transformed all scores from the original grade scale (0, 3, 4, 5) to a more linear scale (0, 1, 2, 3). We then used SPSS software to calculate ICC scores for *absolute agreement* among *randomly selected raters*. We calculated both *single* and *average* measures scores.

ICC values typically range between 0 and 1, where 0 indicates absence of agreement and 1 indicates perfect agreement. Negative ICC values are also possible and indicate systematic disagreement. For traditional psychometric analyses, Cicchetti (1994) suggests that ICC values of 0.00-0.39 indicate poor reliability, 0.40-0.59 fair reliability, 0.60-0.74 good reliability, and 0.75-1.00 excellent reliability. However, many reports on the reliability of rubric assessment "present low reliability coefficients as compared to traditional psychometric requirements", especially if rubrics aim to assess performance on complex, open-ended tasks such as essay writing (Jonsson & Svingby, 2007, p. 139). In addition, studies on rubric reliability report a large variety of reliability measures that are not always comparable (Jonsson & Svingby, 2007). In lack of clear guidelines for what should be considered acceptable reliability scores, we chose to use traditional psychometric requirements as suggested by Cicchetti but to accept reliability levels that are described as "fair" in traditional studies.

Validity analysis

To evaluate the validity of our rubric, we used an *argument-based approach* developed by Kane (1992). According to this approach, evaluating the validity of an assessment instrument includes specifying how test scores should be interpreted, and providing a convincing argument for that interpretation. For our rubric, we suggested the following interpretation: *The rubric assesses engineering students' ability to reason about ten important aspects of* *wicked problems, which in turn is an important part of engineering students' ability to integratively address wicked problems*.

We further suggested that this interpretation is based on the following four assumptions:

1. The ability to reason about important aspects of wicked problems is an important part of the ability to integratively address wicked problems.
2. The rubric measures students' ability to reason about important aspects of wicked problems.
3. Assessment with the rubric is generalizable across different domains of engineering (i.e. different engineering majors), different specific wicked problems, and different raters.
4. There are no sources of systematic error (such as lack of student motivation to demonstrate their ability).

To evaluate these assumptions, we drew on empirical data from expert reviews of the rubric, teachers' observations during the intervention, student feedback during and after the intervention, feedback from teachers who used the rubric to score student texts, student performance data for different parts of the intervention, and information about the educational context in which the rubric was developed and used.

Utility analysis

To evaluate the utility of our rubric, we relied on student performance data for different parts of the intervention, the results of the reliability analysis, teachers' observations during the intervention, and feedback from teachers who had used the rubric to assess student texts. We used the performance and reliability data to establish that the rubric could be used to differentiate between different levels of student performance in summative and formative assessment. Teachers' observations during the intervention provided information about what kinds of information could be obtained during assessment with the rubric, and how that information could be used to inform teaching. Finally, we used feedback from teachers who had used the rubric to assess student texts to analyze how the teachers experienced the utility of the rubric and whether or not they would consider using it in their own teaching.

Results and conclusions

Rubric reliability

We found that reliability scores for all raters varied across the ten criteria in the rubric. We suggested that this result indicates that the criteria in the rubric are not equally clear and well-formulated and that especially those criteria for which we obtained low or negative ICC scores (criteria II, IV and X) should be clarified. We also found that variation between scores for different criteria was considerably lower for trained than for untrained raters, and that reliability generally increased with rater training. Based on these findings, we suggested that rater training to some degree can compensate for unclear criteria description. We also suggested that the findings reiterate conclusions in the literature about the importance of rater training to achieve acceptable levels of reliability in assessment (Brookhart & Chen, 2015; Hallgren, 2012; Jonsson & Svingby, 2007).

We also found that reliability increased considerably when we used average measures analysis rather than single measures analysis. Average measures analysis for trained raters resulted in the highest reliability scores, indicating fair to excellent reliability for all but one criterion (criterion IX). We concluded that it would be preferable to use an average of several teachers' scores when high levels of reliability are important, for example in summative assessment.

Rubric validity

In Paper IV, we provided a detailed argument for each of the four assumptions underlying our validity argument (p. 84). We observed that it seems implausible to assess students' ability to integratively address wicked problems with a single rubric-based assessment. We suggested that it may be necessary to combine multiple forms of assessment to cover all aspects of students' ability to integratively address wicked problems and that our rubric can be used to assess one important aspect of this ability: the ability to reason about ten important aspects of wicked problems.

We further concluded that

> the interpretation that the rubric assesses engineering students' ability to reason about ten important aspects of wicked problems (…) is warranted given five conditions: (1) The specific task that is to be assessed does not introduce construct-irrelevant difficulties, (2) raters are adequately trained so that they fully understand the criteria in the rubric and are able to resist including construct-irrelevant criteria, (3) the rubric is used in an engineering education context, (4) students have access to the rubric when performing the task of responding to a wicked problem, and (5) students are motivated to perform the task. (Lönngren, Adawi, & Svanström, 2017, p. 7)

Rubric utility

In Paper IV, we suggested that rubric utility for summative assessment is possible but that it may be prohibitively time consuming, especially if an average measures approach and rater training are needed to achieve acceptable levels of reliability. We concluded that the rubric's utility for summative assessment may be limited.

However, In line with common conclusions in the literature on the use of analytic rubrics (Jonsson & Svingby, 2007), we found that our rubric could provide high levels of utility for formative assessment. Using the rubric in formative assessment in the rubric-based intervention provided the teachers with detailed information about the students' strengths and learning needs. The teachers could successfully use this information to design contingent scaffolding throughout the remainder of the intervention. We suggested that "if resources are limited and formative assessment is only needed on a collective level, for example for developing contingent scaffolding for the class as a whole, a teacher could choose to assess a subset of students' responses" with the rubric (Lönngren, Adawi, & Svanström, 2017, p. 7).

We also found that the rubric provided opportunities for teacher professional development. All teachers who used the rubric to assess student texts reported that they learned a lot from this activity, for example that it helped them to better understand what students need to learn to be able to integratively address wicked problems. We suggested that the high level of detail in our rubric was important for its utility for teacher professional development. The high level of detail facilitated engaged and critical discussions among teachers, both in collaborative workshops and in one-on-one discussions that I had with some of the teachers.

How do students perform when different scaffolding strategies are used in a rubric-based intervention? (RQ 2.5)

Methods

Empirical material

To address RQ 2.5, we used quantitative and qualitative data from the exams and the workshops. Data from the mid-term exam included 85 student responses to the MQ, 82 responses to the AQ1, and 84 responses to the AQ2 (out of 94 students who participated in the mid-term exam). Data from the final exam included 97 responses to the FQ (out of 101 students who participated in the final exam). Data from the workshops included the results of the students' group assessment of the example response to the MQ, written observations (provided by the course examiner) of what the students did during the workshops and what questions were raised, and workshop leaders' written reflections about how they experienced the workshops. Out of 101 students who completed the course, 53 participated in both workshops, 29 participated in one of the workshops, and 18 did not participate in any of the workshops.

Data analysis

To analyze students' performance in the written exams, I graded all written responses to the MQ and FQ with the last version of the rubric (R4). This analysis resulted in quantitative data in the form of 10*85 + 10*97 grades (one grade for each criterion in the rubric) on grade levels 0, 3, 4, or 5. I also analyzed the AQ1 and AQ2 to see whether, and how, including responses to these questions would change students' grades on the mid-term exam.

I first used Microsoft Excel to calculate overall grades for all written responses (i.e. average of the grades for all ten criteria in the rubric). I also used Excel to obtain descriptive statistics about mean grades and standard errors for the overall grades and the grades on all individual criteria in the MQ, AQ1, AQ2, and FQ. I then used SPSS software to identify significant differences between mean grades. For this purpose, I first transformed the data from the original scale of grades 0, 3, 4, and 5 to a more linear scale of grades 0, 1, 2, and 3. I then used the General Linear Model for repeated measures to analyze differences in performance *within* groups of students (performance on the MQ versus performance on the MQ+AQ1+AQ2 for all students, and performance on the MQ+AQ1+AQ2 versus performance on the FQ for all students), and I used the nonparametric Mann-Whitney U-test to compare mean grade improvement *between* groups of students (mean grade improvement from the MQ+AQ1+AQ2 to the FQ for students who participated in both workshops versus students who participated in none of the workshops). For the purpose of consistency, I report all results in the original grade scale (0, 3, 4, 5).

To get information about students' performance during the workshops, I analyzed the results from students' assessment of the example text during workshops. For this purpose, I used descriptive statistics to get information about the level of agreement between the groups' assessments. An important assumption in this analysis is that it is unlikely that a high level of agreement could have been achieved if the students had not developed a good understanding of the criteria in the rubric, i.e. a high level of agreement is assumed to indicate good performance during the workshops. In addition, in the group of authors, we reviewed and discussed the data from workshop observations and reflections to get an understanding of how well the students were able to perform the other tasks during the workshops.

Results

I found that performance during the *mid-term exams*, without access to the rubric and prior to instruction, was poor. Mean grades for all but one criterion (II) were below the pass threshold (grade 3). If students' responses to the AQ1 and AQ2 were included in the analysis, performance increased slightly but significantly for four of the criteria (VI, VIII, IX, and X) and the overall grade. However, mean grades for all but one criterion (II) were still below the pass threshold (see Figure 3).

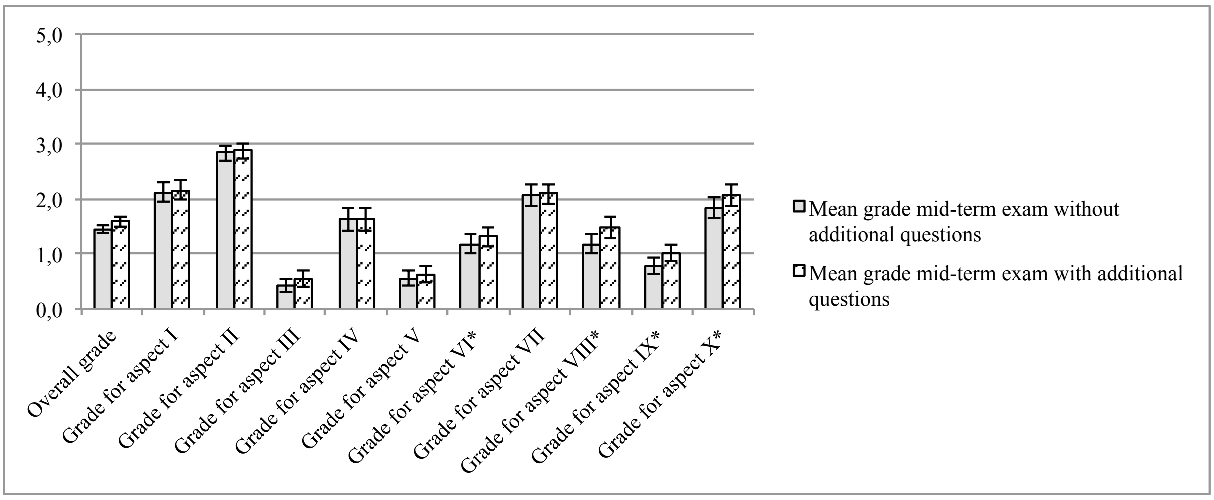

Figure 3. Results for students' performance on the mid-term exam for the MQ only and MQ+AQ1+AQ2 respectively. The difference between the mean grade for the MQ and MQ+AQ1+AQ2 is significant only where indicated with '*'. The significance level is 0.05; error bars indicate standard errors.

In the analysis of students' performance during the *workshops*, I found a high level of agreement among the student groups in their assessment of the example response to the MQ: 81% of perfect agreement among 16 groups that assessed criteria I through V, and 87% of perfect agreement among 15 groups that assessed criteria VI through X. This indicates that the students had developed a good understanding of the criteria in the rubric and how to apply them to an example response. The analysis of the observations and reflections from the workshops further suggested that all present students actively participated throughout the workshops and that they were able to perform all exercises. The students discussed all criteria in the rubric to a reasonable degree and asked relevant questions. We observed that minor difficulties remained in workshop A with respect to criteria III and IV, and in workshop B with respect to criteria VIII and IX. We revised the descriptions of these criteria in the rubric prior to the final exam.

Regarding performance during the *final exam*, I found that the mean grade in the FQ was above the pass threshold (grade 3) for all criteria in the rubric and for the overall grade. In comparing students' performance in the final exam (FQ) to that during the mid-term exam (MQ+AQ1+AQ2), I also found that students' performance had increased significantly for all criteria in the rubric and for the overall grade (see Figure 4). We found a possible floor effect for students' performance on criteria I and II; most students achieved the highest possible grade on these criteria, which indicates that many students may not have been able to demonstrate their full potential in fulfilling these criteria.

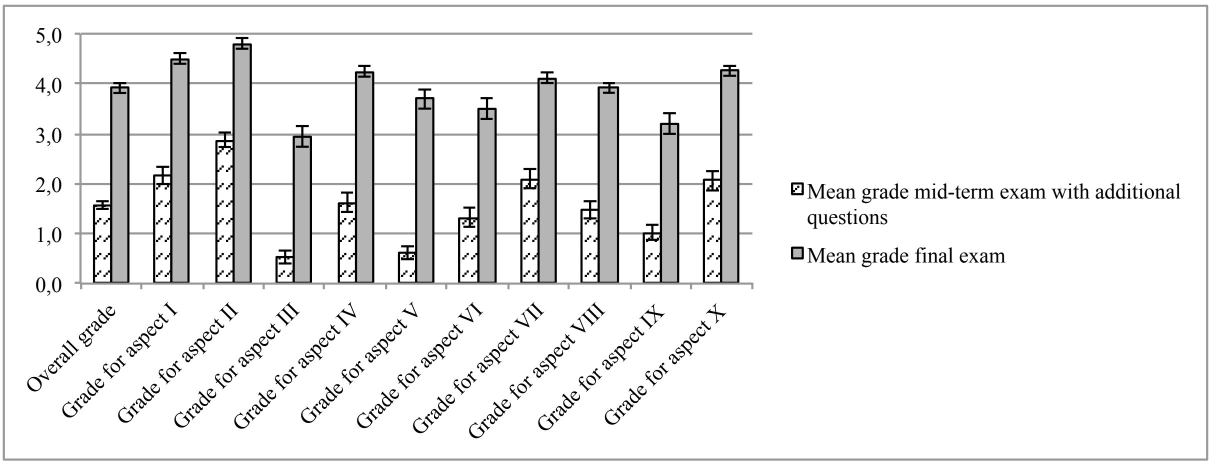

Figure 4. Results for students' performance on the mid-term exam (MQ+AQ1+AQ2) and final exam (FQ) respectively. The difference between the mean grades in the mid-term exam and the final exam is significant for the overall grade and all individual criteria. The significance level is 0.05; error bars indicate standard errors.

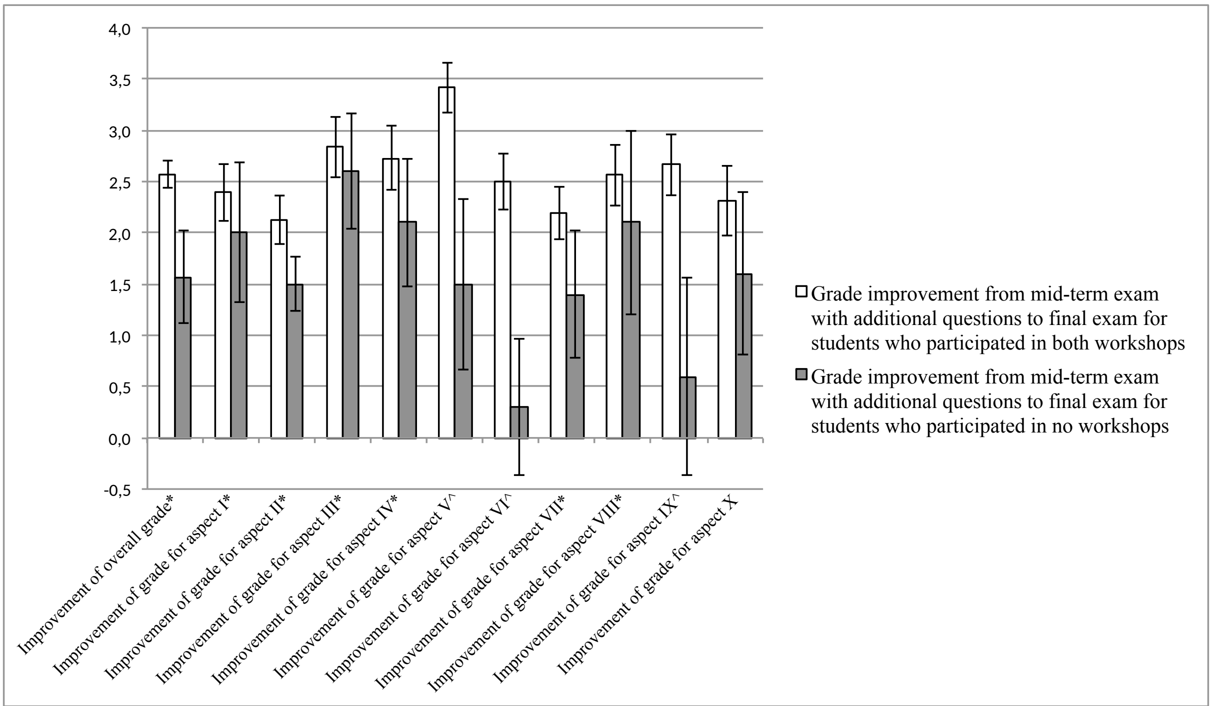

Figure 5. Results for improvement in performance between mid-term exam (MQ+AQ1+AQ2) and final exam (FQ) for students who participated in both workshops and students who participated in none of the workshops, respectively. Grade improvement for students who participated in both workshops is significant for the overall grade and for all individual criteria. Grade improvement for students who participated in none of the workshops is significant only where indicated with '*'. The difference between grade improvement for students who participated in both workshops versus no workshops is significant only where indicated with '^'. The significance level is 0.05; error bars indicate standard errors.

Finally, I compared students' performance on the FQ depending on whether they had participated in both or none of the workshops. For students who had participated in both workshops, I found that performance increased significantly between the MQ+AQ1+AQ2 and the FQ for all individual criteria and for the overall grade. For students who did not participate in any of the workshops, performance increased significantly for the overall grade and for six of the criteria (I, II, III, IV, VII, and VIII). In a direct comparison between the groups' performance, I found that the groups of students who had participated in both workshops performed significantly better on three of the criteria (V, VI, and IX). These are three of the four criteria for which no significant increase in performance was found for the group of students who did not participate in any of the workshops. These results suggest that participation in the workshops had a significant effect on students' performance on criteria V, VI, and IX. The results are summarized in Figure 5.

How do students approach wicked problems when different scaffolding strategies are used in a rubric-based intervention? (RQ 2.6)

Methods

Empirical material

To address RQ 2.6, we used a wide range of qualitative data. Data from the mid-term exams include students' responses to the MQ, AQ1, and AQ2, and their reports on how much time they had spent answering the MQ. Data from the workshops included the course leader's written observations, the workshop leaders' written reflections, and a total of 134 student responses to one-minute papers that were administered and filled in at the end of each of the workshops. In the one-minute papers, the students were asked to complete the following statements: 'The most important thing I've learned today is …', 'Something that is still unclear to me is …', and 'General comments about the workshop: …'. Data from the final exam included students' responses to the FQ.

To complement the data that was collected during the intervention, I conducted three focus group interviews with four, three, and one student(s), respectively, about one week after completion of the course. Participants were selected with the aim to maximize variation in perspectives about the intervention. Out of a total of 40 students who were invited to participate, eleven were invited because they had expressed either positive or negative attitudes towards working with wicked problems in a mandatory course in their educational program. We expected that these students could contribute with different perspectives on the intervention. Four of these students volunteered to participate. The remaining students were invited based on random selection. To avoid bias in exam grading, interview participants' final exams were graded before the interviews, and the resulting grades were not discussed during the interviews. The interviews lasted 30-50 minutes and free lunch was provided. Written, informed consent was obtained in advance from all interview participants; the interviews were audio-recorded for later analysis.

During the interviews, the participants were asked to explain how they had approached the FQ and how they had experienced working with the FQ. The discussions focused particularly on how the participants had used the rubric in the process, but also on how and when they had gathered background information for the FQ and whether they had used any of the other tools that were introduced in the workshops (mind mapping and questioning assumptions). Further questions concerned whether the participants had experienced any difficulties in the process of answering the FQ, and if so, how they had tackled them. Participants were also asked to explain how they had decided that their responses were 'good enough', and how they thought that they would address wicked problems in the future.

Finally, we collected additional data through course evaluation forms and a course evaluation meeting. The evaluation forms were administered electronically to all students directly after the course; 42 students completed the form. The evaluation meeting was held about six weeks after completion of the course and lasted about one hour. I participated in that meeting together with three student representatives from the course, the course examiner, and the program director.

Data analysis

To analyze the data, I first summarized the audio data from the focus group interviews and transcribed selected parts. I then identified and clustered data that provided information about

- how seriously the students took the tasks, e.g. how much time they said that they had spent on a specific task or what they wanted to achieve in a task;
- how the students' experienced the tasks, e.g. whether they reported that they found a task interesting, challenging, or relevant; and
- the tools and strategies that students used to address the tasks, especially how they addressed challenges

In discussions among all authors, we used these data clusters to develop composite accounts of students' approaches during the different parts of the intervention. Where appropriate, we used direct quotes from the focus group interviews to illustrate interesting aspects of students' approaches.

Results

The analysis of students' reports of how they experienced the *MQ* indicates that the students found the scope of the MQ overwhelming and that they were frustrated with not knowing how to adequately respond to the MQ. In focus group interviews and in responses to the AQ2, students also indicated that they did not take the MQ very seriously since they prioritized answering the other questions in the mid-term exam for which they received credit (which they did not for the MQ). At the same time, in their responses to the AQ2, many students indicated that they were reasonably satisfied with their responses to the MQ – a judgment that differs significantly from my assessment of their responses with the rubric (see p. 88).

As mentioned above, all students who participated in the *workshops* were actively engaged in the exercises throughout the workshops. Workshop observations indicate that the students took the tasks seriously and that they wanted to do them as correctly as possible. However, during the workshops and in one-minute papers, students also indicated that they experienced a considerable amount of frustration when they realized that it would not be possible to find perfect solutions to the tasks. This was particularly evident in the students' comments during the workshops when they realized that it would be impossible to perform 100% objective and reliable grading. In all qualitative data, students also reported that they felt stressed during the workshops because they found that there was too little time to adequately perform the tasks. In focus groups, students also mentioned that they had rushed through the tasks, often at the expense of deep engagement with challenging questions.

The analysis of students' approaches during the *final exam* indicated that, similar to the MQ, students felt overwhelmed by the scope and complexity of the FQ. In particular, students reported that they experienced difficulties in identifying IT-related improvement measures.

However, contrary to the MQ, students also indicated that they took the task very seriously and that they were motivated to try to achieve the highest possible grade for the FQ.

The rubric influenced how the students approached the FQ. In focus group interviews, students reported that the rubric provided a good structure for working with the FQ and that they had followed the rubric very closely as they worked with their responses to the FQ. While this structure seems to have helped students to address challenges in responding to the FQ, it may also have limited students' reflections about the quality or relevance of the improvement measures that they suggested in their responses. However, one student also reported that the rubric sparked reflections that he/she had not made without the rubric, e.g. the question of whether or not nature could be seen as a stakeholder in relation to a wicked problem. Several students also mentioned that they had used the rubric to self-assess their responses to the FQ prior to submission.

What affordances for learning do different scaffolding strategies provide in the rubric-based intervention? (RQ 2.7)

Methods

To address RQ 2.7, I deductively applied van de Pol, Volman, and Beishuizen's (2010) framework for analyzing scaffolding strategies to the results obtained for RQs 2.5 and 2.6. I first categorized the scaffolding provided during different parts of the intervention in terms of scaffolding intentions (cognitive, meta-cognitive, or affective). Then, I categorized all empirical material that provided evidence about the presence of absence of affordances for learning to integratively address wicked problems in different parts of the intervention. Finally, I combined these categorizations to draw conclusions about what scaffolding strategies may have supported what affordances for learning (see Table 11 for examples that illustrate how this categorization was done).

Table 11. Examples that illustrate how data were categorized to identify affordances for learning to address wicked problems in the different parts of the intervention.

Part of the intervention	Scaffolding intention	Evidence for (+) and against (-) affor-dances for learning provided through the respective scaffolding strategies	Data sources that provide the evidence
Mid-term exam	Cognitive	- Students reported that they did not understand the purpose of the MQ.	Focus group interviews
Mid-term exam	Meta-cognitive	- Students' evaluation of their performance on the MQ differed strongly from the teacher's evaluation.	Performance data, responses to the AQ2
Workshops A & B: assessment exercises	Affective	+ Students reported that they saw the value of the exercise for the final exam, - but that they did not see the value for learning to address wicked problems.	One-minute papers, focus group interviews, course evaluation forms
Workshops A & B: assessment exercises	Affective	+ Students reported that seeing a written response to a wicked problem from an IT-perspective had helped them to see the relevance of IT for addressing wicked problems, which in turn increased their motivation to try to find a way to use IT to address wicked problems in their future roles as professional IT-engineers.	Focus group interviews
Final exam	Cognitive	+ Students reported that the rubric had helped them in structuring their response to the FQ. - Students reported that they sometimes had followed the rubric 'slavishly' rather than focusing on the quality of the solutions presented in their responses. - Experts and teachers who used the rubric to assess written responses to wicked problems noted that the rubric does not allow assessment of the quality of proposed solutions and that it might encourage students to use the rubric as a recipe for writing a response, which in turn might lead to a surface approach to learning.	Focus group interviews, expert reviews, inter-rater discussions
Final exam	Meta-cognitive	+ Students reported that they had used the rubric to self-assess their responses to the FQ prior to submission.	Focus group interviews

Results

We found that scaffolding during the *mid-term exam* was minimal for all scaffolding intentions. We also found that this minimal scaffolding may have been one of the reasons for students' superficial engagement with the MQ, which in turn may have provided limited affordances for learning to adopt integrative approaches to wicked problems during the mid-term exam. However, students' responses to the MQ, AQ1, and AQ2 provided valuable information about what the students struggled with in addressing the wicked problem. This information enabled the teachers to design contingent scaffolding for the workshops and the FQ, which could be used to create affordances for learning in the intervention as a whole. In addition, in focus group interviews, students indicated that engaging with a wicked problem prior to the workshops may have prepared them for the workshops and thus supported their learning during the workshops. This interpretation is consistent with the description of learning in scaffolding theory (p. 78); the MQ provided an opportunity for the students to experience what it means to engage with a wicked problem in their disciplinary context. While the students in the MQ did not gain experience of how to integratively address wicked problems, the experience of engaging with a wicked problem prepared them for further experiences during the workshops and the FQ.

Overall, we found that the *workshops* provided adequate cognitive and meta-cognitive scaffolding to provide affordances for learning to understand the criteria in the rubric and to apply them to an example response to the MQ. We also found that the workshops provided adequate affective scaffolding for maintaining a high level of motivation during the workshops. However, students also indicated that they were frustrated with the lack of time to complete all exercises and about the impossibility to achieve 100% objective and reliable grading, which indicates a lack of affective scaffolding for dealing with negative emotions. The students' frustration with the lack of time and objectivity may have distracted them from fully utilizing the learning affordances provided during the workshops.

In focus group interviews, some students also indicated that they would have preferred more detailed instructions on how to perform the exercises in the workshops to avoid 'wasting time on doing the exercises wrongly', i.e. they would have preferred stronger cognitive scaffolding. Such an approach might have reduced students' tendency to rush through the exercises (and some of their frustration), which could have supported deeper engagement with difficult questions. However, it might also have eliminated difficult questions by providing too much structure and instead contributed to encouraging a surface approach. We therefore concluded that more cognitive scaffolding probably would not have improved affordances for learning during the workshops.

The evidence regarding affordances for transfer of learning from the workshops to other contexts was less clear. On the one hand, students reported that they were unlikely to use the tools that were introduced during the workshops to address wicked problems in the future. On the other hand, students also reported that they had gotten a better 'feeling' for what is important when addressing wicked problems and that they may be more likely to address for example stakeholder interests when addressing a wicked problem than they were before the intervention. Students also mentioned that the workshops helped them understand that, and how, IT could contribute to addressing wicked problems, something they had not thought about before attending the course. Again, this interpretation is consistent with scaffolding theory in the sense that the workshops provided opportunities for the students to experience what it means to integratively address wicked problems, and thus to gain a better understanding of the processes that are involved in integratively addressing wicked problems.

We further found that the combination of cognitive, meta-cognitive, and affective scaffolding that was provided during the *final exam* was sufficient to allow students to respond to the FQ and to self-assess their responses prior to submission. In particular, we found that the rubric provided significant amounts of cognitive scaffolding that helped students develop a written response to a wicked problem. This finding is in line with suggestions in previous research that assessment and learning are closely related (Baartman, Bastiaens, Kirschner, & van der Vleuten, 2007; Gibbs & Simpson, 2005; Pereira, Flores, & Niklasson, 2016; Scouller, 1998). It is also in line with suggestions that assessment rubrics can provide useful tools for learning (Panadero & Jonsson, 2013).

According to scaffolding theory, engaging in educational tasks provides affordances for learning. Consequently, scaffolding that supports students' engagement in those tasks should increase affordances for learning. However, the fact that students closely followed the rubric as they worked with the FQ indicates that the rubric may have provided *too much* cognitive scaffolding. This conclusion was supported in many of the expert reviews and inter-rater discussions. Thus, the strong cognitive scaffolding provided by the rubric may have provided affordances for learning how to use the rubric to create a written response to a wicked problem and to understand the criteria in the rubric, while at the same time limiting affordances for learning to use a holistic, flexible, and creative approach when addressing a wicked problem, i.e. it may have inadvertently tamed the process of addressing wicked problems in the written exam. This conclusion is in line with common critiques of some scaffolding approaches as overly focused on task completion, at the expense of deep engagement and transferable learning (e.g. Roll, 2012). However, reducing the amount of cognitive scaffolding may increase students' experience of insecurity and frustration as they engage with challenging tasks such as addressing wicked problems. Therefore, reducing the amount of cognitive scaffolding may require teachers to simultaneously increase the amount of affective scaffolding. Increasing the amount of meta-cognitive scaffolding may also be beneficial, especially when cognitive scaffolding is limited (An & Cao, 2014).

DISCUSSION

The aim of the research underlying this thesis was to contribute to improving engineering education practice such that it could better prepare engineering students to contribute to sustainability, with a particular focus on wicked problems in engineering education. To develop a "rich, composite account" (Bell 2004, p. 251), I conducted four studies that complement and build on each other. To match the specific research aims and questions in each of the studies, I used different theoretical perspectives and methodological approaches. I open this chapter with a *discussion on the use of different theoretical perspectives* in my research. For this purpose, I describe similarities and differences between two of these perspectives, a phenomenographic perspective on learning (which was used in Study 2) and sociocultural theory of learning (which was used in Study 4). Based on this description of similarities and differences, I discuss how the use of different perspectives may have influenced the research, both in terms of opening up possibilities for addressing different kinds of research questions and in terms of what knowledge claims can be made on the basis of the research. To complement this broad discussion of the project as a whole, I then discuss potential *limitations in the individual studies*. Finally, I discuss *contributions and implications* of the research described in this thesis and suggest questions for future research. For this purpose, I integrate findings from all studies rather than reporting contributions and implications in relation to each of the 14 individual research questions (see p. 6 and Figure 1) that have been addressed in the context of this Ph.D. project.

Figure 6 provides an overview of the results, conclusions, contributions, and implications of the research described in this thesis. The results and conclusions listed in Figure 6 are largely identical to those listed in Figure 1, but they are organized differently. In Figure 1, results and conclusions were listed as responses to the individual research questions. In Figure 6, the results, conclusions, contributions, and implications are organized according to different themes in the research. The results, conclusions, contributions, and implications in each column in Figure 6 are related to each other, but each column may contain elements that emerged in response to several research questions (for example, the first column contains results, conclusions, and contributions from Studies 1 and 2, as well as from the synthesis of the two studies).

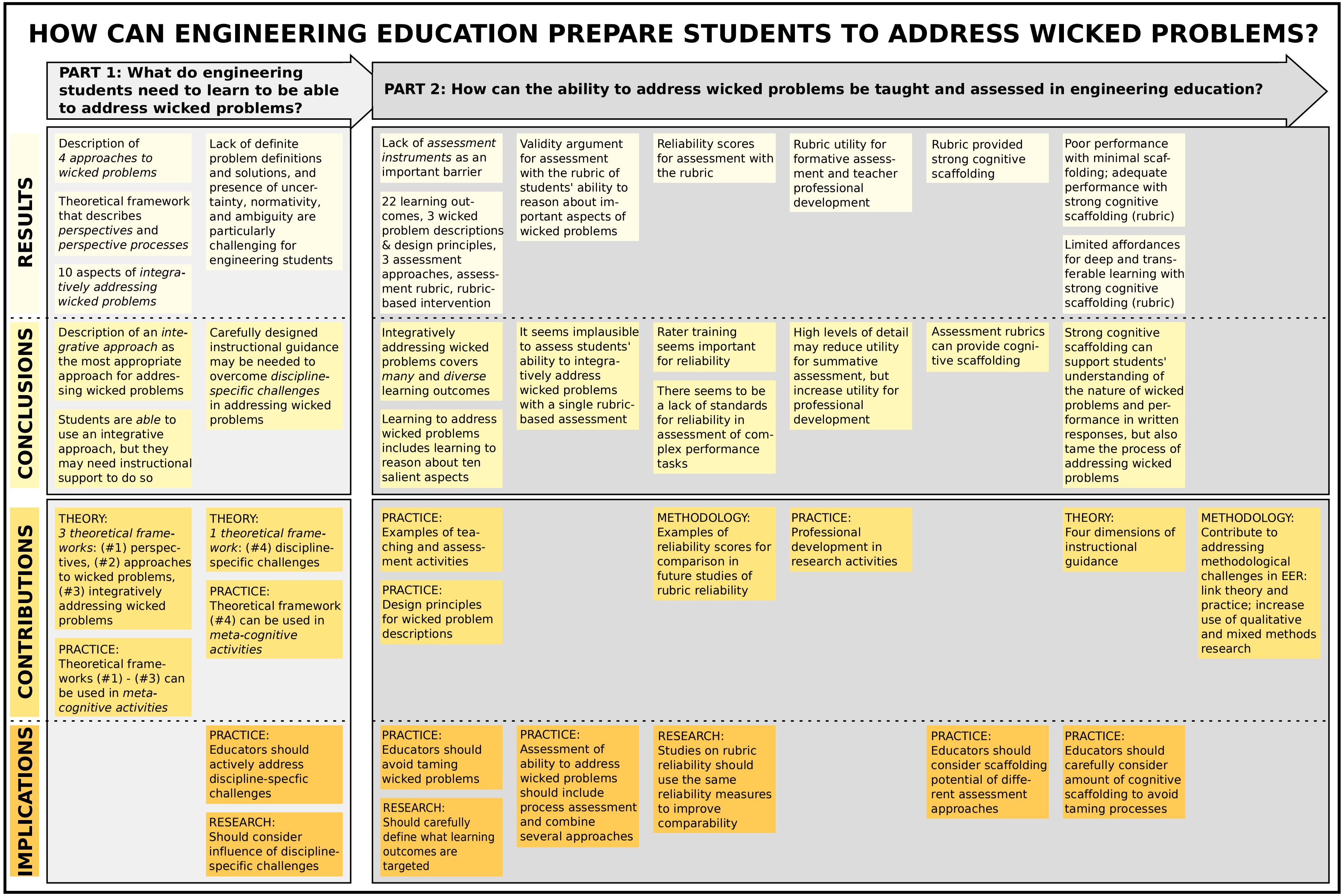

Figure 6. Overview of results, conclusions, contributions, and implications of the research.

On the use of multiple theoretical perspectives[8]

> The range of theories [in educational research] do different sorts of intellectual work for us – at different timescales and units of analysis – and the development of more comprehensive accounts of learning and competence will need to leverage on multiple theoretical perspectives, and thereby different research traditions, to develop *rich, composite accounts.* (Bell 2004, p. 251, emphasis added)

In the EER literature, scholars stress the importance of choosing theoretical perspectives (Borrego, Douglas, & Amelink, 2009; Jawitz & Case, 2009) and methodological approaches (Borrego, Douglas, & Amelink, 2009; Case & Light, 2011; Koro-Ljungberg & Douglas, 2008) based on the nature of the research aims and questions. At the same time, if researchers lack awareness of the "implications [of different theoretical perspectives] for methodology and method", calls for using more diverse research approaches can lead to "naïve pleas for an eclecticism of method which appears uninformed by deeper levels of understanding about epistemological or ontological foundations" (Hart 2006, p. 540). Scholars who have been trained in the dominant post-positivist paradigm of EER may lack awareness of the importance of different theoretical perspectives and methodological approaches. For example, Koro-Ljungberg and Douglas (2008) found that authors of papers that were published in the Journal of Engineering Education between 2005-2006 sometimes applied quantitative analytical methods to qualitative data. Similarly, Borrego, Douglas, and Amelink (2009) found that engineering education practitioners sometimes lament a lack of qualitative studies in EER and a bias towards quantitative approaches in the peer review process, but at the same time, the study participants still enacted a quantitative, post-positivist bias when they were the ones critiquing other researchers' work (see p. 10 in this thesis for a more detailed discussion). In this way, "advocates of methodological eclecticism (…) fail to take seriously the inescapable assumptions and values that accompany the use of a method and the pursuit of practically useful results" (Yanchar & Williams, 2006, p. 3).

In response to the lack of awareness of theoretical perspectives and tendencies of methodological eclecticism in EER, Jawitz and Case (2009) and Koro-Ljungberg and Douglas (2008) called for more explicit engagement with theoretical perspectives and methodological approaches in EER publications. They argued that such explicit engagement could contribute to developing awareness in the field of diverse theoretical perspectives and methodological approaches, help readers to more adequately interpret and evaluate research results, and support the development of a common language among researchers who are interested in similar research phenomena.

Yanchar & Williams (2006, p. 9) provided a similar argument for educational research. They suggested that it is possible to conduct research that avoids both rigid adherence to dominant approaches and uninformed eclecticism by

- making theoretical perspectives explicit;

---

[8] "Theoretical perspectives" in this section can refer to both *research paradigms* (such as post-positivism or interpretivism) and *theories of learning* (such as phenomenographic theory of learning or social constructivist theory of learning). Both types of theoretical perspectives influence what kinds of research questions can be addressed and what kinds of knowledge claims can be made.

- creatively adopting existing methods to suit the particular subject matter and context of the research rather than following existing descriptions of methods through "formulaic or mechanical procedures";
- conducting research "in a reasonably coherent manner, so that theories, questions, methods, data analysis, and criteria for success form meaningful and interpretable accounts"; and
- "viewing the assumptive framework that guides researchers' choices and interpretations as fallible, alterable, and in need of critical examination".

To design and conduct research projects that are both flexible and coherent, Slife's (1987, p. 96) description of a form of "noneclectic eclecticism" in psychology may be helpful. In such an approach, "individual theories are embraced one at a time [based on what is most appropriate in a given situation], and not the best elements of several theories." In other words, researchers do not uncritically combine incompatible theories; rather, they use different, but in themselves coherent, theoretical perspectives one at a time to address specific research aims and questions.

In the research underlying this thesis, I have used such a "noneclectic eclectic" approach. In each unit of research, I have used one theoretical framework at a time. Between these units of research, I have made theoretical shifts that have influenced the overall results in important ways. In the remainder of this section, I will focus on the most important theoretical shift in my research: from a phenomenographic approach in Study 2 to a social constructivist approach in the analysis of scaffolding strategies in Study 4. To discuss this shift, I will first explore the different research foci that are enabled by phenomenographic and social constructivist theories of learning. I will then compare theoretical assumptions in the two theories. Finally, I will discuss how the theoretical shift (in terms of research foci and assumptions) may have influenced the validity of the research results in Study 4.

Complementary research foci in research based on phenomenographic and social constructivist theories of learning

In Study 2, we used a phenomenographic approach to investigate and describe different ways in which engineering students approached a wicked problem (p. 53). Even though this was not the focus of the phenomenographic analysis, we also observed that the same students used different approaches in different parts of the interviews; the students were clearly able to adopt an integrative approach to a wicked problem, but they only did so in some parts of the interviews (p. 56). One explanation for this finding is that different parts of the interviews provided different conditions that enabled or hindered students from expressing an integrative approach to the wicked problem. In the context of the phenomenographic interviews, the most obvious change in conditions was different kinds of questions and comments from the interviewer. However, these changes could not be analyzed with a phenomenographic approach because, in phenomenographic analysis, all data excerpts are combined in a common pool of meaning. Thus, each data excerpt is analyzed in relation to all other data excerpts rather than in relation to questions or comments that preceded or followed an excerpt in the interviews. I concluded that further research with a different methodological approach, based on a different theoretical perspective, was needed to complement the findings from the phenomenographic study. In particular, I concluded that I needed to use a research approach that would allow me to explore how instructional guidance (for example in the form of questions and comments) could facilitate students' use of different kinds of approaches to wicked problems. One theoretical perspective that explicitly focuses on how instructional guidance can facilitate learning is social constructivist theory. I used social constructivist theory to inform my analysis of scaffolding strategies in Study 4 (p. 78).

The different foci in phenomenographic and social constructivist research mean that the two approaches have different explanatory power for different kinds of educational phenomena. They can therefore be used to address different kinds of research questions and to develop different kinds of knowledge claims. In the research underlying this thesis, the two approaches provided complementary tools that together allowed me to develop a "more comprehensive account" (Bell, 2004, p. 251) of what it means to teach and learn to integratively address wicked problems in the context of engineering education than would have been possible if I had relied on a single theoretical perspective. While I have not used different theories of learning within the context of the same study, I have transferred results from Study 2 to Study 4, i.e. from a research context in which a phenomenographic approach was used to a research context in which a social constructivist approach was used. This transfer of results made it necessary to critically examine and address similarities and differences between the theoretical assumptions in the two approaches as well as possible ways in which these differences might have influenced the results in Study 4.

Comparing theoretical assumptions in phenomenographic and social constructivist theories of learning

Phenomenographic and social constructivist theories of learning share common features, but they also differ in important ways. Both approaches have been developed as reactions to then-dominant positivist and behaviorist approaches to educational research (Liu & Matthews, 2005; Svensson, 1997). Both approaches also set out to overcome the Cartesian mind-body dualism (Marton & Booth, 1997; Liu & Matthews, 2005), which is built on the assumption that mind and body are fundamentally different and separable entities. Such an assumption creates theoretical dilemmas in educational research. For example, if mental and physical phenomena are fundamentally different and separate, how could it be possible for them to interact such that learning could occur? To resolve this and similar dilemmas, both phenomenographic and social constructivist theories assume "an inextricable relationship between the individual's learning and the learning context" (Micari, Light, Calkins, & Streitwieser, 2007, p. 462).

While phenomenographers and social constructivists share an ambition to overcome mind-body dualism, they have addressed this challenge in different ways. In the phenomenographic tradition, the main focus has been on developing a non-dualist ontology by dissolving the distinction between the knower and the known (Marton & Booth, 1997; Svensson, 1997), while largely ignoring social aspects of learning (Richardson, 1999). In social constructivism, the main focus has instead been on developing a holistic understanding of the learner in his/her social context (Liu & Matthews, 2005), while retaining a division between the knowers and the known (Marton & Booth, 1997).

The different ways in which phenomenographic and social constructivist researchers have attempted to overcome mind-body dualism have led to different assumptions about the nature of knowledge. In the phenomenographic tradition, learning is viewed as becoming aware of educationally critical aspects of a phenomenon and of different ways in which aspects of a phenomenon are related to each other. It is assumed that students develop a more and more complete understanding of a phenomenon as they learn to differentiate between different aspects and to discern relationships between those aspects (p. 52). Marton and Booth (1997,

p. 64) illustrated the phenomenographic view of learning with the phenomenon of “number” in mathematics education:

> In the *complete* experience of number *all* aspects [of the phenomenon] are present simultaneously. In a lesser way of experiencing number, different aspects are present in a parallel fashion, side by side but separate. In even less advanced ways, certain aspects are held in focus in one case and others in another (emphasis added).

In other words, in phenomenography, it is assumed that the phenomenon is complex in itself and that learning is about experiencing the full extent of this complexity. Learning from this perspective happens in the context of a relationship between a learner and the experienced phenomenon:

> The most fundamental assumption [in phenomenography] is that knowledge and conceptions have a relational nature. Conceptions are dependent both on human activity and the world or reality external to any individual. (Svensson, 1997, p. 165)

Svensson (1997, p. 165) contrasted the relational perspective in phenomenography to other theoretical positions, including constructivist theories of learning:

> The position taken [in phenomenography] differs from empiristic and positivistic assumptions about observations as facts, and knowledge as inductively based on facts. It also differs from rationalistic, mentalistic and constructivistic assumptions about knowledge as rational or mental constructions within a more or less closed rational and/or mental system. Thus the view of knowledge [in phenomenography] is that it is relational, not only empirical or rational, but created through thinking about external reality. (Svensson, 1997, p. 165)

In the above quote, Svensson argued that constructivist theories of learning rely on the assumption that learning is a *rational* process through which knowledge is *internalized* in the mind. This assumption about internalization of knowledge is also present in social constructivist theories, even if the process of individual internalization is preceded by knowledge development in a social context:

> every function in the child's (…) development appears twice: first, on the social level, and later, on the individual level; first between people (interpsychological), and then *inside* the child (intrapsychological). (Vygotsky 1987, p. 57, emphasis added)

The assumption that knowledge can be internalized also entails an assumption that it is possible to abstract knowledge from the specific context in which it was developed and to achieve “generalized or abstracted perception” (Liu & Matthews, 2005, p. 394). In fact, Liu and Matthews (2006, p. 392) argued that “in Vygotsky’s theory, the development of intellect and rationality beyond situations is *the central aim of education*” (emphasis added). Similarly, Akkermann (2001, p. 90) argued that Vygotsky viewed learning as

> proceeding from local to general, from context-bound to context-free, from externally-supported to internally-driven (or ‘mentalised’). Accordingly, cognitive achievements are gauged in terms of three major acts of distancing: 1. The ability to emerge from here-and-now contingencies (characteristic of practical intelligence), 2. The ability to extract knowledge from its substrate (i.e. from contexts of use and personal goals); and 3. The ability to act mentally on virtual worlds, carrying out operations in the head instead of carrying them out externally.

In summary, social constructivist and phenomenographic theories hold different assumptions about the nature of knowledge. Social constructivist theories assume that knowledge is first developed in a *relationship between several minds*, and then *internalized in individual minds*. In this process, knowledge is abstracted and generalized from the specific contexts in which it was developed. More advanced knowledge is thus assumed to be more rational and context-free than less advanced knowledge. In phenomenography, on the other hand, it is assumed that knowledge is developed in a *relationship between an individual mind and an external reality*. Thus, more advanced knowledge is assumed to be equally relational and context-sensitive as less advanced knowledge; more advanced knowledge is instead assumed to be more complex in that it keeps more (or more important) aspects, and relationships between aspects, in focal awareness.

Influences of the theoretical shift on the validity of the results in Study 4

> If theoretically incompatible methods are to be used jointly, then some theoretical integration or subsumption must be performed to render the combined data workable and interpretable. (Yanchar & Williams, 2006, p. 5)

Studies 2 and 4 were based on different theoretical perspectives. In Study 2, we developed a phenomenographic description of four different approaches to a wicked problem (p. 53). In Study 4, we used the description of one of these approaches as a basis for developing an assessment rubric for assessing written responses to wicked problems (p. 79). This transfer of results from Study 2 to Study 4 required theoretical subsumption in which the results from Study 2 were abstracted and decontextualized. Since abstraction and generalizability cannot be assumed for results of phenomenographic research (Svensson, 1997), it was necessary to explore in what ways this process of abstraction may have influenced the validity of the research in Study 4. We found that the abstraction of the phenomenographic results influenced the validity of the results from Study 4 in two ways. First, it reduced the richness of the description of an integrative approach to wicked problems such that important aspects of this approach were lost in the development of the assessment rubric in Study 4. This meant that we had to reduce the *scope* of our validity claim for the rubric. Second, it resulted in different *levels* of validity in assessment with the rubric depending on how similar the assessment context was to the context in which the phenomenographic description of an integrative approach had been developed. In the following section, I will discuss each of these two points in detail.

Reduced richness of description leads to reduced scope of validity claim

In Study 4, we originally wanted to assess students' ability to integratively address wicked problems. In an early version of an abstract for Paper V (which was written before all data from Study 4 had been analyzed), we wrote that we aimed to develop "an analytic rubric for assessing the quality and comprehensiveness of engineering students' written responses to wicked problems". However, as we developed a better understanding of how students and educators interacted with the rubric, we realized that our initial aim was unrealistic. For example, we realized that the rubric does not assess how coherently a text argues for a certain way of addressing a wicked problem. We also realized that the rubric does not assess the quality of proposed improvement measures (such as whether suggested improvement measures are realistic or whether they involve high levels of risk for serious, unintended consequences). Consequently, as we analyzed the validity of the rubric, we found that we needed to adjust our original "validity argument" (Kane 1992). We found that we could not claim to completely assess students' ability to integratively address wicked problems; instead, we argued that the rubric could be used to assess "students' ability to reason about ten important aspects of [wicked problems], which in turn is an important part of engineering students' ability to integratively address [wicked problems]" (Lönngren, Adawi, & Svanström, 2017, p. 4). The new validity argument is less comprehensive than the argument we originally strove for, which indicates that important aspects of an integrative approach to wicked problems may have been lost in the theoretical transition from Study 2 to Study 4. However, even with a less comprehensive validity argument, the results of Study 4 provide valuable results that contribute to the aim of this thesis.

The level of similarity of context influences the level of validity that can be achieved

In Study 4, we also noted that the validity of the rubric seemed to depend on the context in which it was used, which indicates that the process of abstraction from Study 2 to Study 4 was incomplete. In particular, we noted that the validity of the rubric for analyzing responses to the *mid-term* exam question might have been lower than the validity of the rubric for analyzing responses to the *final* exam question. In Paper V, we discussed two possible sources of systematic error in the mid-term exam that could have compromised validity in that context: (1) that the students might not have addressed all criteria in the rubric because they did not know that they were supposed to, and (2) that the students might not have been motivated to perform well since they did not receive grade credit for their responses. One explanation for potentially compromised validity in the assessment of the mid-term exam responses is that the context in the mid-term exam differed significantly from the context in the phenomenographic interviews in Study 2. For example, in the analysis of the phenomenographic interviews, students' initial reflections about the wicked problem were excluded "since expressing an understanding of a phenomenon requires that one has already established a relationship to it. In the case of problems as complex as [a wicked problem], establishing a relationship can be expected to take some time" (Lönngren, Ingerman, & Svanström, 2016, p. 6). Since the students did not receive any instruction about wicked problems prior to the mid-term exam (p. 82), it is likely that they did not establish a stable relationship to the wicked problem during the exam. Therefore, the mid-term exams might have elicited responses that were similar to the initial reflections in the phenomenographic study rather than similar to the excerpts from the phenomenographic interviews that were actually included in the analysis in Study 2. An important conclusion is that abstraction of the results from Study 2 may have been incomplete and that the rubric retains contextual aspects of the original description of an integrative approach to wicked problems. Therefore, the validity of the rubric needs to be evaluated for each specific context in which it is used.

Limitations in the individual studies

Limitations related to participant selection

In all studies, I have striven to maximize relevant variation among informants. However, for ethical and practical reasons, I did not always achieve maximum variation. Most importantly, in all studies, I only interviewed students and educators who explicitly volunteered to participate. While this was a necessary procedure to ensure ethical conduct, it may have limited the variation in participants' opinions about, and interest in, ESE in engineering education. For example, in Study 4, I tried to involve an equal number of students in each of the three focus group interviews (i.e. an equal number of students who had participated in both, one, and no workshops respectively). However, it was difficult to find volunteers among students who had not participated in any of the workshops, possibly because these students on average were more critical to the intervention as a whole. Thus, I had focus group interviews with four respectively three students who had participated in two or one workshops, but only one student who had not participated in any of the workshops. This may have reduced the richness of the findings with regard to students who had not participated in any of the workshops, and possibly for students who were critical of the intervention. Similarly, in Study 3, it would have been beneficial if more educators from different engineering disciplines had participated in the workshops during which participants developed assessment activities.

Limitations related to analysis processes

All studies have been conducted in collaboration between several researchers, which allowed us to triangulate interpretations among the researchers. Ideally, each researcher would have independently analyzed each data set before discussing the results among the research team. We followed this approach in Study 4 for evaluating inter-rater reliability for the rubric. However, for all other analyses, my co-researchers did not have the time and resources to analyze all of the material. Instead, I identified empirical examples that I found representative of our developing interpretations and empirical examples that did not seem to fit these interpretations. We used these examples in meetings in the research teams to discuss and refine the interpretations.

Limitations related to the scope of the studies

All four studies described in this thesis were case-based studies (Svensson 2016), i.e. the purpose of the studies was to explore specific cases of teaching and learning rather than to develop decontextualized and broadly generalizable results. According to Svensson, "an important result of the investigation [in case-based studies] is to give a basis for both comparisons with descriptions of other cases and an improvement of the delimitation and description of cases in new investigations" (2016, p. 283). To explore how the descriptions of our cases might compare to other cases, it would have been beneficial if each of the four studies had been repeated in several different contexts. For example, Studies 1 and 2 could have included interviews with students from several educational institutions and several engineering disciplines, at different times in the students' education, and about several different wicked problems. Similarly, in Study 4, it would have been beneficial to implement and evaluate the rubric and the rubric-based intervention in several different educational contexts and with several different wicked problems. Repeating the studies in other contexts was beyond the scope of this thesis but could be added in future research projects.

In Study 4, we drew conclusions about affordances for learning provided in differently scaffolded educational activities. However, we were not able to draw conclusions about how sustainable and transferable that learning was. To be able to draw such conclusions, we would have needed to include longitudinal data about the impact of the intervention. Again, such a study was beyond the scope of my project.

Finally, in Study 4, we achieved reasonable levels of reliability for the rubric under certain conditions. However, the confidence intervals for the reliability scores were quite large. This reduced the strength of conclusions drawn from the reliability data. To achieve smaller confidence intervals, each rater would have needed to assess a considerably larger number of student texts (probably around 20 texts each instead of three to five as was the case in our study). However, raters reported that they spent an average of one to two hours on each of the texts they assessed. We did not have the resources to compensate raters for their work; therefore, we could not ask them to assess a large number of texts.

Limitations related to methodological approaches

In the interviews for Studies 1 and 2, I as an interviewer took an active part in the discussions about the wicked problem; I actively and repeatedly challenged the participants' approaches to the problem and tried to encourage the use of many different perspectives on the problem. This interview procedure may have favored an *isolate and succumb* approach (Study 2; see p. 53 in this thesis). It may also have favored *perspective shifting* processes (Study 1) since it involved repeatedly encouraging participants to consider new perspectives rather than exploring each perspective in more depth. The fact that participants used non-favored approaches, such as an *integrate and balance* approach or *perspective integration* processes, indicates that the interview procedure did not limit the participants to exclusively use favored approaches. However, it is possible that a different interview procedure could have allowed the participants to express *additional* perspectives, perspective processes, and approaches to the wicked problem. Future studies could attempt to elicit these additional approaches and thus extend the descriptions developed in Studies 1 and 2.

In Study 4, we encountered methodological challenges related to the dual focus of design-based research on developing theoretical and practical contributions. To maximize internal validity in the analysis of how student performance changed between the mid-term and final exam, we would have needed to use a traditional pre- and post-test design in which the mid-term and final exams would have been as similar to each other as possible. As discussed in Paper V, this was not possible since the study was embedded in a real educational context. On the other hand, a traditional pre- and post-test design would not have allowed us to study affordances for learning in a naturalistic context.

Similarly, the dual focus of the design-based research approach meant that two of the researchers involved in Study 4 were also involved in teaching during the rubric-based intervention. Thus, the researchers had dual roles in the study, which may have influenced their interpretation of the results. The risk of undue influence on the interpretation of results was reduced by involving a third researcher in the study who was not involved in teaching. The risk could have been reduced even more if the intervention had been delivered by teachers who were not involved in the research. However, that would simultaneously have reduced the researchers' familiarity with the research context and their ability to develop rich and contextualized interpretations of the empirical data.

Contributions to, and implications for, EER

There is a relatively large body of literature in which the wicked problems concept has been applied and refined, and in which strategies for addressing wicked problems have been suggested and discussed. However, much of this literature is primarily focused on raising awareness about the nature of wicked problems and the need for alternative approaches to addressing them (p. 23). Xiang suggested that many of the descriptions of strategies for addressing wicked problems are not based on empirical investigations and that there is a general lack of "well-grounded theoretical explorations or empirical investigations" (2013, p. 2) in the literature on wicked problems. There is also a general lack of research on teaching and learning to address wicked problems in the context of engineering education, as illustrated by the low number of publications that specifically refer to wicked problems and engineering education (see Table 3 on p. 24).

The research described in this thesis contributes to the literature on wicked problems in engineering education by adding descriptions of four empirical studies that can contribute to a better understanding of what it means to be able to address wicked problems and how such an ability to address wicked problems could be taught and assessed in engineering education. Specific contributions to research include four theoretical frameworks, a description of four dimensions of instructional guidance, concrete examples of how theory and practice can be linked in EER, concrete examples of qualitative and mixed-methods research in EER, and reliability data that can contribute to developing a basis for reliability standards in assessment of complex competencies.

Theoretical frameworks

In Study 1, we identified *perspectives* as an important concept in discussions about the multidimensionality of sustainability problems. Previous descriptions of the concept focused on specifying *thematic* aspects of perspectives that need to be considered in addressing (wicked) sustainability problems (Seager, Selinger, & Wiek, 2012; Wals & Blaze Corcoran, 2006). However, as Bengtsson and Kronlid (2016) argued, it is impossible to specify *all* thematic aspects that should be considered in the context of ESE. Our *framework for perspectives* (framework #1 in Figure 6) extends previous descriptions in the literature by describing *non-thematic* characteristics of perspectives (coverage, depth, complexity), suggesting that Brentano's concept of *intentionality* (Brentano, 1874/2009) can be used as a philosophical basis for defining the nature of perspectives, and describing *meta-perspective* as a kind of perspective that can support meta-cognitive approaches to perspectives.

Previous descriptions of perspectives indicated the importance of *perspective shifting processes* in addressing wicked problems (Wals & Blaze Corcoran, 2006). Our framework refines and extends these descriptions by contrasting perspective shifting processes with *five additional perspective processes* and arguing that, for addressing wicked problems, some of these processes (perspective integration, reflection, and evaluation) may be even more important than perspective shifting processes. Our *framework for perspectives* can potentially be used as an analytic tool, for example to analyze the quality of students' perspectives and the prominence of different kinds of perspective processes in their reasoning about a wicked problem. In these ways, the research in Study 1 contributes to the literature that is focused on suggesting strategies for addressing wicked problems (p. 21).

Many of the strategies for addressing wicked problems that have been suggested in the literature address only a subset of the characteristics of wicked problems (Duckett, Feliciano, Martin-Ortega, & Munoz-Rojas, 2016). These descriptions can be used to identify specific tools and techniques that students can learn, which would enable them to address *some* of the characteristics of wicked problems, but not all of them. The results of Study 2 extend these previous descriptions by describing and empirically illustrating approaches that students used when they attempted to simultaneously address *all* characteristics of a wicked problem (p. 53). For example, the descriptions of the *divide and control* and *isolate and succumb* approaches provide empirically based descriptions of approaches in which students *tame* wicked problems (p. 22) and thus do *not* address all characteristics of wicked problems. The description of the *integrate and balance* approach, on the other hand, provides an empirically based description of an approach in which students succeed in simultaneously addressing all characteristics of wicked problems.

The descriptions of the four approaches may also provide a useful *framework for approaches to wicked problems* (framework #2 in Figure 6), for example as an analytic tool for identifying the relative amount of time students spend using each of the approaches as they attempt to address a wicked problem. Similarly, the *framework for integratively addressing wicked problems* (p. 57; framework #3 in Figure 6) may provide an analytic tool for exploring different processes that engineering students engage in when they use an integrative approach to addressing wicked problems. Finally, the *description of discipline-specific challenges* in addressing wicked problems (p. 32; framework #4 in Figure 6) can provide an analytic framework for exploring how different characteristics of educational contexts relate to students' ability to integratively address wicked problems.

Four dimensions of instructional guidance

Learning to address wicked problems is challenging for engineering students. Therefore, it is reasonable to assume that engineering students need well-designed instructional guidance to support this challenging process. Unfortunately, educational researchers are divided about how instructional guidance can support different kinds of learning. Specifically, the research community is divided about the value of explicit instructional guidance. On one side of the debate are those whose primary argument is that students have limited cognitive resources and that instruction, therefore, should eliminate task-irrelevant processes to maximize the cognitive resources available for learning. Proponents of this view argue that instructional approaches with high levels of direct instruction are generally more effective and efficient than instructional approaches with low levels of direct instruction (Kirschner, Sweller, & Clark, 2006). On the other side of the debate are those who argue that there are many different forms of instructional guidance and many different ways of reducing cognitive load during learning. Proponents of this view argue that direct instruction is just *one* of many useful techniques to provide instructional guidance. They also argue that high levels of direct instruction may support certain kinds of learning outcomes at the expense of others and that it may limit transfer of learning to other contexts (Hmelo-Silver, Duncan, & Chinn, 2007).

Wise and O'Neill attempted to reframe this debate into a more multidimensional discussion of how instructional guidance can support different kinds of learning outcomes. They argued that, for both sides of the debate, "the *quantity* of guidance is just one dimension along which guidance can be usefully characterized". They further argued that other important concerns are "the *context* in which guidance is delivered and the *timing* with which guidance is delivered" (2009, p. 82, italics in original).

In study 4, we found that even the *kind* of guidance (i.e. different forms of scaffolding) may be important to consider. In particular, we found that more *cognitive* guidance may not always promote deep and transferable learning of complex learning outcomes such as the ability to address wicked problems, but that more *affective* or *meta-cognitive* guidance could be beneficial for that kind of learning. Thus, we suggest that there are at least four dimensions of instructional guidance that can influence in what ways instructional guidance can support student learning.

Methodological contributions

The research described in this thesis contributes to addressing methodological challenges in EER. One important challenge in EER is *to link theory and practice* such that research is grounded in, and applicable to, practice (Borrego, Streveler, Miller, & Smith, 2008; Jesiek, Borrego, & Beddoes, 2010). Two methodological approaches that are commonly described as research approaches that can link theory and practice are *action research* and *DBR* (Barab, 2014; Johansson & Lindhult, 2008; McKenney & Reeves, 2012; van den Akker, Gravemejer, McKenney, & Nieveen, 2006). In the research underlying this thesis, we have used an action research approach in Study 3 and a DBR approach in Study 4. The descriptions of these studies provide concrete examples of how these approaches can be used in EER and can thus provide "model publications" for other researchers who want to use similar approaches (Borrego, Foster, & Froyd, 2014, p. 46).

Another methodological approach that can contribute to linking theory and practice is phenomenography. In fact, phenomenography has been developed with the explicit aim to contribute to educational development (Bowden, 2000; Collier-Reed & Ingerman, 2013; Marton, 2015; Marton & Booth, 1997). Phenomenographic results are commonly used to design learning studies in which variation identified through phenomenographic analysis is used to design and evaluate different approaches to teaching (see e.g. Pang & Ling, 2012; Pang & Marton, 2003). Results from phenomenographic research have also been used to design program evaluation instruments, but typically not as a basis for developing instruments for assessing individual students' performance (Micari, Light, Calkins, & Streitwieser, 2007). In Study 4, we have successfully used the phenomenographic results from Study 2 for this purpose. Thus, the research illustrates a novel way of linking phenomenographic research and practice in the context of engineering education.

A second methodological challenge in EER is to *increase the number and quality of qualitative and mixed-methods studies* (p. 13). The research described in this thesis contributes with descriptions of four studies in which qualitative and mixed-methods approaches have been used. Indirectly, these descriptions can also inspire other researchers in the field to use similar methodological approaches, which would further increase the share of qualitative and mixed-methods research in EER.

Finally, the research described in this thesis also contributes to addressing the challenge of developing standards of reliability in assessment of complex competencies; the results of the reliability analysis in Study 4 provide a concrete example of reliability scores that could be achieved in assessment with a highly detailed analytic assessment rubric. In future research, these scores can be compared with scores that are achieved in similar studies and thus serve as a basis for discussing what levels of reliability can be expected in assessment of complex competencies.

Recommendations for research on wicked problems in engineering education

The presence of discipline-specific challenges in addressing wicked problems (p. 32) indicates that the disciplinary context matters for research on how students learn to address wicked problems. Thus, I suggest that researchers in engineering education should carefully consider how the presence of these challenges might influence research processes and results. For example, transferability of results to and from other disciplinary contexts may not be straightforward.

In Study 3, we found that the ability to integratively address wicked problems covers many and diverse learning outcomes; in Study 4, we found that we needed to adjust our validity claim to a more limited learning outcome than we had initially attempted. These results suggest that it is important to carefully define learning outcomes in research on students' ability to integratively address wicked problems.

In Study 4, we observed a lack of comparability among studies that aim to evaluate the reliability of assessment rubrics for complex performance tasks. To improve comparability and eventually establish accepted standards for reliability, the research community should agree on one or two kinds of reliability scores and use these to report results of reliability analyses. Based on the literature that I reviewed for Study 4, I suggest that ICC scores and Krippendorff's alpha may provide reliability scores that could be accepted by most researchers in the field.

Contributions to, and implications for, engineering education practice

Teaching and learning to integratively address wicked problems is challenging for both students and educators (p. 5). In addition, educators experience a lack of instructional tools (such as assessment instruments, see p. 66) that could support them in teaching their students to address wicked problems. The research described in this thesis contributes to engineering education practice through concrete examples of teaching and assessment activities, design principles for wicked problem descriptions, and frameworks that can be used in meta-cognitive activities in engineering education aimed at developing students' ability to address wicked problems.

Examples of teaching and assessment activities

In the context of the research described in this thesis, we have developed a variety of concrete examples of teaching and assessment activities that engineering educators can use in, or adapt to, their teaching:

- 22 intended learning outcomes related to the ability to integratively address wicked problems (p. 69),
- three assessment approaches that can be used to assess (some of) these learning outcomes (p. 73),
- an assessment rubric that can be used to assess engineering students' ability to reason about ten important aspects of wicked problems (p. 79),
- three descriptions of wicked problems that can be used in teaching and assessment activities (p. 39), and
- an educational intervention in which the rubric and two of the wicked problem descriptions were used in an engineering education context (p. 82).

In many seminars and presentations about the research described in this thesis, educators commented that the concrete examples inspired them to improve their own teaching, for example by developing more concrete learning outcomes or assessment activities that are better matched to the learning they want to support. One of the descriptions of wicked problems (*water-shortage in Jordan*, p. 39) is currently used in teaching and research at the University of British Columbia in Vancouver, Canada (Nesbit, 2017).

Design principles for wicked problem descriptions

As discussed throughout this thesis, engineering education is today characterized by a dominance of well-structured story problems and a lack of engagement with ill-structured problems such as wicked problems (Jonassen, Strobel, & Beng Lee, 2006). One contributing factor may be the lack of examples of descriptions of ill-structured and wicked problems (a notable exception is a collection of non-traditional thermodynamics problems developed by Riley, 2012). The research described in this thesis contributes three examples of wicked problem descriptions (p. 39). However, to achieve adequate representation of wicked problems in engineering education, many more examples will be needed and educators will need to be able to adapt these problems to their own contexts and/or develop their own problem descriptions.

Unfortunately, designing wicked problem descriptions is a challenging process and there has been a lack of concrete guidelines to support this process. I have tried to address this lack by providing a set of design principles for wicked problem descriptions (p. 41) that can help educators to ensure that they are indeed working with untamed, wicked problems. Indirectly, these design principles can also contribute to addressing the general lack of examples of wicked problem descriptions for use in engineering education.

Frameworks for use in meta-cognitive activities in engineering education practice

As described above (p. 107), the research described in this thesis contributes four theoretical frameworks (##1-4 in Figure 6) that can be used as analytic tools in EER. These frameworks can also provide tools for engineering education practice. Educators can, for example, use them to support students' meta-cognitive development.

According to a recent review of research on meta-cognition in science education (Zohar & Barzilai, 2013, p. 121f), "metacognitive instruction is applied for improving students' metacognitive thinking, for improving students' skills (such as reading skills, problem-solving skills or higher-order thinking skills) or for improving students' knowledge and conceptual understanding". Zohar and Barzilai argued that "the call for teaching metacognition is considered one of the main three recommendations for improving instruction that emerged from over three decades of research about how people learn". Flavell, Miller, and Miller (2002) distinguished between meta-cognitive knowledge, skills, and experiences. They described meta-cognitive *knowledge* as "knowledge, beliefs, ideas and theories about people as 'cognitive creatures' and about their diverse interactions with cognitive tasks and strategies" (*ibid.,* p. 123). They further described three sub-categories of meta-cognitive *knowledge*: *knowledge of persons*, which refers to knowledge of the personal variables that influence cognitive activity; *knowledge of tasks*, which refers to an understanding of how the nature of different tasks influences cognitive activity; and *knowledge of strategies*, which refers to "knowledge about thinking, learning and problem-solving strategies that students might use in order to achieve goals". Meta-cognitive *skills* are described as "the skills and processes used to guide, monitor, control and regulate cognition and learning". Meta-cognitive knowledge and skills are primarily described in cognitive terms. Meta-cognitive *experiences*, on the other hand, are described as containing both cognitive and affective elements; they include "feelings, judgements or estimates" that "may be conscious and analytic, but often are non-conscious and non-analytic" (*ibid.*, p. 123).

Educators can use our theoretical frameworks to support the development of students' meta-cognitive knowledge and students' ability to deal with meta-cognitive experiences. First, educators can use our *framework for approaches to wicked problems* (p. 53; framework # 2 in Figure 6) to engage students in meta-cognitive discussions about different ways of addressing wicked problems. In fact, in many of our research presentations and seminars in which results from Study 2 have been presented and disseminated, educators commented that they would like to use the descriptions of the four phenomenographic categories in that way (without us prompting them to think about this possibility). Second, educators can use the *framework for integratively addressing wicked problems* (p. 57; framework #3 in Figure 6) to discuss the nature of an integrative approach in more detail. Third, educators can use the *framework for perspectives* (p. 49; framework #1 in Figure 6) to engage students in meta-cognitive discussions about the value of different kinds of perspectives and perspective processes for addressing wicked problems. In this way, the three frameworks can support educators in developing students' meta-cognitive knowledge of *tasks* (i.e. the specific requirements of the task of addressing wicked problems) and *strategies* (i.e. what strategies students can use in addressing wicked problems). Fourth, educators can use the *description of discipline-specific challenges* (p. 32; framework #4 in Figure 6) to explicitly discuss and address discipline-specific challenges together with their students. In this way, educators can help students develop meta-cognitive knowledge of *persons* (i.e. the characteristics of engineering (education) culture that may create discipline-specific challenges in addressing wicked problems). Finally, educators can use the phenomenographic *description of the isolate and succumb approach to wicked problems* (which is part of framework #2 in Figure 6) to develop students' ability to deal with meta-cognitive *experiences*. In particular, educators can use this description to explicitly discuss situations in which students experience wicked problems as impossible to address. Thus, educators may help students to *not* succumb but persist in trying to integratively address wicked problems through an iterative process.

Teacher professional development

In Studies 3 and 4, we chose to directly involve engineering educators in the research process. Through this involvement, our research contributed to teacher professional development. For example, in Study 4, educators who had used the rubric to assess student texts commented that using the rubric helped them develop a better understanding of the characteristics of wicked problems and of what students need to learn to be able to address wicked problems (Lönngren, Adawi, & Svanström, 2017). This finding is in line with previous research about the value of assessment activities for teacher professional development. For example, Parr, Glasswell, and Aikman (2007) found that teachers who engaged with rubrics for assessing students' writing also developed their own understanding of the range of functions of writing, developed a shared language about writing, and improved their teaching practice. Similarly, Ash and Levitt found that teachers who "strategically and intentionally participate in formative assessment practices can undergo profound transformation in their professional growth" (2003, p. 23; see also McKenney & Reeves, 2012).

The results of the research described in this thesis can also be used to design activities for teacher professional development. For example, our assessment rubric can be a valuable tool for developing teachers' ability to develop their own assessment instruments for specific educational contexts. The results from Studies 3 and 4 further suggest that intensive, collaborative workshops can provide fruitful opportunities for educators to develop their own understanding of wicked problems and of how to they could teach their students to address wicked problems. Such workshops can also provide opportunities to develop new examples of wicked problem descriptions, intended learning outcomes, assessment approaches, and contingent scaffolding. Thus, the activities can serve dual functions of teacher professional development and the development of a resource bank that can be used to facilitate other educators' attempts to include wicked problems in their teaching.

Recommendations for designing teaching and assessment activities

In engineering education practice, educators should try to *avoid taming wicked problems*. Our design principles for wicked problem descriptions (p. 41) can provide a useful tool to ensure that the problems one is using are indeed wicked. Educators should also carefully consider how much and what kinds of scaffolding they should provide to adequately support students' engagement with the challenging task of addressing wicked problems *without inadvertently taming the process of addressing wicked problems*.

Educators should also pay attention to discipline-specific challenges and provide adequate scaffolding that can help students overcome these challenges. Apart from using our *framework for discipline-specific challenges* in meta-cognitive activities (see p. 111), educators can also attempt to address *disciplinary egocentrism*, which Richter and Paretti (2009) described as the inability to see connections between one's own discipline and interdisciplinary problems (such as wicked problems), and an inability to understand and value differences in disciplinary perspectives. Richter and Paretti suggested that educators can help students overcome the challenges of disciplinary egocentrism by explicitly discussing the discipline's "modes of thinking and methodologies" and possible ways in which the specific discipline could contribute to addressing an interdisciplinary problem (2009, p. 40). In this way, educators can support students in developing more complex views of the nature of knowledge and more nuanced understandings of – and engagement with – social and political aspects of engineering.

In developing teaching and assessment activities, educators should evaluate how different kinds of assessment approaches can support learning. In particular, educators should evaluate the scaffolding potential of different assessment approaches and consider prioritizing approaches that have a large potential to support the intended learning – even if those approaches may be more difficult to implement and lead to somewhat lower levels of reliability in assessment (see also Gibbs & Simpson, 2005).

Finally, since the ability to integratively address wicked problems covers a wide range of learning outcomes (p. 69), educators should combine different kinds of assessment approaches to assess students' ability to address wicked problems. This conclusion is in line with studies on assessment in higher education that suggested that a variety of assessment methods should be used rather than exclusively relying on written tests (Baartman, Bastiaens, Kirschner, & van der Vleuten, 2007; Pereira, Flores, & Niklasson, 2016). Educators need to carefully consider what kinds of learning outcomes can be assessed with different assessment approaches. In particular, educators should ensure that they not only assess the *product* of students' attempts to address wicked problems, but also the *processes* that students use to arrive at these products (King & Kitchener, 1994).

Further research

The research described in this thesis contributed to addressing the questions of what engineering students need to learn to be able to address wicked problems, and how that ability could be taught and assessed in engineering education. However, many questions remain.

For example, there is a general lack of consolidation and convergence in the literature on wicked problems (Duckett, Feliciano, Martin-Ortega, & Munoz-Rojas, 2016; Turnbull & Hoppe, 2017). To advance research on how education can prepare students to address wicked problems, it would be valuable to consolidate the existing research from different fields such as EER and ESER. However, in the face of discipline-specific challenges in addressing wicked problems, it is unclear whether, and in what form, such a consolidation could be valid. To better understand the conditions under which results can be consolidated, future research could for example explore discipline-specific challenges in other disciplines.

To better understand how the ability to integratively address wicked problems can be *taught* in engineering education, future research should investigate how the four dimensions of instructional guidance (quantity, context, timing, and kind of guidance) interact in different educational contexts. Such research could provide a better understanding of how different approaches to instructional guidance can support different kinds of learning. For example: Are different kinds of guidance more useful in certain educational contexts or at certain points in time during an intervention? Do different kinds of guidance require more careful adjustments of the quantity and timing of the guidance? Or should different amounts or kinds of guidance be used in different kinds of educational contexts? More specifically, considering engineering students' discipline-specific challenges in addressing wicked problems, do engineering students require more and/or different kinds of instructional guidance than students in other majors when they learn to address wicked problems?

More research is also needed to better understand how the ability to integratively address wicked problems can be *assessed* in engineering education. For example, while we could suggest that educators should combine different assessment approaches for assessing students' ability to integratively address wicked problems, it remains unclear exactly how different assessment approaches should be combined. What kinds of assessment approaches should be used? How could they be combined? And how do different assessment approaches interact when they are combined in different ways? It also remains unclear how different assessment approaches (both formative and summative) can provide support for learning. For example, what kinds of scaffolding can different kinds of assessment approaches provide? And what kinds of learning do these different kinds of scaffolding in assessment activities support?

CONCLUSION

Current practice in engineering education does not adequately prepare students to address wicked problems, such as many sustainability problems. This thesis provides engineering education-specific descriptions of what it means to be able to address wicked problems, and of how the ability to address wicked problems can be taught and assessed. The results of the research include:

- a description of engineering education-specific challenges in addressing wicked problems;
- 3 wicked problem descriptions and design principles for wicked problem descriptions for use in engineering education;
- description of four different approaches that engineering students have used in addressing a wicked problem;
- 22 intended learning outcomes, 3 assessment approaches, an analytic assessment rubric, and a rubric-based intervention for developing students' ability to integratively address wicked problems;
- validity, reliability, and utility evaluations of the assessment rubric; and
- insights about students' performance, their approaches to wicked problems, and affordances for learning in differently scaffolded activities during the rubric-based intervention.

Important conclusions based on these results include the following:

- An integrative approach can be considered most appropriate for addressing wicked problems; engineering students may be able to use such an integrative approach but they may need instructional support to do so.
- Carefully designed instructional guidance may be needed to overcome engineering education-specific challenges in addressing wicked problems.
- Strong cognitive scaffolding can support students' understanding of the nature of wicked problems and students' performance in written responses to wicked problems, but *excessive* cognitive scaffolding with a highly detailed assessment rubric may not

be the ideal form of instructional guidance since it may inadvertently tame the process of addressing wicked problems and thus lead to limited affordances for deep and transferable learning. Instead, more affective and meta-cognitive scaffolding may be needed.

- High levels of detail in an analytic assessment rubric may reduce rubric utility for summative assessment, but increase utility for teacher professional development.
- The ability to integratively address wicked problems covers many and diverse learning outcomes. Therefore, assessing students' ability to integratively address wicked problems may require a combination of different assessment approaches.

Contributions of the research include four theoretical frameworks that describe perspectives, approaches to wicked problems, an integrative approach to wicked problems, and discipline-specific challenges in addressing wicked problems. These frameworks can be used as analytic tools in EER as well as practical tools in engineering education practice. The research also contributes to addressing methodological challenges in EER through examples of qualitative and mixed-methods research studies that link theory and practice. Contributions to practice further include concrete examples of teaching and assessment activities for developing engineering students' ability to integratively address wicked problems, design principles for wicked problem descriptions, and teacher professional development in and through the research activities in the project. Through these contributions, the research can, directly and indirectly, contribute to improving engineering education practice such that it can better prepare students to contribute to sustainability.

REFERENCES

AAES (1994). *The Role of Engineering in Sustainable Development.* American Association of Engineering Societies and the World Engineering Partnership for Sustainable Development, Washington, DC.

ABET (2016). *Criteria for Accrediting Engineering Programs.* Baltimore, Maryland: Engineering Accreditation Commission.

Åkerlind, G. S. (2005). Variation and commonality in phenomenographic research methods. *Higher Education Research & Development, 24* (4), pp. 321–334.

Akkermann, E. (2001). Piaget's Constructivism, Papert's Constructionism: What's the difference? *Constructivism: Uses and Perspectives in Education. 1 & 2*, pp. 85-94. Geneva: Research Center in Education.

An, Y.-J. (2010). Scaffolding Wiki-Based, Ill-Structured Problem Solving in an Online Environment. *MERLOT Journal of Online Learning and Teaching, 6* (4), pp. 723-734.

An, Y.-J., & Cao, L. (2014). Examining the Effects of Metacognitive Scaffolding on Students' Design Problem Solving and Metacognitive Skills in an Online Environment. *MERLOT Journal of Online Learning and Teaching, 10* (4), pp. 552-568.

Andersson, C., Törnberg, A., & Törnberg, P. (2014). Societal systems - Complex or worse? *Futures, 63*, pp. 145-157.

Ardoin, N. M., Clark, C., & Kelsey, E. (2013). An exploration of future trends in environmental education research. *Environmental Education Research, 19* (4), pp. 499-520.

Ash, D., & Levitt, K. (2003). Working within the Zone of Proximal Development: Formative Assessment as Professional Development. *Journal of Science Teacher Education, 14* (1), pp. 23-48.

Atman, C. J., Yasuhara, K., Adams, R. S., Barker, T. J., Turns, J., & Rhone, E. (2008). Breadth in Problem Scoping: a Comparison of Freshman and Senior Engineering Students. *International Journal of Engineering Education, 24* (2), pp. 234-245.

Australian Public Service Commission. (2007). *Tackling wicked problems: A public policy perspective. Commonwealth of Australia.* Retreived from http://www.apsc.gov.au/publications-and-media/archive/publications-archive/tackling-wicked-problems.

Baartman, L. K., Bastiaens, T. J., Kirschner, P. A., & van der Vleuten, C. P. (2007). Evaluating assessment quality in competence-based education: A qualitative comparison of two frameworks. *Educational Research Review, 2*, pp 114–129.

Baillie, C., & Armstrong, R. (2013). Crossing Knowledge Boundaries and Thresholds: Challenging the Dominant Discourse Within Engineering Education. In J. Lucena (Ed.), *Engineering Education for Social Justice: Critical Explorations and Opportunities; Philosophy of Engineering and Technology 10* (pp. 135-152). Dordrecht: Springer Science+Business Media.

Balint, P. J., Stewart, R. E., Desai, A., & Walters, L. (2011). *Wicked Environmental Problems: Managing Uncertainty and Conflict.* Washington, DC: Island Press.

Baltic 21 (2002). *An Agenda 21 for Education in the Baltic Sea Region - Baltic 21E.* Baltic 21 Series No 1/2002.

Barab, S. (2014). Design-Based Research: A Methodological Toolkit for Engineering Change. In R. K. Sawyer, *The Cambridge Handbook of the Learning Sciences* (Vol. 2, pp. 151-170). Cambridge University Press.

Batie, S. S. (2008). Wicked problems and applied economics. *American Journal of Agricultural Economics, 90* (5), pp. 1176-1191.

Bell, P. (2004). On the Theoretical Breadth of Design-Based Research in Education. *Educational Psychologist, 39* (4), pp. 243-253.

Bendixen, L. D., Dunkle, M. E., & Schraw, G. (1994). Epistemological Beliefs and Reflective Judgment. *Psychological Reports, 75*, pp. 1595-1600.

Bengtsson, S., & Kronlid, D. (2016). Infinite Environmental (Ethical) Education. *Symposium paper presented at the European Conference on Educational Research. 30 SES 13: Teaching and Learning in the Face of Wicked Socio-Ecological Problems (Part 1): Exploring Theoretical and Methodological Frameworks.* Dublin: European Educational Research Association.

Berkes, F. (2011). Implementing exosystem-based management: evoluation or revolution? *Fish and Fisheries, 13*, pp. 465-476.

Bernhard, J., & Baillie, C. (2016). Standards for Quality of Research in Engineering Education. *International Journal of Engineering Education, 32* (6), pp. 2378–2394.

Bielefeldt, A. R., & Canney, N. E. (2016). Changes in the Social Responsibility Attitudes of Engineering Students Over Time. *Science and Engineering Ethics, 22*, pp. 1535-1551.

Biesta, G. J. (2010). Why 'What Works' Still Won't Work: From Evidence-Based Education to Value-Based Education. *Studies in Philosophy and Education, 29* (5), pp. 491–503.

Biggs, J. B. (1996). Enhancing teaching through constructive alignment. *Higher Education, 32*, pp. 1-18.

Biggs, J. B. (2014). Constructive alignment in university teaching. *HERDSA Review of Higher Education, 1*, pp. 5-22.

Biglan, A. (1973). The Characteristics of Subject Matter in Different Academic Areas. *Journal of Applied Psychology, 3*, pp. 195-203.

Blackman, T., Elliot, E., Greene, A., Harrington, B., Hunter, D. J., Marks, L., et al. (2006). Performance Assessment and Wicked Problems: The Case of Health Inequalities. *Public Policy and Administration, 21* (2), pp. 66-80.

Bloom, B. S., Engelhart, M. D., Furst, E. J., Hill, W. D., & Krathwohl, D. R. (1956). *Taxonomy of educational objectives: The classification of educational goals. Handbook I: Cognitive domain.* New York: David McKay Company.

Borrego, M., & Bernhard, J. (2011). The emergence of engineering education research as a globally connected field of inquiry. *Journal of Engineering Education, 100*, pp. 14-47.

Borrego, M., Douglas, E. P., & Amelink, C. T. (2009). Quantitative, Qualitative, and Mixed Research Methods in Engineering Education. *Journal of Engineering Education, 98* (1), pp. 53-66.

Borrego, M., Foster, M. J., & Froyd, J. E. (2014). Systematic Literature Reviews in Engineering Education and Other Developing Interdisciplinary Fields. *Journal of Engineering Education, 103* (1), pp. 45-76.

Borrego, M., Streveler, R. A., Miller, R. L., & Smith, K. A. (2008). A New Paradigm for a New Field: Communicating Representations of Engineering Education Research. *Journal of Engineering Education*, pp. 147-162.

Bowden, J. A. (2000). The nature of phenomenographic research. In J. A. Bowden, E. Walsh, & G. Fulcher (Ed.), *Phenomenography.* Melbourne Victoria, Australia: MIT Publishing.

Bozic, M., Escalas Tramullas, M. T., Čizmić, S., & Pavlović, D. S. (2014). Engineering practice: Teaching ill-structured problem solving in an internship-like course. *Proceedings of the 2014 IEEE Global Engineering Education Conference (EDUCON)*, pp. 721-726. Istanbul: IEEE.

Braxton, J. M., & Hargens, L. L. (1996). Variation in academic disciplines: Analytical frameworks and research. In J. S. Smart (Ed.), *Higher Education: Handbook of Theory and Research* (Vol. XI, pp. 1-46). New York: Agathon Press.

Brentano, F. (1874/2009). Psychology from an Empirical Standpoint. (L. L. McAlister, Ed., A. C. Rancurello, D. B. Terrell, & L. L. McAlister, Trans.) London and New York.

Brinkmann, S., & Kvale, S. (2015). *InterViews: learning the craft of qualitative research interviewing* (3rd ed.). Los Angeles, CA: Sage Publications.

Brookhart, S. M., & Chen, F. (2015). The quality and effectiveness of descriptive rubrics. *Educational Review, 67* (3), pp. 343–368.

Brown, V. A., Harris, J. A., & Russell, J. Y. (Eds.). (2010). *Tackling Wicked Problems - Through the Transdisciplinary Imagination.* London: Earthscan.

Brundiers, K., & Wiek, A. (2011). Educating Students in Real-world Sustainability Research: Vision and Implementation. *Innovations in Higher Education, 36*, pp. 107-124.

Buchanan, R. (1992). Wicked Problems in Design Thinking. *Design Issues, 8* (2), pp. 5-21.

Burge, J., & McCall, R. (2014). Diagnosing wicked problems. *Proceedings of the 6th International Conference on Design Computing and Cognition.* New York: Springer.

Carroll, M. S., Blatner, K. A., Cohn, P. J., & Morgan, T. (2007). Managing Fire Danger in the Forests of the US Inland Northwest: A Classic “Wicked Problem” in Public Land Policy. *Journal of Forestry, 105* (5), pp. 239-244.

Case, J. M., & Light, G. (2011). Emerging Methodologies in Engineering Education Research. *Journal of Engineering Education, 100* (1), pp. 186-210.

Cech, E. A. (2012). Great Problems of Grand Challenges: Problematizing Engineering’s Understandings of Its Role in Society. *International Journal of Engineering, Social Justice, and Peace, 1* (2), pp. 85–94.

Cech, E. A. (2014). Culture of Disengagement in Engineering Education? *Science, Technology, & Human Values, 39* (1), pp. 42-72.

Chapin Ill, F. S., Trainor, S. F., Huntington, O., Lovecraft, A. L., Zavaleta, E., Natcher, D. C., et al. (2008). Increasing Wildfire in Alaska’s Boreal Forest: Pathways to Potential Solutions of a Wicked Problem. *BioScience, 58* (6), pp. 531-540.

Cherry, E. (1999). *Programming for Design: From Theory to Practice.* New York: Wiley.

Cho, K., & Jonassen, D. H. (2002). The Effects of Argumentation Scaffolds on Argumentation and Problem Solving. *Educational Technology Research and Development, 50* (3), pp. 5-22.

Cicchetti, D. V. (1994). Guidelines, criteria, and rules of thumb for evaluating normed and standardized assessment instruments in psychology. *Psychological Assessment, 6* (4), pp. 284-290.

Cizek, G. J. (1997). Learning, achievement, and assessment: Constructs at a crossroads. In G. D. Phye (Ed.), *Handbook of classroom assessment: Learning, achievement, and adjustment* (pp. 1-32). San Diego: Academic Press.

Cohen, S. A. (1987). Instructional alignment: searching for a magic bullet. *Educational Researcher, 16* (8), pp. 16-20.

Cohen, L., Manion, L., & Morrison, K. (2011). *Research Methods in Education* (Vol. 7). London and New York: Routledge.

Collier-Reed, B. I., & Ingerman, Å. (2013). Phenomenography: From Critical Aspects to Knowledge Claim. In J. Huisman, & M. Tight, *Theory and Method in Higher Education Research (International Perspectives on Higher Education Research* (Vol. 9). Emerald Group Publishing Limited.

Conklin, J. (2005). Wicked Problems & Social Complexity. In J. Conklin, *Dialogue Mapping: Building Shared Understanding of Wicked Problems.* Wiley.

Conklin, J. (2010). *Wicked Problems & Social Complexity.* Retrieved from http://cognexus.org/wpf/wickedproblems.pdf on 28 11 2013

Connell, S. (1997). Empirical-Analytical Methodological Research in Environmental Education: response to a negative trend in methodological and ideological discussions. *Environmental Education Research, 3* (2), pp. 117-132.

Coyne, R. (2004). Wicked problems revisited. *Design Studies, 26*, pp. 5-17.

Creswell, J. W. (2003). A Framework for Design. In J. W. Creswell, *Research Design: Qualitative, Quantitative and Mixed Methods Approaches* (2nd ed., pp. 3-26). Thousand Oaks, CA: Sage.

Cross, N. G. (1984). *Developments in Design Methodology.* Chichester: Wiley.

DeFries, R., & Nagendra, H. (2017). Ecosystem management as a wicked problem. *Science, 356*, pp. 265–270.

DeGrade, P., & Stahl, L. (1990). *Wicked Problems, Righteous Solutions: A Catalog of Modern Engineering Paradigms.* Upper Saddle River, NJ: Prentice Hall.

Devaney, J., & Spratt, T. (2009). Child abuse as a complex and wicked problem: Reflecting on policy developments in the United Kingdom in working with children and families with multiple problems. *Children and Youth Services Review, 31*, pp. 635-641.

Dillon, J., & Wals, A. E. (2006). On the danger of blurring methods, methodologies and ideologies in environmental education research. *Environmental Education Research, 12* (3-4), pp. 549-558.

Dobson, H. E., & Tomkinson, C. B. (2010). Creating Sustainable Development Change Agents through Problem Based Learning: From rearranging the deckchairs on the Titanic to inciting social revolution – where should student PBL projects be pitched? *Proceedings of the fifth Biannual Conference on Engineering Education in Sustainable Development, September 19-22.* Gothenburg, Sweden.

Dorn, N. (2011). Exploring Responsibility Rationales in Research and Development (R&D). *Science, Technology, & Human Values, 37* (3), pp. 180-209.

Dorst, K. (2004). On the problem of design problems - problem solving and design expertise. *Journal of Design Research, 4* (2).

Dorst, K. (2006). Design Problems and Design Paradoxes. *Design Issues, 22* (3), pp. 4-17.

Douglas, E. P., Koro-Ljungberg, M., McNeill, N. J., Malcolm, Z. T., & Therriault, D. J. (2012). Moving beyond formulas and fixations: solving open-ended engineering problems. *European Journal of Engineering Education, 37* (6), pp. 627-651.

Downey, G. L. (2005). Are Engineers losing Control of Technology? From 'Problem Solving' to 'Problem Definition and Solution' in Engineering Education. *Chemical Engineering Research and Design, 83* (A6), pp. 583–595.

Downey, G. L., & Lucena, J. C. (1997). Engineering Selves: Hiring Into a Contested Field of Education. In G. L. Downey, & J. Dumit (Eds.), *Cyborgs and Citadels: Anthropological Interventions in Emerging Sciences and Technologies* (pp. 117-142). Santa Fe, New Mexico: School of American Research Press.

Duckett, D., Feliciano, D., Martin-Ortega, J., & Munoz-Rojas, J. (2016). Tackling wicked environmental problems: The discourse and its influence on praxis in Scotland. *Landscape and Urban Planning, 154*, pp. 44–56.

EESD Conference Scientific Committee. (2004). *Declaration of Barcelona*. Retrieved 02 28, 2017, from http://eesd15.engineering.ubc.ca/declaration-of-barcelona/.

Farrell, R., & Hooker, C. (2013). Design, science and wicked problems. *Design Studies, 34* (6), pp. 681-705.

Fenner, R. A., Cruickshank, H. J., & Ainger, C. (2014). Sustainability in civil engineering education: why, what, when, where and how. *Proceedings of the Institution of Civil Engineers - Engineering Sustainability*, *167 (ES5)*, pp. 228-237.

Fernandes, R., & Simon, H. A. (1999). A study of how individuals solve complex and ill-structured problems. *Policy Sciences, 32*, pp. 225-245.

Flavell, J. H., Miller, P. H., & Miller, S. A. (2002). *Cognitive development* (Vol. 4). Upper Saddle River, NJ: Prentice Hall.

Fortenberry, N. L. (2006). An Extensive Agenda for Engineering Education Research. *Journal of Engineering Education, 95* (1), pp. 3-5.

Francis, H. (1993). Advancing phenomenography: Questions of method. *Nordisk Pedagogik, 13*, pp. 68-75.

Gambetta, D., & Hertog, S. (2016). *Engineers of Jihad: The Curious Connection between Violent Extremism and Education.* Oxfordshire, United Kingdom: Princeton University Press.

Gardner, A., & Willey, K. (2016). Academic identity reconstruction: the transition of engineering academics to engineering education researchers. *Studies in Higher Education*.

Gibbs, G., & Simpson, C. (2005). Conditions Under Which Assessment Supports Students' Learning. *Learning and Teaching in Higher Education, 1*, pp. 3-31.

Gibson, I. S. (2002). Assessment in Engineering Education - A European Perspective. *International Journal of Engineering Education, 18* (4), pp. 465-471.

Gough, A. (2013). The Emergence of Environmental Education Research: A "History" of the Field. In R. B. Stevenson, M. Brody, J. Dillon, & A. E. Wals (Eds.), *International Handbook of Research on Environmental Education* (pp. 13-22). New York and London: Routledge.

Grand Challenges. (2003-2016). *Grand Challenges: Solving global health and development problems for those most in need*. Retreived from https://grandchallenges.org/#/map on 30 6 2017

Guba, E. G., & Lincoln, Y. S. (1994). Competing Paradigms in Qualitative Research. In N. K. Denzin, & Y. S. Lincoln (Eds.), *Handbook of qualitative research* (pp. 105-117). Thousand Oaks, CA: Sage.

Guerra, A., & Kolmos, A. (2012). Assessing Learning Outcomes and Engineering PBL Project Reports. *Conference proceedings of the SEFI annual conference* (pp. 171-177). Lisbon, Portugal: SEFI.

Haase, S. (2014). Engineering students' sustainability approaches. *European Journal of Engineering Education, 39* (3), pp. 247-271.

Hallgren, K. A. (2012). Computing Inter-Rater Reliability for Observational Data: An Overview and Tutorial. *Tutorials in Quantitative Methods for Psychology, 8* (1), pp. 23-34.

Harris, L. R. (2011). Phenomenographic perspectives on the structure of conceptions: The origins, purposes, strengths, and limitations of the what/how and referential/structural frameworks. *Educational Research Review, 6*, pp. 109–124.

Hart, P. (2006). Requisite variety: the problem with generic guidelines for diverse genres of inquiry. *Environmental Education Research, 12* (3-4), pp. 539-548.

Hart, P., & Nolan, K. (1999). A Critical Analysis of Research in Environmental Education. *Studies in Science Education, 34*, pp. 1-69.

Hattie, J. A. (1987). Identifying the salient facets of a model of student learning: a synthesis of meta-analyses. *International Journal of Educational Research, 11*, pp. 187-212.

Hattie, J. A., & Timperley, H. (2007). The Power of Feedback. *Review of Educational Research, 77* (1), pp. 81-112.

Head, B. (2008). Wicked problems in public policy. *Public Policy, 3* (2), pp. 101-118.

Head, B. (2016). Forty years of wicked problems literature: the need for closer links to policy studies. *Proceedings of Workshop on Wicked Problems.* University of Wageningen, 27-28 October.

Hearnshaw, E. J., Tompkins, J., & Cullen, R. (2011). Addressing the wicked problem of water resource management: An ecosystem services approach. *55th Annual AARES National Conference.* Melbourne, Victoria.

Hess, J. L., Brownell, S. A., & Dale, A. T. (2014). The Wicked Problems in Sustainable Engineering (WPSE) Initiative: Pilot Results of a Cross-Institutional Project-Based Course Offering. *Proceedings of the 121st ASEE Annual Conference & Exhibition, June 15-18.* Indianapolis: ASEE.

Högskoleverket (2010). *Högskoleverkets system för kvalitetsutvärdering 2011-2014.* Retrieved from http://hsv.se/download/18.4afd653a12cabe7775880003715/1022R-system-kvalitetsutv.pdf.

Högskoleverket (2012). *Beslut om riktlinjer för val av mål vid utvärdering av utbildningar som leder till generell examen (omgång 3–2012).* Retrieved from http://www.uk-ambetet.se/download/18.197eccc1140ee238b58b27/12-4013-10-beslut-mal-generell-examen.pdf.

Hmelo-Silver, C. E., Duncan, R. G., & Chinn, C. A. (2007). Scaffolding and Achievement in Problem-Based and Inquiry Learning: A Response to Kirschner, Sweller, and Clark. *Educational Psychologist, 42* (2), pp. 99-107.

Hsieh, H.-F., & Shannon, S. E. (2005). Three Approaches to Qualitative Content Analysis. *Qualitative Health Research, 15*, pp. 1277-1288.

Hughes, T. P., Huang, H., & Young, M. A. (2012). The Wicked Problem of China's Disappearing Coral Reefs. *Conservation Biology, 27* (2), pp. 261-269.

Jamison, A., Hyldgaard Christensen, S., & Botin, L. (2011). *A Hybrid Imagination: Science and Technology in Cultural Perspective* (Synthesis Lectures on engineers, technology, and society (Lecture #12). Morgan and Claypool.

Jawitz, J., & Case, J. (2009). Communicating your findings in engineering education: the value of making your theoretical perspective explicit. *European Journal of Engineering Education, 34*, pp. 149-154.

Jehng, J. J. (1993). Schooling and Students' Epistemological Beliefs about Learning. *Contemporary Educational Psychology, 18*, pp. 23-35.

Jentoft, S., & Chuenpagdee, R. (2009). Fisheries and coastal governance as a wicked problem. *Marine Policy, 33*, pp. 553-560.

Jervis, L. M., & Jervis, L. (2005). What is the Constructivism in Constructive Alignment? *Bioscience Education, 6* (1), pp. 1-14.

Jesiek, B. K., Borrego, M., & Beddoes, K. (2010). Advancing global capacity for engineering education research: relating research to practice, policy and industry. *European Journal of Engineering Education, 35* (2), pp. 117-134.

Jesiek, B. K., Newswander, L. K., & Borrego, M. (2009). Engineering Education Research: Discipline, Community, or Field? *Journal of Engineering Education, 98* (1), pp. 39-52.

Johansson, A. W., & Lindhult, E. (2008). Emancipation of workability? Critical versus pragmatic scientific orientation in action research. *Action Research, 6* (1), pp. 95-115.

Jonassen, D. H. (1997). Instructional Design Models for Well-Structured and Ill-Structured Problem-Solving Learning Outcomes. *Educational Technology Research and Development, 45* (1), pp. 65-94.

Jonassen, D. H. (2000). Toward a Design Theory of Problem Solving. *Educational Technology and Research Development, 48* (4), pp. 63-85.

Jonassen, D. H., Strobel, J., & Beng Lee, C. (2006). Everyday Problem Solving in Engineering: Lessons for Engineering Educators. *Journal of Engineering Education, 92* (2), pp. 139-151.

Jonsson, A., & Svingby, G. (2007). The use of scoring rubrics: Reliability, validity and educational consequences. *Educational Research Review, 2*, pp. 130-144.

Kabo, J. (2010). Seeing through the lens of social justice: A threshold for engineering. Kingston, Canada: Queen's University.

Kahneman, D. (2011). *Thinking, Fast and Slow.* New York: Farrar Straus Giroux.

Kane, M. T. (1992). An Argument-Based Approach to Validity. *Psychological Bulletin, 112* (3), pp. 527-535.

Kanematsu, H., & Barry, D. M. (2016). Eco Car Problem-Based Learning Project in Second Life. i H. Kanematsu, & D. M. Barry, *STEM and ICT Education in Intelligent Environments*. Vol. 91, pp. 159-165. Switzerland: Springer International Publishing.

Kates, R. W., Parris, T., & Leiserowitz, A. A. (2005). What is sustainable development? *Environment, Science and Policy for Sustainable Development, 47* (3), pp. 8–21.

King, P. M., & Kitchener, K. S. (1994). *Developing Reflective Judgment.* San Francisco, CA: Jossey-Bass.

Kinnunen, P., & Simon, B. (2012). Phenomenography and grounded theory as research methods in computing education research field. *Computer Science Education, 22* (2), pp. 199-218.

Kirschner, P. A., Sweller, J., & Clark, R. E. (2006). Why Minimal Guidance During Instruction Does Not Work: An Analysis of the Failure of Constructivist, Discovery, Problem-Based, Experiential, and Inquiry-Based Teaching. *Educational Psychologist, 41* (2), pp. 75-86.

Kitchener, K. S. (1983). Cognition, Metacognition and Epistemic Cognition: A Three-Level Model of Cognitive Processing. *Human Development, 26*, pp. 222-232.

Koh, K. H. (2011). Improving teachers' assessment literacy through professional development. *Teaching Education, 22* (3), pp. 255-276.

Koro-Ljungberg, M., & Douglas, E. P. (2008). State of Qualitative Research in Engineering Education: Meta-Analysis of JEE Articles, 2005-2006. *Journal of Engineering Education, 97* (2), pp. 163-175.

Kovacic, S. F., & Sousa-Poza, A. (2013). *Managing and Engineering in Complex Situations.* New York: Springer.

Krasny, M. E. (2013). Tales of a Transdisciplinary Scholar. In M. E. Krasny, & J. Dillon, *Trading Zones in Environmental Education* (Vol. 1). New York: Peter Lang Publishing Group.

Krathwohl, D. R. (2002). A Revision of Bloom's Taxonomy: An Overview. *Theory Into Practice, 41* (4), pp. 212-218.

Krefting, L. (1991). Rigor in Qualitative Research: The Assessment of Trustworthiness. *The American Journal of Occupational Therapy, 45* (5), pp. 214-222.

Kuhn, T. (1962/1970). *The structure of scientific revolutions* (2nd ed.). Chicago: The University of Chicago Press.

Kunz, W., & Rittel, H. W. (July 1970). Issues as Elements of Information Systems. *Working Paper No. 131*, pp. 1-9.

Lane, S., & Tierney, S. T. (2008). Performance Assessment. In T. L. Good (Ed.), *21st Century Education: A Reference Handbook* (Vol. 1, pp. 461–470). Los Angeles, CA: SAGE.

Lattuca, L. R., & Stark, J. S. (1995). Modifying the major: Discretionary thoughts from ten disciplines. *The Review of Higher Education, 18* (3), pp. 315-344.

Levin, K., Cashore, B., Bernstein, S., & Auld, G. (2009). Playing it Forward: Path Dependency, Progressive Incrementalism, and the “Super Wicked” Problem of Global Climate Change. *IOP Conference Series: Earth and Environmental Science*, *6*, p. 502002.

Liu, S., Lin, C., & Tsai, C. (2011). College Students’ Scientific Epistemological Views and Thinking Patterns in Socioscientific Decision Making. *Science Education, 95* (3), pp. 497–517.

Liu, C. H., & Matthews, R. (2005). Vygotsky’s philosophy: Constructivism and its criticisms examined. *International Education Journal, 6* (3), pp. 386-399.

Lohmann, J. R. (2008). Global Engineering Excellence: The Role of Educational Research and Development. *Revista de Ensino de Engenharia, 27* (3), pp. 33-44.

Lönngren, J. (2014). *Engineering Students' Ways of Relating to Wicked Sustainability Problems.* Chalmers University of Technology, Department of Applied IT. Gothenburg: Chalmers.

Lönngren, J., Adawi, T., & Svanström, M. (2017). Assessing engineering students’ written responses to wicked problems: Reliability, validity and utility of an analytic rubric. *Proceedings of the 7th Research on Engineering Education Symposium, 6-8 July*. Bogota, Colombia.

Lönngren, J., Ahrens, A., Deppert, K., Hammarin, G., & Nilsson, E. (2010). Sustainable Development in Nano-Perspectives – An Innovative Student Initiative. *Proceedings of fifth Engineering Education in Sustainable Development Conference, 19-22 September.* Gothenburg, Sweden.

Lönngren, J., Ingerman, Å., & Svanström, M. (2017). Avoid, Control, Succumb, or Balance: Engineering Students’ Approaches to a Wicked Sustainability Problem. *Research in Science Education, 47*, pp. 805-831.

Lönngren, J., Jacobsson, D., Mårsell, E., & Nilsson, E. (2012). Breaking Catch-22 of Engineering Education for Sustainable Development: An Example of Parallell Learning of Teachers and Students. *Improving Student Learning Through Research and Scholarship: 20 Years of ISL, 29-31 August* (p. 28). Lund: Oxford Centre for Staff and Learning Development.

Lönngren, J., & Svanström, M. (2015). Assessing "Wicked Sustainability Problem"-Literacy in Engineering Education. *Proceedings of the 122nd ASEE Annual Conference & Exhibition.* Seattle: American Society for Engineering Education.

Lönngren, J. & Svanström, M. (2016). Systems thinking for dealing with wicked sustainability problems: Beyond functionalist approaches. In W.L Filho & S. Nesbit (eds.) *New Developments in Engineering Education for Sustainable Development*. First edition. World Sustainability Series. Springer International Publishing.

Lönngren, J., Svanström, M., Ingerman, Å., & Holmberg, J. (2016). Dealing with the multidimensionality of sustainability through the use of multiple perspectives - a theoretical framework. *European Journal of Engineering Education, 41* (3), pp. 342-352.

Lundqvist, U., & Svanström, M. (2008). Inventory of content in basic courses in environment and sustainable development at Chalmers University of Technology in Sweden. *European Journal of Engineering Education, 33* (3), pp. 355-364.

Malmi, L., Adawi, T., Curmi, R., de Graaf, E., Duffy, G., Kautz, C., et al. (2016). How authors did it - a methodological analysis of recent engineering education research papers in the European Journal of Engineering Education. *European Journal of Engineering Education Research*.

Marton, F. (1979). Skill as an aspect of knowledge. *The Journal of Higher Education, 50* (5), pp. 602-614.

Marton, F. (2015). *Necessary Conditions of Learning.* New York: Routledge.

Marton, F., & Booth, S. (1997). *Learning and Awareness.* Mahwah, NJ: Lawrence Erlbaum Associates.

Marton, F., & Pong, W. Y. (2005). On the unit of description in phenomenography. *Higher Education Research & Development, 24* (4), pp. 335–348.

Marton, F., & Säljö, R. (1976). On qualitative differences in learning: I – Outcome and process. *British Journal of Educational Psychology, 46*, pp. 4-11.

Mayring, P. (2000). Qualitative Content Analysis. *Forum: Qualitative Social Research, 1* (2), p. Art. 20.

McCall, R., & Burge, J. (2016). Untangling wicked problems. *Artificial Intelligence for Engineering Design, Analysis and Manufacturing, 30*, pp. 200–210.

McKenney, S., & Reeves, T. C. (2012). *Conducting educational design research.* London: Routledge.

McKenzie, M. (2016-08-17). Personal communication.

Menkhaus, K. J. (2010). State Fragility as a Wicked Problem. *Prism, 1* (2), pp. 85-100.

Micari, M., Light, G., Calkins, S., & Streitwieser, B. (2007). Assessment Beyond Performance Phenomenography in Educational Evaluation. *American Journal of Evaluation, 28* (4), pp. 458-476.

Ministers of Education (2000). *Haga Declaration.* The Ministers of Education of Denmark, Estonia, Finland, Germany, Latvia, Lithuania, Norway, Poland, The Russion Federation and Sweden or their Representatives. Stockholm, Sweden: 1st CBSS Ministers of Education meeting.

Ministry of Education and Research, Sweden (1992). *The Swedish Higher Education Act.*

Ministry of Education and Research, Sweden (1993). *Higher Education Ordninance.*

Moskal, B. M., & Brown, S. (2014). Results of the JEE Readership Survey. *Journal of Engineering Education, 103* (4), pp. 517-524.

Moskal, B. M. (2000). Scoring Rubrics: What, When and How? *Practical Assessment, Research & Evaluation, 7* (3), pp. 1-5.

Mulder, K. F., Segalàs-Coral, J., & Ferrer-Balas, D. (2010). How to Educate Engineers for/in Sustainable Development? 10 years of discussion, what challenges remain? *Engineering Education in Sustainable Development, September 19-22.* Gothenburg, Sweden.

Nagel, J. K., Pierrakos, O., Zilberberg, A., & McVay, S. (2012). Understanding industry experiences: From problem solving to engineering students' learning gains. *2012 Frontiers in Education Conference Proceedings*, pp. 1-6. Seattle, WA: FIE.

Nelson, W. A. (2003). Problem Solving Through Design. *New Directions for Teaching and Learning, 95*, pp. 39-44.

Nesbit, S. (28 04 2017). personal communication.

Noble, D., & Rittel, H. W. (1988). Issue Based Information Systems for Design. *Computing in Design Education [ACADIA Conference Proceedings], 28-30 October*, (pp. 275-286). Ann Arbor, United States.

Noffke, S. E. (1997). Chapter 6: Professional, Personal, and Political Dimensions of Action Research. *Review of Research in Education, 22*, pp. 305-343.

Nordegraaf, M., Douglas, S., Geuijen, K., & van der Steen, M. (2016). Weaknesses of Wickedness: A Critical Perspective on Wickedness Theory. *Proceedings of Workshop on Wicked Problems, 27-28 October.* University of Wageningen.

Norman, D. A. (2013). *The Design of Everyday Things: Revised and Expanded Edition.* New York: Basic Books.

Norton, B. G. (2012). The Ways of Wickedness: Analyzing Messiness with Messy Tools. *Journal of Agricultural Environmental Ethics, 25*, pp. 447–465.

Novak, J. D., & Gowin, D. N. (1984). *Learning How to Learn.* Cambridge, UK: Cambridge University Press.

Osorio, N. L. (2005). What every engineer should know about engineering education. *Proceedings of the IL/IN Sectional Conference of the Americal Association for Engineering Education.* DeKalb, Illinois : Northern Illinois University.

Palmer, J. A. (1998). *Environmental Education in the 21st century: Theory, practice, progress and promise.* London: Routledge.

Palmer, J. Al (2012). Risk governance in an age of wicked problems: lessons from the European approach to indirect land-use change. *Journal of Risk Research, 15* (5), pp. 495-513.

Panadero, E., & Jonsson, A. (2013). The use of scoring rubrics for formative assessment purposes revisited: A review. *Educational Research Review, 9*, pp. 129–144.

Pang, M. F. (2003). Two Faces of Variation: on continuity in the phenomenographic movement. *Scandinavian Journal of Educational Research, 47* (2), pp. 145-156.

Pang, M. F., & Ling, L. M. (2012). Learning study: helping teachers to use theory, develop professionally, and produce new knowledge to be shared. *Instructional Science, 40*, pp. 589-606.

Pang, M. F., & Marton, F. (2003). Beyond “lesson study”: Comparing two ways of facilitating the grasp of some economic concepts. *Instructional Science, 31*, pp. 175–194.

Paretti, M. C., & McNair, L. D. (2012). Analyzing the intersections of institutional and discourse identifies in engineering work at the local level. *Engineering Studies, 4* (1), pp. 55-78.

Parr, J. M., Glasswell, K., & Aikman, M. (2007). Supporting Teacher Learning and Informed Practice in Writing through Assessment Tools for Teaching and Learning. *Asia-Pacific Journal of Teacher Education, 35* (1), pp. 69-87.

Paulsen, M. B., & Wells, C. T. (1998). Domain differences in the epistemological beliefs of college students. *Research in Higher Education, 39* (4), pp. 365-384.

Pawley, A. L. (2009). Universalized Narratives: Patterns in How Faculty Members Define “Engineering”. *Journal of Engineering Education, 98* (4), pp. 309–319.

Pereira, D., Flores, M. A., & Niklasson, L. (2016). Assessment revisited: a review of research in Assessment and Evaluation in Higher Education. *Assessment & Evaluation in Higher Education, 41* (7), pp. 1008–1032.

Raisio, H. (2010). *Embracing the wickedness of health care: Essays on reforms, wicked problems and public deliberation.* Finland: Acta Wasaensia, 228, Universitas Wasaensis.

Reason, P. (2006). Choice and Quality in Action Research Practice. *Journal of Management Inquiry, 15* (2), pp. 187-203.

Reason, P., & Bradbury, H. (Ed.). (2001). *Handbook of action research: Participative inquiry and practice.* London: Sage.

Reddy, Y., & Andrade, H. (2010). A review of rubric use in higher education. *Assessment & Evaluation in Higher Education, 35* (4), pp. 435–448.

Reid, A. (2014). *Environmental Education Research general guidlines.* Retrieved 02 16, 2017, from http://www.tandf.co.uk/journals/pdf/ceer-general-guidelines.pdf

Reid, A., & Scott, W. (2006). Researching education and the environment: retrospect and prospect. *Environmental Education Research, 12* (3-4), pp. 571-587.

Richardson, J. T. (1999). The Concepts and Methods of Phenomenographic Research. *Review of Educational Research, 69* (1), pp. 53-82.

Richter, D. M., & Paretti, M. C. (2009). Identifying barriers to and outcomes of interdisciplinarity in the engineering classroom. *European Journal of Engineering Education, 34* (1), pp. 29-45.

Ricker-Gilbert, J., Jayne, T., & Shively, G. (2013). Addressing the “Wicked Problem” of Input Subsidy Programs in Africa. *Applied Economic Perspectives and Policy, 35* (2), pp. 322-340.

Rickinson, M. (2001). Learners and Learning in Environmental Education: a critical review of the evidence. *Environmental Education Research, 7* (3), pp. 207-320.

Riley, D. (2008). *Engineering and Social Justice.* Morgan & Claypool.

Riley, D. (2012). *Engineering Thermodynamics and 21st Century Energy Problems: A textbook companion for student engagement.* Morgan & Claypool.

Rith, C., & Dubberly, H. (2007). Why Horst W.J. Rittel matters. *Design Issues, 22* (4), pp. 1-20.

Rittel, H. W., & Webber, M. W. (1973). Dilemmas in a General Theory of Planning. *Policy Sciences, 4*, pp. 155-169.

Roberts, N. (2000). Wicked Problems and Network Approaches to Resolution. *International Public Management Review, 1* (1), pp. 1-19.

Robottom, I., & Hart, P. (1995). Behaviourist EE research: environmentalism and individualism. *Journal of Environmental Education, 26* (2), pp. 5-9.

Roll, I., Holmes, N. G., Day, J., & Bonn, D. (2012). Evaluating metacognitive scaffolding in Guided Invention Activities. *Instructional Science, 40*, pp. 691-710.

Royal Academy of Engineering (2005). *Engineering for sustainable development: guiding principles.* London, UK.

Sauvé, L. (2005). Currents in Environmental Education: Mapping a Complex and Evolving Pedagogical Field. *Canadian Journal of Environmental Education, 10*, pp. 11-37.

Schön, D. A. (1987). *Educating the Reflective Practitioner: Toward a New Design for Teaching and Learning in the Professions* (Vol. 1). Jossey-Bass.

Schommer-Aikins, M., & Hutter, R. (2002). Epistemological beliefs and thinking about everyday controversial issues. *The Journal of Psychology, 136* (1), pp. 5-20.

Schraw, G., Dunkle, M. E., & Bendixen, L. D. (1995). Cognitive Processes in Well-Defined and Ill-Defined Problem Solving. *Applied Cognitive Psychology, 9*, pp. 523-538.

SCNEER (2006). The Research Agenda for the New Discipline of Engineering Education. *Journal of Engineering Education, 95* (4), pp. 259-261.

Scouller, K. (1998). The influence of assessment method on students' learning approaches: Multiple choice question examination versus assignment essay. *Higher Education, 35*, pp. 453-472.

Seager, T., Selinger, E., & Wiek, A. (2012). Sustainable Engineering Science for Resolving Wicked Problems. *Journal of Agricultural Environmental Ethics, 25*, pp. 467–484.

Seely, B. E. (2005). Patterns in the History of Engineering Education Reform: A Brief Essay. In *Educating the Engineer of 2020: Adapting Engineering Education to the New Century.*

Sharman, M. (2009). *The Wicked Problem of Biodiversity. Targets of sustainability - that is the question.* Retreived from Biodiversity, sustainability and targets: http://www.azoresbioportal.angra.uac.pt/files/noticiasThewickedproblemofbiodiversity2.pdf on 20 03 2013

Shin, N., Jonassen, D. H., & McGee, S. (2003). Predictors of well-structured and ill-structured problem solving in an astronomy simulation. *Journal of Research in Science Teaching, 40* (1), pp. 6–33.

Shindler, B. A., & Cramer, L. A. (1999). Shifting public values for forest management: Making sense of wicked problems. *Western Journal of Applied Forestry, 14* (1), pp. 28-34.

Shuell, T. J. (1986). Cognitive conceptions of learning. *Review of Educational Research, 56*, pp. 411-436.

Simon, H. A. (1981). *The sciences of the artificial* (Vol. 2). Cambridge, MA: MIT Press.

Sin, S. (2010). Considerations of Quality in Phenomenographic Research. *International Journal of Qualitative Methods, 9* (4), pp. 305-319.

Slife, B. (1987). The Perils of Eclecticism as Therapeutic Orientation. *Theoretical and Philosophical Psychology, 7* (2), pp. 94-103.

Stevenson, R. B., Wals, A. E., Dillon, J., & Brody, M. (2013). Introduction: An Orientation to Environmental Education and the Handbook. In R. B. Stevenson, M. Brody, J. Dillon, & A. E. Wals (Eds.), *International Handbook of Research on Environmental Education* (pp. 1-6). New York and London: Routledge.

Stiggins, R. J. (1995). Assessment literacy for the 21st century. *Phi Delta Kappan, 77* (3), pp. 238-245.

Stone, C. A. (1998). The Metaphor of Scaffolding: Its Utility for the Field of Learning Disabilities. *Journal of Learning Disabilities, 31* (4), pp. 344-364.

Sund, P., & Greve Lysgaard, J. (2013). Reclaim "Education" in Environmental and Sustainability Education Research. *Sustainability, 5*, pp. 1598-1616.

Svanström, M., Lozano-García, F. J., & Rowe, D. (2008). Learning outcomes for sustainable development in higher education. *International Journal of Sustainability in Higher Education, 9* (3), pp. 339-351.

Svensson, L. (1976). *Study skill and learning.* Gothenburg: Acta Universitatis Gothoburgensis.

Svensson, L. (1997). Theoretical Foundations of Phenomenography. *Higher Education Research & Development, 16* (2), pp. 159-171.

Svensson, L. (2016). Towards an Integration of Research on Teaching and Learning. *Scandinavian Journal of Educational Research, 60* (3), pp. 272-285.

Tatham, P., & Houghton, L. (2011). The wicked problem of humanitarian logistics and disaster relief aid. *Journal of Humanitarian Logistics and Supply Chain Management, 1* (1), pp. 15-31.

The EESD Observatory (2009). *The EESD Observatory: Status of engineering education for sustainable development in European higher education.* The Alliance for Global Sustainability. Gothenburg: Informtrycket.

Thompson, G. (2002). Status and Prospects of Sustainable Engineering Education in Some American Universities. *Proceedings of the Engineering Education in Sustainable Development Conference, 24-25 October*, (pp. 1-8). Delft.

Tian, X. (2007). Do Assessment Methods Matter? A Sensitivity Test. *Assessment & Evaluation in Higher Education, 32* (4), pp. 387-401.

Turnbull, N., & Hoppe, R. (2017). Problematizing 'wickedness': a critique of the wicked problems concept, from philosophy to practice and beyond. *Proceedings of the International Research Society for Public Management Conference, April 19-21*, (pp. 1-31). Corvinus University Budapest.

Tyler, R. W. (1949). *Basic principles of curriculum and instruction.* Chicago: University of Chicago Press.

UNCED (1992a). *Agenda 21.* United Nations Conference on Environment & Development, Rio de Janeiro, Brazil, 3 to 14 June 1992. New York: United Nations Sustainable Development.

UNCED (1992b). *Agenda 21, Chapter 36, Promoting Education, Public Awareness and Training.* United Nations Conference on Environment & Development, Rio de Janeiro, Brazil, 3 to 14 June 1992. New York: United Nations Sustainable Development.

UNCED (1992c). *The Rio Declaration on Environment and Development.* United Nations Conference on Environment & Development, Rio de Janeiro, Brazil, 3-14 June. New York: United Nations Sustainable Development.

UNESCO (1978). *Tbilisi Declaration.* Intergovernmental conference on environmental education. Paris: UNESCO.

UNESCO (2005). *United Nations Decade of Education for Sustainable Development (2005-2014).* Paris: UNESCO.

UNESCO (2010). *Teaching and Learning for a Sustainable Future: a multimedia teacher education programme.* Retrieved from http://www.unesco.org/education/tlsf/mods/theme_d/mod24.html.

UNESCO (2014). *UNESCO Roadmap for Implementing the Global Action Programme on Education for Sustainable Development.* Paris: UNESCO.

University of Michigan (2016). *A Taxonomy for the Field of Engineering Education Research.* Retrieved 03 17, 2017, from EER Taxonomy Version 1.0: http://taxonomy.engin.umich.edu/taxonomy/eer-taxonomy-version-current/.

van de Pol, J., Volman, M., & Beishuizen, J. (2010). Scaffolding in Teacher–Student Interaction: A Decade of Research. *Educational Psychology Review, 22*, pp. 271–296.

van den Akker, J., Gravemejer, K., McKenney, S., & Nieveen, N. (2006). *Educational design research.* London, New York: Routledge.

Voss, J. F. (1987). Learning and Transfer i Subject-Matter Learning: A Problem-Solving Model. *International Journal of Educational Research, 11* (6), pp. 607-622.

Voss, J. F., Greene, T. R., Post, T. A., & Penner, B. C. (1983). Problem-Solving Skills in the Social Sciences. *The Psychology of Learning and Motivation, 17*, pp. 165-213.

Vygotsky, L. S. (1987). Thought and Language. In L. S. Vygotsky, R. Rieber, & A. Carton (Eds.), *Collected works* (N. Minick, Trans., Vol. 1, pp. 39–285). New York: Plenum.

Wagner, J. (1993). Ignorance in Educational Research - Or, How Can You Not Know? *Educational Researcher, 22* (5), pp. 15-23.

Wald, H. D., Borkan, J. M., Scott Taylor, J., Anthony, D., & Reis, S. P. (2012). Fostering and Evaluating Reflective Capacity in Medical Education: Developing the REFLECT Rubric for Assessing Reflective Writing. *Academic Medicine: Journal of the Association of American Medical Colleges, 87* (1), pp. 41-50.

Wals, A. E., & Blaze Corcoran, P. (2006). Sustainability as an outcome of transformative learning. Education for Sustainable Development in Action. In J. Holmberg, & B. E. Samuelsson, *Drivers and Barriers for Implementing Sustainable Development in Higher Education.* Paris: UNESCO.

Wankat, P. S., Felder, R. M., Smith, K. A., & Oreovicz, F. S. (2002). The Scholarship of Teaching and Learning in Engineering. In M. T. Huber, & S. Morreale, *Disciplinary Styles in the Scholarship of Teaching and Learning: Exploring Common Ground.* Washington: AAHE/Carnegie Foundation for the Advancement of Teaching.

Weber, R. P. (1990). *Basic content analysis.* Beverly Hills, CA: Sage.

Whyte, K. P., & Thompson, P. B. (2012). Ideas for how to take wicked problems seriously. *Journal of Agricultural and Environmental Ethics, 25* (4), pp. 441-445.

Wiek, A., Withycombe, L., & Redman, L. (2011). Key competencies in sustainability: a reference framework for academic program development. *Integrated Research System for Sustainability Science, 6*, pp. 203–218.

Wise, A. F., & O'Neill, K. (2009). Beyond more versus less: A reframing of the debate on instructional guidance. In S. Tobias, & T. M. Duffy (Eds.), *Constructivist instruction: Success or failure?* (pp. 82-105). New York, NY, US: Routledge/Taylor & Francis Group.

Wisnioski, M. H. (2012). *Engineers for change: competing visions for technology in 1960s America.* Cambridge, Massachusetts: MIT Press.

Wood, D., Bruner, J. S., & Ross, G. (1976). The role of tutoring in problem solving. *Journal of child psychology and psychiatry, and allied disciplines, 17* (2), pp. 89-100.

Wragg, T. (2013). *An Introduction to Classroom Observation (Classic Edition).* Oxon: Routledge.

Xiang, W.-N. (2013). Working with wicked problems in socio-ecological systems: Awareness, acceptance, and adaptation. *Landscape and Urban Planning, 110*, pp. 1-4.

Yanchar, S. C., & Williams, D. D. (2006). Reconsidering the Compatibility Thesis and Eclecticism: Five Proposed Guidelines for Method Use. *Educational Researcher, 35* (9), pp. 3-12.

Zohar, A., & Barzilai, S. (2013). A review of research on metacognition in science education: current and future directions. *Studies in Science Education, 49* (2), pp. 121-169