Design research with a focus on learning processes: an overview on achievements and challenges

Susanne Prediger · Keno Gravemeijer · Jere Confrey


Abstract. Design research continues to gain prominence as a significant methodology in the mathematics education research community. This overview summarizes the origins and the current state of design research practices focusing on methodological requirements and processes of theorizing. While recognizing the rich variations in the foci and scale of design research, it also emphasizes the fundamental core of understanding and investigating learning processes. That is why the article distinguishes two archetypes of design research, one being focused on curriculum innovations, one being focused on developing theories on the learning processes, which is the main focus of the thematic issue. For deepening the methodological discussion on design research it is worth to distinguish aims and quality criteria along the archetypes and elaborate achievement and challenges for the future.

1. Introduction

Design research is an evolving methodology with substantial variation among concrete approaches, and yet with common features. We briefly sketch the origins of design research (in Section 2.1) and identify dimensions of variability (Section 2.2). Distinguishing two archetypes, ‘design research with a focus on curriculum products’, and ‘design research with a focus on learning processes’ (in Section 2.3), this thematic issue and the overview on achievements and challenges mainly concerns the elaboration of the latter archetype for mathematics education, where it is the dominant approach.

We first characterize this approach by grounding it in background theories that require an active role of students in constructing their own knowledge (Section 3). This is followed by a methodological reflection on this type of design research (Section 4), which takes its starting point in some of the critique on design research. As generating theories is crucial for design research with a focus on learning processes, we discuss the variety of theories that play a role, and the process of developing theories in design research. We close with an overview of how those considerations play out, elucidated with reference to the contributions to this thematic issue (Section 5).
2. **Design research as a research program of increasing importance in general education and mathematics education**

2.1 Different origins of design research

Design research emerged in different places, under different names (e.g., “design research”, “design-based research”, “design experiments”, “design theories”, “educational design research”, and “developmental research”), and – what is more important – to serve different needs, resulting in different characteristics. In this article we use the term “design research” as a generic, symmetric term, which does not prioritize research over design or vice versa, even if the variety of terms still exists today (see for instance, Kelly, Lesh, and Baek 2008; van den Akker, Gravemeijer, McKenney, and Nieveen 2006). We sketch some different lines of origin in order to make understandable different priorities.

One origin of design research goes back to the *tradition of curriculum innovation*: In the 1960s and 1970s, the so-called RDD model (research-development-dissemination-model) was the primary model for curricular innovation. According to this model, fundamental research precedes and informs the development of curricula, which then are disseminated and implemented. However, in reality in schools, ready-made curricula were not implemented as intended (Fullan and Pomfret 1977). Furthermore, it became evident that the so-called first-generation instructional design theories (Merrill, Li, and Jones 1990) of that time presupposed the use of scientific knowledge that was not available to support innovative curriculum development projects.

Against this background, design research emerged as an alternative. Design research combines instructional design and educational research. Instead of executing those activities in sequence, design-researchers perform both simultaneously, often while involving teachers from the beginning.

![Fig. 1 Former linear causal models relating basic research and innovation (Stokes 1997, p. 10)](image)

This shift may be clarified with the perspective of Stokes (1997) on the relation between theory and practice in general. When analyzing the relation between technology and fundamental research in science from a historical perspective, Stokes found that the standard model of the relation between basic research and practical use is too narrow to reflect what happened in reality. The standard model suggests that all innovations start with pure basic research that informs applied research which is used in development and finally translates into production and operation (Fig. 1). He came to the conclusion that the conduct of science is better described by a model that involves two dimensions, “quest for understanding”, and “considerations of use”. Those together constitute a matrix with four quadrants (Fig. 2).
Stokes’ model may be used to describe the situation in educational research, and many authors have advocated that besides Bohr’s and Edison’s quadrant, also Pasteur’s quadrant of use-inspired basic research must be filled (Burkhardt and Schoenfeld 2003; Reeves 2000). This is the quadrant where design research fits, as it pairs a quest for understanding to considerations of use.

The aforementioned first-generation instructional design theories were tailored to the linear model and assumed fixed learning goals, ample academic knowledge, and directly applicable general theories. Design research ideas emerged in situations where first-generation instructional theories fell short; they are tailored to more innovative learning environments, where learning goals are to be refined in the process, little academic knowledge is available, and general theories do not yet offer much help (cf. Section 4.1 for further discussion).

Historically, the ideas of design research often emerged when researchers involved in instructional design projects felt the need for strengthening the scientific basis of their projects, while existing knowledge to build on was very scarce or completely lacking (cf. Confrey 2006 for a historical reconstruction). In these processes, the awareness emerged that carefully conducted and analyzed design projects also could generate theory. This was for example the case in the Madison-Project (Romberg 1973), UCSMP (Usiskin 1986), the work of Brown (1992), who coined the term “design experiments”, Collins (1992), and others in the US.

Also in Europe, various traditions arose: As Margolin and Drijvers (2015) describe, the research program of “Didactical Engineering” came to fruition in France, which was centered around the Theory of Didactical Situations (Artigue 1992, 2015; Brousseau 1983, 1997). In the Netherlands, there were two different strands of “developmental research”, one linked to the development of RME (realistic mathematics education) at Utrecht University (see for instance Gravemeijer 1994, 1998; Streefland 1991), and one that had its roots in general educational technology at the University of Twente (van den

---

**Fig. 2** Stokes’ (1997, p. 73) model of four quadrants with two dimensions of goals

<table>
<thead>
<tr>
<th>Quest for fundamental understanding</th>
<th>Consideration of use?</th>
</tr>
</thead>
<tbody>
<tr>
<td>Yes</td>
<td>Use-inspired basic research (Pasteur)</td>
</tr>
<tr>
<td>No</td>
<td>Pure applied research (Edison)</td>
</tr>
<tr>
<td>Pure basic research (Bohr)</td>
<td></td>
</tr>
</tbody>
</table>

---

Research is inspired by

- No
- Yes

Consideration of use?

- No
- Yes
Akker 1999). Although the Dutch discussion started in the late 1980s (cf. Gravemeijer and Koster 1988 for early Dutch proceedings), it only reached a wider attention in the mid-1990s (Gravemeijer 1994) and changed the name to the internationally more common “design research” (Gravemeijer and Cobb 2006). In Germany, Wittmann (1995), one of the authors of the primary curriculum Mathe 2000, advocated for a “design science” by emphasizing the need for didactical engineering based on both, insights into student thinking and into mathematical structures relevant for learning. In the last years, the Dortmund tradition was elaborated by a systematic emphasis on empirical investigation and theory building in “topic-specific didactical design research” (Prediger et al. 2012; Prediger and Zwetzschler 2013).

Design research further has its roots in engineering (Edelson 2002), and in one-on-one teaching experiments (Cobb and Steffe 1983; Confrey and Lachance 2000; Steffe 1983; Thompson 1979). According to Edelson, design research expands on ordinary design—which uses knowledge on design procedures, problem analyses, and design solutions to create a successful design product—by adding the goal of developing useful, generalizable theories. In one-on-one teaching experiments, researchers tried to come to understand how students think and provide instructional tasks to elicit student thinking as the experiment evolves. This orientation on student thinking subsequently fed into design research focusing on student thinking (cf. Confrey 2006).

Design research became more widely known in 2003 and 2004 with special issues of the Educational Researcher (Cobb, Confrey, diSessa, Lehrer, and Schauble 2003; Design-based research collective 2003) and the Journal of the Learning Sciences (Barab and Squire 2004). Especially the article by Cobb et al. (2003) became a reference point, thanks to their exposition on general characteristics of design research (which they refer to as design experiments). They discern the following five common characteristics: (1) interventionist, (2) theory generative, (3) prospective and reflective, (4) iterative, and (5) ecologically valid and practice-oriented.

Although a more elaborate discussion of those characteristics can be found in the contribution of Confrey and Maloney (2015), we briefly explain these common characteristics: (1) The intent of design research is to create and study new forms of instruction, in this sense, it must be interventionist rather than naturalistic. (2) The goal of design research is to generate theories about the process of learning and the means of supporting that learning. Generating theories here means both, developing and refining theories (but rarely “testing” in the narrow sense of experimental psychology). Individual design experiments aim at pragmatic, humble theories; humble in the sense of being concerned with topic-specific learning processes, and pragmatic in that they effectively inform prospective design. The authors further
stress the central role of prior research as a means to both, specify the design and justify the differentiation of central and ancillary conditions. In a chain of projects, theories are amalgamated beyond the humble, pragmatic ones (cf. Section 3).

(3) The connection between theory and experiment is twofold, namely prospective and reflective: Theory prospectively informs the design for the design experiment, and is further developed in the retrospective reflection on deviances between the expected and the observed teaching and learning processes.

(4) Typical for design research studies are the iterative cycles of invention and revision; when conjectures are refined during an experiment or between experiments (Confrey and Lachance 2000), in respectively micro-design cycles and macro-design cycles (Gravemeijer and Cobb 2013). The micro-design cycles take place within one design experiment, when the researches try and adapt both, the instructional activities and the theory that underpins them. Each micro-design cycle consists of an anticipatory thought experiment, the enactment of instructional activities, and analysis, leading to adaptation or revision of subsequent activities. Macro-design cycles occur when design experiments are repeated while building on one or more preceding design experiments. In practice, design research may vary in the extent that it builds on either micro- or macro-cycles, or both. And, as will show later on, the contributions to this issue show that these are not the only ways in which individual trials can build on each other. Moreover, knowledge is not only gained in the various iterations, but also in the retrospective analysis.

(5) The emphasis on ecological validity and practice-orientation “reflects its pragmatic roots: […] The theory must do real work” (Cobb et al. 2003, p. 11). Because the research is situated within real classrooms, the conditions of the study already represent the complexity of conditions of practice. Hence, care has to be taken that research reports describe those conditions, and the way they may have influenced the learning process. In addition, the theories are closely tied to the activities of students and teachers and tested locally, and revised repeatedly.

As a caveat we may note that although the research is carried out in real classrooms, and in this sense is closely linked to practice, the experimental reform classrooms differ immensely from everyday practices in regular schools due to the interventionist characteristics. How this gap might be bridged is described in the contribution of Cobb and Jackson (2015). Another solution is presented by Stephan (2015) who shows how carrying out design research in close collaboration with teachers can be instrumental in avoiding the aforementioned gap.

2.2 Variability of design research approaches today

Although most design research approaches can be subsumed under the previously identified five characteristics, they still present a large variability in forms, depending on its
origin, its actual context and the specific needs it is supposed to fulfill. Hence, literature of the last ten years pays tribute to this variety (e.g. Kelly et al. 2008; Plomp and Nieveen 2013; van den Akker et al. 2006, for educational research in different domains; and Bakker and van Eerde 2015 for mathematics education).

The above-mentioned book of Plomp and Nieveen (2013), which comprises 51 illustrative cases of design research, marks the interest in and use of design research in general. The same can be said for mathematics education, where especially in the last years, the fields of actions have been extended to all age levels and institutions, from Kindergarten (Sarama and Clements 2002), primary schools (e.g. Confrey and Maloney 2015; de Beer et al. 2015), secondary schools (e.g. Gresalfi 2015; Lobato et al. 2015; Prediger and Krägeloh 2015; Stephan 2015) to university (e.g. Kwon et al. 2015; Rasmussen 2001).

Surveying the field, we may observe that design research not only varies with respect to age groups, but also with respect to:

- the reasons for doing design research,
- the type of results,
- the intended role of the results in educational innovation one is aiming for,
- the scale of the design project, and
- the background theory.

For some researchers, the main reason for doing design research is the wish to account for the messiness of real classrooms. For others, the necessity of experimenting primarily emerges from their desire to come to grips with forms of instruction that do not yet exist. In the former case, practical usefulness may be an important goal, whereas in the latter case, understanding what is going on and what is made possible in the classroom as well as developing corresponding theories are paramount.

In terms of results, the former may aim at producing artifacts that can be used directly in classrooms. In contrast, the latter may aim at local instruction theories or more general insights that cannot directly be applied. However, they are considered informative for practitioners, instructional designers and researchers.

Depending on what role the results are meant to play in educational innovation, the proceeds of a design experiment may be ready to use, translated into a prototypical instructional sequence, or framed as a paradigm case. In relation to the latter two, we may observe a continuum that varies from approaches that are closely related to the method of a one-on-one teaching experiment (Confrey 2006), to design research in which the one-on-one teaching experiment methodology is paired to the development of prototypical instructional sequences and the underlying local instruction theories (Gravemeijer and Cobb 2013). An additional distinction concerns the question whether the design experiment is basically self-contained, or part of a larger research program.
Van den Akker (2013, p. 55) describes differences in scale by listing different levels to which the research and design refers. From the nano level (of individuals and single tasks), micro level (classrooms and teaching units), meso level (e.g. school-specific curriculum), macro level (e.g. national syllabi or core objectives) up to the supra level (international or internationally comparative aspects). This ZDM-issu mainly refers to the micro level, with one exception (Cobb and Jackson 2015) on the macro level.

Finally, implicit or explicit background theories on teaching and learning will strongly influence both the conception and the results of the design experiment. Here a self-evident example is that of socio-constructivism, which orients design-researchers to think through and analyze how students construct new knowledge (cf. Section 3 and 4).

2.3 Two archetypes of design research

The variation on the above aspects results in a variety of approaches to design research. However, this variety may be smaller than expected, since the aforementioned characteristics are not independent from each other. In fact, we may discern two archetypes of design research, one that primarily aims at direct practical use, and one that primarily aims at generating theory on teaching learning processes. Both aim at generating theory and practical products, but there is a difference in focus and time span. This becomes visible with respect to what the research is expected to produce, and what role those products are meant to play; respectively:

• curriculum products and design principles, ready to be used by practitioners and instructional designers;
• local theories and paradigm cases that are meant to inform practitioners and researchers.

The second approach is the focus of this introduction and special issue. We briefly discuss the first approach, which is exemplified by the work of Van den Akker, Plomp, and colleagues.

Nieveen, McKenney, and Van den Akker (2006) characterize their approach as design research from a curriculum perspective, which has as its leading objective to improve understanding of how to design for implementation: “(...) insights are sought on how to build and implement consistent, harmonious, and coherent components of a robust curriculum (...)” (ibid, 72). They envision three types of outputs of design research: design principles, curricular products, and professional development of the participants. The design principles constitute the theoretical yield; often generalizations on the basis of the characteristics of the actual intervention. The approach consists of successive approximations of ideal interventions which are indicated as prototypes.

These prototypes are subject to an elaborated process of formative evaluation cycles in a series of phases, which are summarized by Plomp and Nieveen (2013) as: preliminary research, development or prototyping phase, and assessment phase. The preliminary re-
search phase concerns a problem analysis and the development of a conceptual framework. The prototyping phase encompasses various iterations of the prototype, with formative evaluation as the main research activity. The assessment phase also may comprise a number of iterations but the evaluation has a more summative character. The evaluation criteria that are used in the various phases, reflect the focus on ready-to-use products, they encompass the quality criteria of relevance, consistency, practicality, and effectiveness. Using a similar set of cycles, curricular development can be described in phases of pilot studies, field-tests and evaluation of full implementation, leading to the articulation of program theory and implementation study (Schoenfeld 2007).

3. General characteristics of design research focusing on learning processes in mathematics education

The focus of this volume is on the second archetype, design research which focuses on learning processes and generating local theories. This section provides an overview of the background theory and research that serve as the foundation for these studies

3.1 Relevance of focus on learning processes and basic assumptions

Background theories (like e.g. genetic epistemology, constructivism, constructionism, socio-cultural approaches and situated learning) act as a fundamental core of design research approaches in mathematics education. Most often chosen are those theories which articulate an active role by students with support from teachers in constructing their own knowledge (Brousseau 1997; Cobb and McClain 2004; Confrey and Kazak 2006; Steffe 1991; Voigt 1985). These theories compel design researchers to design and create classrooms where students are provided rich tasks to work with and ample opportunities to participate, individually and collectively. Once these conditions are met, the research concentrates on the emergence of students’ thinking over time and seeks to identify both, productive moments and moments of failure, refining the relevant designs in light of them. The ramifications of modifications are tested and explored both during the teaching phase and in the retrospective analysis (Cobb et al. 2003).

Design research requires a different approach to curriculum and its supports needed by teachers than is often assumed in direct studies of curriculum (Grouws et al. 2010). The goal is not to study the implementation of a “finished piece of curriculum” where measures of fidelity of implementation apply (Huntley 2009; Tarr, Grouws, Chávez, and Soria 2013). Rather the curricular materials in design research are used to stimulate and leverage active, anticipated and unanticipated forms of learning, and therefore revisions during the experiment are expected (Ball and Cohen 1996). The materials (tasks, selection of tools, sequencing) are designed to elicit student thinking and to see where they take the class or rather where the class takes the materials and ideas (Watson and Ohtani in press). Design research requires that teachers are prepared and supported in developing
a knowledge base sufficient to support the constructive process, both for individual stu-
dents and for students-to-student and student-teacher interactions (Lehrer, Carpenter,
Schauble, and Putz 2000).

Most design research with a focus on mathematics learning processes is informed by
grand learning theorists such as Piaget, Vygotsky and Dewey and their intellectual heirs
such as Freudenthal (1968), Brousseau (1983), Treffers (1987), Collins (1992), Brown
and Campione (1996), and Vergnaud (1996). These theorists focus on what it means to
know and understand as it relates to learning over time and among communities of learners.
From these foundational thinkers, a set of three broad assumptions about learning can
be derived and are used by most designer researchers.

Firstly, design researchers with this background draw on learning theories that treat
students as epistemic agents of their own who bring to bear their own experience and
resources. Students are not incomplete adults (Kaput 1999); they have their own points of
view, which must be carefully unearthed, often to the surprise and delight of the research
teams (Ackermann 1995; Duckworth 1996; Maher 2005). These points of view demon-
strate how children are actively working to make sense of the phenomena or circum-
stances and do so within the limits of their experience (Confrey 1991). While such re-
stricted perspectives can lead to “misconceptions”, closer examination often reveals roots
in justified, albeit limited, reasoning and/or potential alternative conceptions (Brousseau
1983; Confrey 1990; Lehrer et al. 2000; Minstrell 2001). The foundational learning theo-
ries seek out the actions taken by students to solve problems and study how they develop
into cognitive operations and schemes that support building rich concepts over time
(Ackermann 1995; Confrey 1990; Kamii 1985; Prediger and Schnell 2014; Thompson
1979). In the building of knowledge, students use representations to express their ideas in
ways that may differ and shed light on conventional representations and evolve over time
(Janvier 1987; Kaput 1987). In addition, children bring to school their own personal and
cultural experiences, which profoundly influence what and how they learn (Moll et al.

Secondly, design research with this background is typically conducted over an extend-
ed time period because of the benefit of examining gradual and steady or rapid and dra-
matic changes as students learn substantial ideas, conceptions, or strategies. One means
of study of the conditions students are likely to change their thinking was labeled “con-
ceptual change” (Posner et al. 1982), and sought to examine how changes in student
thinking could be informed by historic changes and epistemological practices in the dis-
ciplines during the evolution of the ideas.

Thirdly, based on the key theories of Vygotsky and his successors, design researchers
with this background recognize that thought and action are intricately interconnected and
influence each other. This implies that design researchers have to closely attend to the
discourse in the studied classrooms. Student discourse fosters learning to the degree the
participants learn to listen and seek to understand the perspective of others and as teach-
ers scaffold student thinking (Cazden 2001; Hufferd-Ackles, Fuson, and Sherin 2004). Vygotskian and related frameworks point to the mediational role played by the choice, availability and affordances of tools and their semiotic character (Gravemeijer, Lehrer, van Oers, and Verschaffel 2002). Consequently, design researchers are advised to pay close attention to the variety of representations developed by the students, the tools and manipulatives used, and how these interact with the class’s developing ways of talking, explaining, and justifying their thinking. Thus many design researchers document how the use of these representations change over the course of the study as students build on each others inventions and ways of talk are assimilated into ongoing dialogue. For these reasons, artifacts from students’ seatwork and board work are a key source of evidence in the studies (Simon 2000, Lehrer, Giles, and Schauble 2002).

3.2. Consequences for design research studies with a focus on learning processes

The design research in the current issue shows direct evidence of such a focus on learning. During all three stages of conduct of planning, executing and retrospectively analyzing data, this process focus leads to a number of methodological commitments from the researchers. They require the investigator to observe classes, to listen to students and teachers carefully recording their uses of language, inscriptions, and explanations, and, to collect and substantial amounts of student work samples and most fundamentally, be deeply curious about what occurs within the context of learning mathematical ideas. That is why design research with a focus on learning really has to investigate the process of learning, not only its inputs and outputs. Thus, a process-orientation of the investigations is a major characteristic (Gravemeijer and Cobb 2006; Prediger et al. 2012; Steffe and Thompson 2000), which becomes increasingly adopted by other research approaches outside design research, e.g. control trials which start to account for the processes in terms of implementation quality or explaining effects.

This focus on learning as a process in which students are actively engaged in making sense of mathematical ideas remains a hallmark of the type of design research that is central here. Just as in the past, in this special issue, the topics still vary widely in grade range (from primary to tertiary level), and in content (from equipartitioning, speed, division and ratio reasoning, to integers, statistics, arithmetic problem solving, functions, and multivariable calculus). But what is also noticeable is that nearly all of the papers combine a primary design challenge with a secondary one. Besides restructuring a mathematical topic, a second challenge involves more general instructional issues like inquiry-based learning (Kwon et al. 2015), strategic scaffolding (Prediger and Krägeloh 2015), engagement (Gresalfi 2015), model-based reasoning (de Beer et al. 2015), and the role of diagnostic assessment (Confrey and Maloney 2015).

Design research studies with a focus on learning processes can take place in laboratory (interview) settings and in classroom settings. Design experiments in laboratory (interview) settings with small groups of students (n=1-6) are mostly used for early design ex-
periment cycles, focusing on specifying and structuring the learning progression (e.g., Prediger and Krägeloh 2015 start with laboratory settings). For grasping the classroom dynamics and classroom routines and norms, (later cycles of) design research are situated in classroom settings.

Because the focus on process extends over the duration of the study, researchers have time to test ongoing and evolving conjectures about student reasoning and classroom interactions, by trying mini-experiments and exploring the robustness and dispersion of ideas and their influence on each other. When an unexpected result arises during the conduct of the research, the team will devise a way to adjust the instructional plan to allow them to test the meaning and extent of the result. This may require the development of an additional formative assessment, the modification of a task assignment to ask for more explanation, or preparing the teacher to watch for and bring out further examples. For example, Lobato et al. (2015) report on how, within a design study on quadratics, they recognize a critical need to understand how students were reasoning about division and to address it through a small intervention within a study of quadratics. Simultaneously, there was a critical need to help students change their understanding of explanations, so as to use division meaningfully in the context of an evolving understanding of rate and its representation on a number line. It was the retrospective analysis of the second study that led them to revisit and reinterpret the first study. This illustrates how a conjecture may evolve in a study, or across study iterations.

Even projects with a strong focus on curriculum innovation like Kwon et al. (2015) sometimes adopt a process focus: In their design research project on inquiry-based learning for a multivariable calculus course, the quality criteria for which the inquiry-based learning is optimized required the process focus: When quality of innovation is related to the deepening the structures of students’ argumentation processes, it is necessary to develop analytic tools for investigate deeply the students’ processes (Kwon et al. 2015).

4. Theories, theorizing and methodological issues

4.1 Critical perspectives on design research and methodological standards

Design research is still a contested approach (e.g. critique by Dede 2004; Kelly 2004; Philips and Dolle 2006; and others). One main criticism is that design research is not one well-defined research method. And it is true that there is a large variety of research approaches labeled design research. However, rather than to force all variants of design research into one straightjacket, we propose to acknowledge the variations, and demand specification in each study of how their methods were tailored to the individual purpose and context. In Section 2.3, we already discerned two main archetypes; that of design research from the curriculum perspective, and design research with a focus on learning
processes. The first does have a rather well articulated research method—which has some similarities with formative evaluation (cf. e.g. Plomp and Nieveen 2013).

In contrast, design research projects in the second archetype with a focus on learning processes apply a larger variety of methods and data analysis procedures. When discussing methodology, the bottom line for a research project is in the end: “Why should one believe its claims?” A standard quality criterion for achieving such an accountability is to make transparent the process by which the results are produced (cf. Freudenthal 1991). We may observe that in classic research approaches, accountability is achieved by standardized research methods. This does not mean, however, that a sound research approach cannot exist without some pre-accepted standard procedure. In contrast, we would argue that the deeper analysis of learning processes requires adaptations and new inventions of data analysis procedures, which match to the specific topics and learning contexts. Still, even though each design research project has to produce its own methodological justification of its methods, it seems reasonable to formulate general directions on what such a justification would have to look like—a point we will come back to later.

Some researchers have proposed means to reconcile design research with other types of curriculum study. For instance, Schoenfeld (2007) suggested a three-phase model, which would start with small-scale design experiments, then midscale exploratory studies aiming at understanding how and why instruction works, and finally expanded studies replicating the effects discovered in phase 2, or comparative in-depth evaluations of any number of promising curricula (involving random assignment of participants) (Schoenfeld 2007, p. 97f). This three-phase model is one possible means to use design studies, but if it implies all design research studies are pilot studies or preliminary to the real study, then such a phased model fails to recognize the full contribution of the approach of design research. Instead, small-scale design experiments can inform researchers more deeply about the mechanism of teaching and learning in a way that explains how learning occurs. This sheds light on theory, simultaneously. Furthermore, in the three-phase model, there is a danger that particularly in the later phases, the teacher becomes the person charged with fidelity of standardized implementation, i.e. loosing the necessary agency of adaptations to local contexts and students.

A fundamental question that underlies much of the critique on design research is, if it is (in principle) possible or impossible, to simultaneously pursue a design goal and a theoretical yield. Philips and Dolle (2006) are quite outspoken in this respect by claiming that it is not possible to freely pursue design possibilities and at the same time make theoretically sound conclusions around theory. They apparently reason from the classic experimental perspective of the natural sciences, while seeking a controlled means of identifying the set of independent variables which influences the output variables. We may counter, however, that design research is grounded in the assertion that classrooms are complex ecologies where one cannot sufficiently control all variables to draw causal conclusions. This calls for a different approach than the reductionist approach of natural science.
like psychology in which phenomena are disassembled in individual variables whose interdependencies can be researched systematically—essentially by testing hypotheses.

The underlying idea of grasping how things work resonates with Maxwell’s (2004) process-oriented conception of causal explanation, which he juxtaposes with a regularity conception of causality. The regularity conception of causality is based on the logic that, if A follows B frequently, then B is the cause of A. The process-oriented conception of causality refers to the “mechanisms through which and the conditions under which that causal relationship holds” (Shadish, Cook, and Campbell 2002, p. 9). This is a more holistic conception that fits with the objective of design research, to come to understand how things work. Which—as we argued before—may have more practical value for teachers, than knowing that A works better than B.

Many criticizers seem to conceptualize educational innovation from a fidelity perspective, in which teachers are expected to use instructional materials in prescribed ways in order to achieve a specified result. This conception of educational innovation is at odds with the background theories we described earlier (Section 3.1), based on an active role of the students in constructing their own knowledge. This requires the teachers to continuously adapt to how the students of their classroom act and reason rather than simply use ready-made instructional materials in prescribed ways. For a more adaptive and interactive kind of teaching, a different kind of teacher support is needed, enabling teachers to interpret how their students act and reason, and to react adequately. The here discussed archetype of design research aims at offering theories which inform teachers about the anticipated learning processes, and the potential means of supporting those learning processes. Thus, the objective is not—as Philips and Dolle (2006) presume—to design a finished piece of curriculum. Instead, the design aspect concerns design and creation of learning environments as a means for coming to understand innovative learning processes and the way in which those learning processes can be supported. This understanding may take the form of local instruction theories, as well as theoretical frameworks that address more encompassing issues. Following the Research Advisory Committee of NCTM (1996), we may speak of a change in research goals—from “What works?” to, “How does it work?” And, as they go on to say, this implies a shift in the norms of justification. Hence guaranteeing methodological standards requires a deep reflection on the role of theories in design research.

In spite of the variations between forms and aims of design research with a focus on learning processes, common criteria have been established, especially the role of theories, relevance of the questions and results for classrooms and the educational system; (ecological) validity, reliability, transparency, intersubjectivity, for the empirical methods (cf. Bakker and van den Eerde 2015; Kelly 2003). In the next subsections, we discuss the role of theories as quality criteria for design research.
4.2 Different levels of theories

Two central criteria help to distinguish high quality design research from more or less research-inspired design: the coherent *use of theories* as a base of design research, and relevant contributions to *further theorizing* (Confrey and Lachance 2000; Design-Based Research Collaborative 2003; Gravemeijer 1994). Naming these both criteria refers to the double role of theories in mathematics education research (Assude et al. 2008; Prediger and Bikner-Ahsbahs 2010):

(a) theory as a framework for research, and,
(b) theory as a major outcome of research.

Sometimes the framework and the outcome belong to the same kind of theories or even the same theories, but with adaptations and with a different level of empirical support or further refinements. Hence, the roles of theory in design research are rather complex and dynamic.

The role of theories as a precondition for methodologically sound design research has often been emphasized (e.g., Bakker and van Eerde 2015; Cobb et al. 2003; Confrey 2006). Often mentioned are different levels of generality (Confrey and Kazak 2006; diSessa and Cobb 2004; Gravemeijer 1998), listed from most general level to the most specific:

- orienting frameworks or background theories
- domain-specific instruction theories as frameworks for action
- local instruction theories / humble theories / hypothetical learning trajectories

Unfortunately, there is no widely accepted nomenclature; not only are various labels used to describe the same levels, but also the same labels are used to denote different levels, even by the same authors. If possible, we will use the meaning that was intended by the authors who first coined the terms.

*Orienting frameworks* (diSessa and Cobb 2004) or *background theories* (Gravemeijer and Cobb 2013), such as socio-constructivism or cognitive theory, function as the foundation of the research and significantly influence both, the design and the way data are interpreted. They do not change or emerge in a research project. Once the researchers have committed themselves to certain grand theories and background theories, those theories are basically treated as givens in a design research project. The article of Abrahamson (2015) shows, however, that a series of design research projects can lead researchers to switch to another background theory.

*Domain-specific instruction theories* are specific for the school subject, in our case mathematics education and offer a general framework for action. One example of such a domain-specific instruction theory is Brousseau’s (1997) Theory of Didactical Situations which is instrumental in the French Didactical Engineering as described by Margolinas and Drijvers (2015). Another example is the theory of realistic mathematics education,
RME, which is used in various contributions to this special issue (de Beer et al. 2015; Kwon et al. 2015; Margolinans and Drijvers 2015; Stephan 2015). RME is typical in that it was defined as “a domain-specific instruction theory” (Treffers 1987) from the outset. Treffers construed RME retrospectively by analyzing the results of various design research projects trying to fulfill Freudenthal’s (1973) adagio of “mathematics as a human activity”. Although being the product of design research, Treffers’ RME theory also started to function as a guideline for new design research projects. Gravemeijer (1999) recasts RME in terms of instructional design heuristics: guided reinvention, didactical phenomenology and emergent modeling, where emergent modeling may be considered an addition to Treffers’ work, which again was the product of retrospective analysis of a number of design research projects. As a caveat we should add that it is not always clear whether the classroom culture is considered a part of the domain-specific theory or not. However, enacting RME requires a specific classroom culture (as clearly comes to the fore in Stephan 2015).

diSessa and Cobb (2004) have emphasized that even the domain-specific instruction theory is mostly too general for deriving designs. One of the RME heuristics, emergent modeling for instance, advises the designer to look for a model that models the students’ experientially reality, may support students in developing mathematical relations, and can become a model for more formal reasoning when the model starts to derive its meaning from the emerging network of mathematical relations (Gravemeijer 1999). When designing a learning trajectory for a given topic, however, the design heuristic does not tell you what model to choose.

A local instruction theory hence addresses the learning of a specific topic such as addition and subtraction of numbers up to ten, of the multiplication of fractions, offering theories about a possible leaning process, together with theories about possible means of supporting that learning process (Cobb et al. 2003; Gravemeijer 1998). These means of support include the classroom social norms and the socio-mathematical norms that have to be in place. The means of support encompass potentially productive instructional activities and (computer) tools (discussed by Hoyles and Noss 2015) as well as an envisioned classroom culture and the proactive role of the teacher. This broader perspective fits with the notion of a learning ecology:

“Elements of a learning ecology typically include the tasks or problems that students are asked to solve, the kinds of discourse that are encouraged, the norms of participation that are established, the tools and related material means provided, and the practical means by which classroom teachers can orchestrate relations among these elements.” (Cobb et al. 2003, p. 9).

Whereas the local instruction theory is formulated on the level of an instructional sequence, the hypothetical learning trajectory (HLT) was originally defined on a more specific level, that of instructional activities, the anticipated mental activities in which students will engage when they participate in envisioned instructional activities, and considerations on how the anticipated mental activities relate to the end goals one is aiming for (Simon 1995). Speaking of teachers doing the anticipating, Simon emphasizes the hypo-
thetical character of these learning trajectories; the teachers are to analyze the reactions of
the students in light of the stimulated learning trajectory to find out in how far the actual
learning trajectory corresponds with what was envisioned. Based on this information the
teacher has to construe new or adapted instructional activities in connection with a re-
vised learning trajectory.

Over time, the meaning of the term hypothetical learning trajectory has shifted. Simon
(1995) introduced the term in the context of the design of one or two lessons, nowadays,
the term hypothetical learning trajectory, or “learning trajectory” for short (Confrey
2006), is also used for learning processes that evolve during an instructional sequence for
a certain topic and can extend beyond a single class, unit or even school year. Understood
this way, the notion of a learning trajectory is very similar to that of a local instruction
theory. Next to those terms, the appellation “learning progression” is often used nowa-
days. However, this term comes with a different connotation, in that it primarily refers to
an empirically grounded sequence of steps that describe how students are likely to reason
as they move from naïve to more sophisticated reasoning (Daro, Mosher, and Corcoran
2011). Finally Stephan (2015) distinguishes between a classroom learning trajectory and
an individual learning trajectory, highlighting that individual trajectories will vary while
Confrey (2006) argues that a learning trajectory always contains an understanding of
multiple pathways in relation to certain landmarks and obstacles.

The above tidy categorization of theories does not cover all theoretical aspects that
come to the fore in design research. The various studies that are presented in this issue
show that intermediate theoretical gains often concern design details, which truly deserve
the title “humble theories”, for example the design elements of the strategic scaffolding
tool underlying the word problem cracker developed by Prediger and Krägeloh (2015)
which has also led to contributions with a wider range, for example the specification of
necessary strategies.

4.3 Theorizing as a process

All high quality based research is based on theory, but all high quality research also aims
at contributing to developing theories, i.e. to theorizing. But what exactly does develop-
ing theories mean? Some design researchers talk about developing theories in terms of
“validation” (Plomp and Nieveeen 2013, p. 16) which is a usual term for example in the
context of randomized control trials in which hypothesis are validated or falsified. This
way of theorizing is highly valued and in some countries even claimed with exclusive-
ness (Slavin 2002).

However, within design research with a focus on learning processes, theorizing is not
restricted to validating theories formulated in hypotheses but about emergent theorizing
in an evolutionary sense of refining and revising categories and conjectures (Confrey and
Lachance 2000). Metaphorically expressed, this requires to “put theories in harm’s way”,

The process of theorizing often takes place on the level of local instruction theories: The type of design research project we discuss here usually starts with a conjectured local instruction theory, which then evolves by the design experiments. Although also design researchers use terms like ‘put to the test’, the role of such theories in design research is very different from the role theory plays in (quasi-)experimental research. The aim is not to either accept or reject theoretical elements —as is the case in experimental research—but to revise, refine or improve them. This comes to the fore in many of the papers in this special issue.

These papers also show the power of iterations and trials of different elements of a local instruction theory in iterative experiment cycles. Often it is not one design experiment that leads to a local instruction theory, nor a series of design experiments in which a complete theory is iteratively adapted and improved. More often various pieces of understanding gained in individual experiments are put together in one coherent whole (e.g. de Beer et al. 2015; Stephan 2015). In all its variation, accumulation and synthesis strongly shapes the process of theorizing and then contributes to the persuasive power.

The investigation of the students’ learning pathways often transcends the initial conjectures and is open to unanticipated findings (de Beer et al. 2015), or even failures (Lobato et al. 2015). In this respect, design research adheres to Smaling’s (1992) conception of methodological objectivity. He argues that doing justice to the object of research encompasses both, “avoiding distortions”, and, “letting the object speak”. The former corresponds with the usual requirements of reliability and validity. The latter involves looking for signs that may indicate possibilities for promising alternatives.

Accumulation and openness to unanticipated findings can lead to what diSessa and Cobb (2004) call an “ontological innovation”. They argue that a series of design experiments can serve as the context for the development of theories or theoretical frameworks that entail new scientific categories, which can serve as lenses contributing to making sense of what is happening in the actual instructional setting. To fulfill this role, the new categories will have to be embedded in and supported by a corresponding theoretical framework. Note that ontological innovations can play a dual role. On the one hand, they can serve as lenses for making sense of what is happening in the complex, more-or-less real world instructional setting in which a design study is conducted. On the other hand, ontological innovations can function as guidelines or heuristics for instructional design. Social norms and socio-mathematical norms may function as an example. On the one hand, the concepts of social norms and socio-mathematical norms offer an interpretative framework for analyzing classroom discourse and communication (Yackel and Cobb 1996). The interpretative framework then plays a role in interpreting data (as in Lobato et al. 2015), which can be seen as pendant of operationalizing in classical educational research. On the other hand, the same framework reveals what norms have to be established
in the classroom to make certain design experiments successful. This is exemplified in this issue by Gresalfi’s (2015) study on the role of critical engagement.

Coming back to quality criteria for theorizing, most important for us seem to be those criteria which refer to the interplay of experiment and the process of theorizing. Ecological validity requires that the applied theories and the resulting theoretical contributions have to take into account the complexity of classrooms. This asks for a different approach than the reductionist approach of natural science like psychology in which phenomena are disassembled in individual variables whose interdependencies can be researched systematically—especially by testing hypotheses. In this respect, we may refer to Gould (2004) who depicts a complementary way of knowing; the more holistic approach of the humanities, in which, in his view, concilience plays a large part: “the validation of a theory by the ‘jumping together’ of otherwise disparate facts into a unitary explanation” (ibid, p. 192). He presents Darwin’s theory of evolution as a key example of the latter; the credibility of his evolution theory resides in the fact that it explains the large variety of species in one stroke.

Cobb, Jackson, and Dunlap (2015) take this process-oriented conception as a basis for elaborating the scheme of argumentation that underpins the methodology of this type of design research. He adopts the term “argumentative grammar” that is introduced by Kelly (2004). According to Cobb et al. (2015), the argumentative grammar that links data to analysis and to final claims and assertions, requires design-researchers to show:
(a) that the students developed specific forms of reasoning,
(b) that this has to be attributed to the students’ participation in the design experiment,
(c) how these forms of reasoning emerged, and
(d) how the evolving learning ecology ensured this process.
The articles in this issue show a broad range of ways to meet those requirements.

5. Outlook on the papers of the special issue

The studies in this volume, drawing on the sketched theories of learning, and situated in the broad evolution of an engineering based use of design studies with a strong focus on theorizing, share a number of the characteristics described heretofore:

Explicit articulation of a broad or grand theory under which the work is conducted and the development of local instructional theories (Gravemeijer and Cobb 2006; 2013), humble theories (Cobb et al. 2003) or bridging theories (Confrey and Kazak 2006). This volume begins with a report from Margolinas and Drijvers (2015), which provides an overview on two traditions of research that are frequently used as the grand theories for European design research: didactical engineering from the French research community based on Guy Brousseau (1983; 1997), and Realistic Mathematics Education from the Dutch tradition evolving from the fundamental ideas of Freudenthal (1968; 1973). In
each, the scholars describe the need to study learning in context and in real classrooms led to the need to evoke an engineering orientation to design research. They further described how such research would be approached through the careful design of tasks followed by observations and analysis to further refine tasks. Other studies in this special issue track back to teaching experiments (Abrahamson 2015; Confrey and Maloney 2015; Stephan 2015) and early design research (Gresalfi 2015; Labato 2015).

In the contribution of Confrey and Maloney (2015), the authors report on two conjectures and how they were investigated over the course of the study they conducted on the use of diagnostic assessments with second, third and fourth graders. In one case, they show how the theory had to be substantially reevaluated and revised, and in the other, how the conjecture flourished and was elaborated. In neither of these two cases, the results of the study could be anticipated and hypothesized in advance. A design research approach was selected to allow the results to emerge and be carefully documented due to the method in which the study was conducted.

A careful delineation of tasks and how they were designed. Design researchers devote considerable thought to the design of their tasks and how those tasks can elicit rich threads of student reasoning. De Beer et al. (2015), interested in how students combine their understanding of intuitive experience of rate with a meaning of rate as a quantitative description of the relationship between distance and time, report on how they used variety of shapes of glass containers to elicit student thinking. Gresalfi (2015), interested in understanding student engagement in relation to reasoning about measures of central tendency places her tasks in the context of city planning, strengthening the game parameters and raising the stakes.

As has been noted previously in Cobb et al. (2003), design research studies are uniquely suited to exploring the use of rich tasks using new technologies to support innovation and intervention. In the special issue, authors demonstrate ways of designing and redesigning technologies such as the LYPP-Sync, a diagnostic system delivered on smart phones (Confrey and Maloney 2015) or Seeing Chance, Kinemathics, and Giant Steps (Abrahamson 2015). In one submission (Gresalfi 2015) the author works to bring together the use of a game, Quest Atlantis, and their compelling forms of engagement together with what is known from research about the relative merits of different conceptions of central tendency. The combination of design studies reported leaves one with a very contemporary view of education.

Examples clearly demonstrate articulated evolution of the conjectures over the course of the cycles and variations of design and how they affect the phenomenology of the classroom events. Gresalfi (2015), Lobato et al. (2015), and Prediger and Krägeloh (2015) most clearly demonstrate the process of revising iterations of design based on the ongoing results. Gresalfi (2015) does this over a series of three years of development and
revision within her ongoing research program, while Lobato et al. (2015) demonstrate how conditions of failure can be leveraged to find, explore and eventually weave in unexpected results. Prediger and Krägeloh (2015) present a project with six design experiment cycles in laboratory settings with a total of 70 students. The problems appearing in the third design experiment cycle gave rise for their work on comprehension strategies.

**Articulation of the critical role of classroom milieu and related sociomathematical norms in the conduct of the instruction.** Because the studies often introduce new types of tasks, tools, and ideas, there is often a need to explicitly address and refine the norms of classroom behavior, interaction patterns and expectations about student reasoning. This is evident in the second part of the article of Lobato et al. (2015) when student weakness in understanding division is accompanied by a need to set new expectations for explanation and understanding. It is also evident in the way in which Gresalfi (2015) orients to define engagement and evidence. In Cobb and Jackson (2015), their primary claim “most teachers therefore have to fundamentally reorganize rather than merely extend or elaborate their current instructional practices if they are to enact the products of classroom design studies as intended” requires changes in classroom norms and expectations.

**Careful attention to argumentation, explanation and warrant with attention to generalizability of findings.** Confrey and Maloney (2015) report on the extensive analysis that was required to explain the data from their study and how it led them to a variety of outcomes: revising materials, checking the accuracy of the algorithm operating in the software and reformulating aspects of the equipartitioning learning trajectory. The contribution of Prediger and Krägeloh (2015) articulate a variety of outcomes and show how these are explicitly related to their original conjecture and assumptions.

**Thoughtful consideration and intentional consideration of the relationship of the research-practice interface.** While design research has a prima facie relationship to practice in comparison to laboratory-based studies, design researchers often attend to this in multiple ways. Typically this surfaces in relation to the role of the teacher in the study. These studies vary in that regard. In the report of Confrey and Maloney (2015), the teacher is a researcher and the study is conducted in a special summer session, a decision largely influenced by the novelty of the technology used. Describing the development of design research, Margolinas and Drivers (2015) articulate how in didactical engineering, the role of the teacher has evolved. It has evolved from teacher as partner, to teacher as collaborator, to studying teaching in its reciprocal relation to student learning. In contrast, In contrast, in the paper of Stephan (2015), the researcher is the teacher in some instances, but she collaborates with a group of teachers as she merges design research with lesson study. She reports on how the consistent involvement of teachers influences the ecological validity of the work. Cobb and Jackson (2015) take on the issue of “transfer” of de-
sign research most directly and articulate a set of considerations that need to be recog-
nized and tackled for the results of these studies to influence practice at scale. They call for a set of design studies that “investigate dissemination processes by focusing on the development of the capacity for instructional improvement at the level of individual schools and of school systems.”

We welcome our readers to this Special Issue and hope that you find the compilation of articles useful and provocative for further elaboration of design research methodolo-
gies.

References


