Educational Design Research

Part A: An introduction

SLO • Netherlands institute for curriculum development

Editors:
Tjeerd Plomp & Nienke Nieveen

slo

Educational Design Research

Part A: An introduction

Editors:
Tjeerd Plomp & Nienke Nieveen

slo

Colophon

Authors: Jan van den Akker
Brenda Bannan
Anthony E. Kelly
Nienke Nieveen
Tjeerd Plomp

Editors: Tjeerd Plomp
Nienke Nieveen

Design: SLO, Enschede

ISBN: 978 90 329 2334 1

SLO P.O. Box 2041
7500 Enschede
the Netherlands

http://international.slo.nl/publications/edr/

Enschede, November 2013

Contents


WHO
WHAT
WHERE
WHEN
WHY
HOW
QUESTIONS
ANSWERS

Preface

This book is the result of a seminar on 'educational design research' organized from November 23-26, 2007, by Prof Zhu Zhiting (Department of Educational Technology) of the College of Educational Sciences at the East China Normal University in Shanghai (PR China).
The primary goal of the seminar was t*o introduce a group postgraduate students and lecturing staff in China to educational design research as a research approach.* The second goal of the seminar was to prepare, based on the contributions of a number international experts, *proceedings of the seminar* written in such a way that they can be used in postgraduate seminars on educational design research across China.

About 75 people with backgrounds mainly in instructional technology, curriculum and instructional design participated in the seminar. Most of them were working in teacher education, in schools as instructional technologist and/or in distance education. Although participants had (through their studies) already knowledge and some experience in instructional or course design and in research methods, they were eager to be introduced to design research as a relatively new research approach for addressing complex problems in educational practice.

The seminar staff consisted of Profs *Brenda Bannan* and *Eamonn Kelly* (both George Mason University, Fairfax, VA, USA) and Prof *Jan van den Akker* (University of Twente and National Institute for Curriculum Development [SLO], Enschede, The Netherlands), and the two editors of this book Dr *Nienke Nieveen* (National Institute for Curriculum Development [SLO], Enschede) and Prof *Tjeerd Plomp* (University of Twente, Enschede, The Netherlands). As can be seen from the table of content of this book, they are reflecting the background of the participants, as they represented experience in conducting design research in the domains of curriculum development, instructional technology and mathematics and science education. Experts were consciously invited from both Europe (The Netherlands) as well as the USA, so as to ascertain that variation in background and perspective on design research was represented in conducting the seminar.

The chapters in this book are based on the presentations and the small group discussions during this seminar. Although the book does not provide a 'how to do guide' for designing and conducting design research, the chapters have been written in such a way that they reflect both the conceptual underpinning and practical aspects of the 'what' and 'how' of doing design research (chapters by Plomp, Kelly and Nieveen), as well as provide the reader an insight in the specifics of doing design research in the domain of curriculum (chapter by Van den Akker) and instructional technology (chapter by Bannan).

To assist the readers in finding their way in the abundance of literature on design research, we have added a chapter with references and sources on educational design research. This bibliography is far from complete and reflects very much the background and the biases of the editors of this book. Yet we trust that this chapter will assist the interested reader in getting introduced to this exciting and promising research approach.

We want to thank Prof Zhu Zhiting from the East China Normal University for taking the initiative for this seminar. Similarly we want to thank our colleagues for contributing to this book.

But above all, we like to express our hope that this book will stimulate and support many (future) researchers to engage themselves in educational design research.

Jan van den Akker
Director General SLO

Tjeerd Plomp and Nienke Nieveen
Editors

Additional note about the revised edition

As stated in the Preface to the 2009 edition, this book is the result of a seminar on 'educational design research' organized in 2007 by Prof Zhu Zhiting of the East China Normal University in Shanghai (PR China). When we met Prof Zhu in 2011, he asked for a number of illustrative cases (15-20 cases) of successful educational design research (EDR) to be used in combination with the 2009-book *with the purpose that graduate students and novice researchers could also learn from examples by others about how to design and conduct a research project utilizing EDR.*

We, editors, decided to take up the challenge after SLO had indicated to support such an initiative and be willing to publish the new book. This resulted in a major project (starting in December 2011) comprising not only of editing a book with 'illustrative cases of educational design research', but also an update of our 2009 book. So in the end the project resulted in a two-volume book.

This book, Part A: *Educational Design Research: An Introduction* is a revision of our 2009 book. We invited the authors whether they wanted to revise their chapter. This resulted in major revisions of the Chapters 1, 2 and 6 and some updates and minor revisions of the remaining authored chapters. In addition, we felt that that one of the possible foci of design research was underexposed, namely design research with the aim to develop new theories, such as new instructional, pedagogical or learning theories. To compensate for this, we invited Koeno Gravemeijer and Paul Cobb to contribute to our book and we are glad with their chapter 'Design research from the learning design perspective'.
Finally, we revised the chapter on 'References and Sources' somewhat by adding references to a few articles and books that have been published since the first edition of our book, and in addition we added the URL of a number of PhD theses. Moreover, we refer the reader interested in more references to the 'key sources' for the research reported in each of the illustrative cases in the second volume.

Part B: *Educational Design Research: Illustrative Cases* is a varied collection of 51 examples of successful educational design research. The case chapters are written with the aim to enable graduate students and novice researchers to learn how to design and conduct a project utilizing educational design research.

The collection reflects a number of dimensions, such as

- various domains in the field of education, such as curriculum, learning and instruction, subject related pedagogy (for example math education, language education, science education), instructional technology, ICT in education
- various purposes of design research, such as developing an innovative intervention or developing or a new instructional or learning theory
- representing all educational levels: (pre-)primary education, junior and senior secondary education, teacher education, other higher education, and also workplace learning
- having been conducted in more than 20 countries.

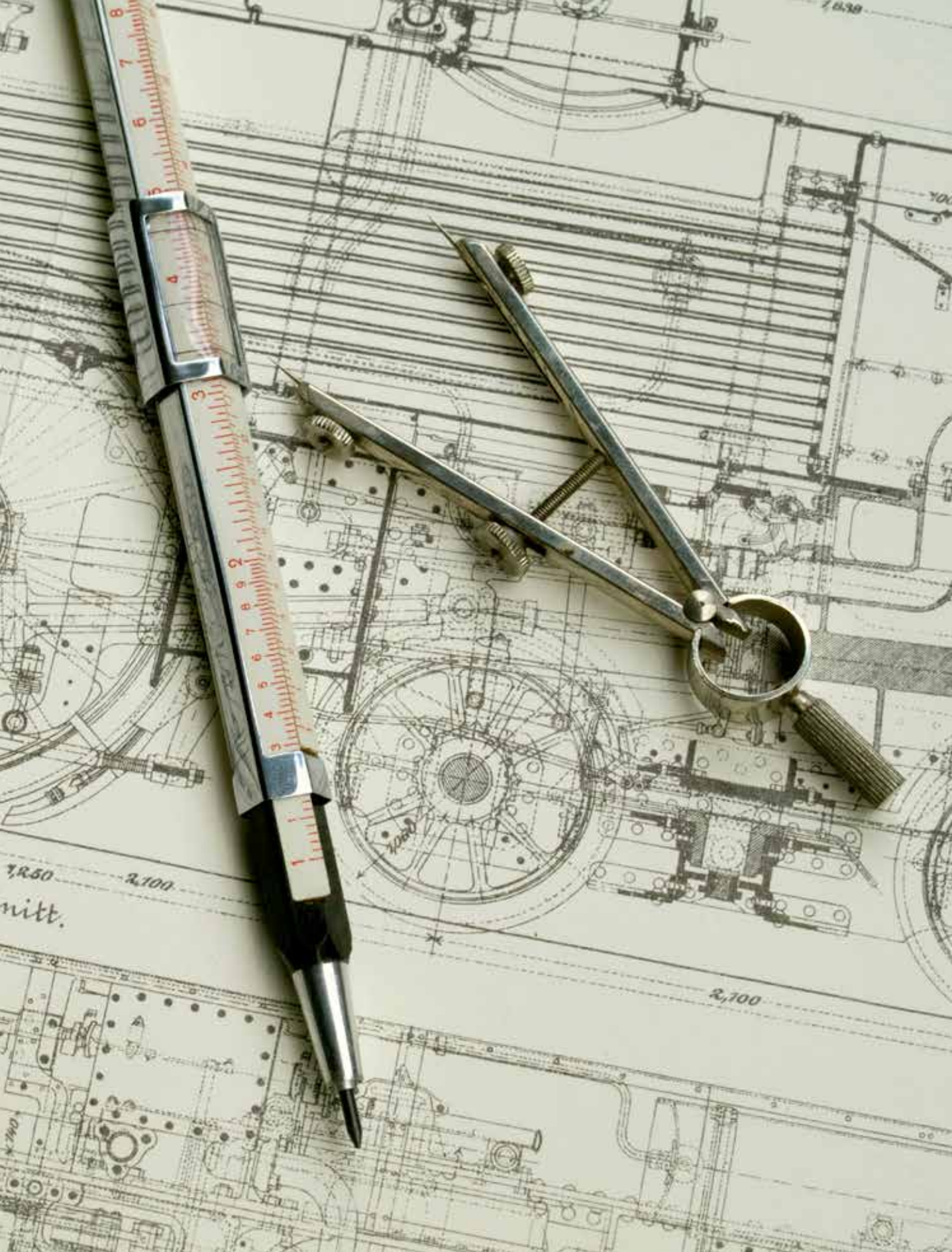
The resulting two-volume book has been published electronically by SLO, the Netherlands Institute for Curriculum Development. Part B, the collection of case studies is a 'supra-book' with each case chapter separately electronically published, which allows readers or users to make their own selection of chapters given their specific purpose of use.
Pdf files of both volumes can be freely downloaded from http://international.slo.nl/edr
This website comprises also a 'case selection tool' to assist users in selecting cases for their intended use.

Professor ZHU Zhiting (East China Normal University in Shanghai, PRChina) and associate professor Wang Qiyun (National Institute of Education, Singapore) will prepare a Chinese edition of this book for which they will be co-editors.

We are very grateful for the support of SLO in preparing and publishing this revised edition of our book.

We hope that this book will become a source of inspiration and good ideas for many (also future) researchers who want to address important problems in educational practice!

Tjeerd Plomp and Nienke Nieveen
Editors



1. Educational Design Research: An Introduction

Tjeerd Plomp

Introduction

The purpose of this chapter is to provide an introduction to educational design research as a research design appropriate to develop research-based solutions to complex problems in educational practice or to develop or validate theories about learning processes, learning environments and the like.
This twofold purpose will be reflected in the definition of design research discussed later in this chapter. However, regardless of the purpose, design research encompasses the systematic study of designing, developing and evaluating educational interventions[1] - such as programs, learning processes, learning environments, teaching-learning materials, products and systems.
The need for a research design that addresses problems in educational practice has been argued by researchers in various 'corners' of the domain of education because of an apparent lack of relevance of much educational research for educational practice.
For example, the Design-Based Research Collective (2003, p.5) states that "educational research is often divorced from the problems and issues of everyday practice – a split that resulted in a credibility gap and creates a need for new research approaches that speak directly to problems of practice and that lead to the development of 'usable knowledge'."

From his background in research in the domain of curriculum development and implementation, van den Akker (1999, p.2) writes "that 'traditional' research approaches such as experiments, surveys, correlational analyses, with their emphasis on description hardly provide prescriptions that are useful for design and development problems in education". He claims that an important reason for design research[2] stems from the complex nature of educational reforms worldwide. Ambitious reforms cannot be developed at the drawing tables in government offices, but call for systematic research, supporting the development and implementation processes in a variety of contexts.

1) intervention is used as a 'container' term referring to all entities that can be designed and developed.
2) which he calls 'development research' in his 1999 publication.

In his review of the state of educational research, and more specifically educational technology research, Reeves (2006, p.57) concludes that there is "a legacy of ill-conceived and poorly conducted research that results in no significant differences or, at best, in modest effect sizes". He also argues for the domain of educational technology that it would be better that educational technologists instead of doing more studies comparing whether in a certain context method A is better than method B, undertake design research aimed at developing an optimal solution for the problem in that context.

In the field of learning sciences, the belief that context matters lead to the conclusion that research paradigms simply examining learning processes as isolated variables within laboratory settings, will necessarily lead to an incomplete understanding of their relevance in more naturalistic settings (Barab & Squire, 2004, p.3; with reference to Brown, 1992). In this field, design-based research "was introduced with the expectation that researchers would systematically adjust various aspects of the designed context so that each adjustment served as a type of experimentation that allowed the researchers to test and generate theory in naturalistic contexts".

These authors illustrate the need for design research as an alternative research approach. Before elaborating on design research, the next section will first discuss more generally possible functions of research and how research functions are related to research designs. After this, design research is defined and a number of characteristics of this research design are presented. This is followed by a discussion on how design research may vary in focus resulting in differentiation between development studies aimed at design principles, and validation studies aimed at theory development and validation. The next section takes a closer look at design research by discussing the main research question in design research and the quality criteria for interventions. This section also presents a simplified research model for design research. After reflecting on the issue of generalizability of design research the following section presents some aspects of conducting design research, such as formative evaluation as the prominent research activity in design research, and how the design stages can become micro-cycles of research. The next section illustrates that conducting design research puts researchers in a situation in which they have to face a number of challenges and how to manage these. These will be discussed before ending the chapter with a section presenting a few concluding remarks and an outline of this book.

A final note on terminology, following van den Akker, Gravemeijer, McKenney, and Nieveen (2006) design research is used as a common label for a 'family' of related research designs which may vary somewhat in goals and characteristics – examples are design

experiments (e.g. Brown, 1992; Cobb, Confrey, diSessa, Lehrere, & Schauble, 2003), design studies (Shavelson, Phillips, Towne, & Feuer, 2003; Walker, 2006), design-based research (Design-Based Research Collective, 2003), developmental research (Lijnse, 1995; van den Akker, 1999) and engineering research (Burkhardt, 2006). In addition to these, other members of this 'family' are participatory action research (Eilks & Ralle, 2002; Marks & Eilks, 2010) and design-based implementation research (Penuel, Fishman, Cheng, & Sabelli, 2011).

Research functions and research designs

Before elaborating on the meaning of design research, this section positions design research as a research design in conjunction with other research designs.

The primary function of scientific research is the search for 'understanding' or for 'knowing' with the aim of contributing to the body of knowledge or a theory in the domain of the research. Other broad aims could be to provide insights and contributions for improving practice, and to inform decision making and policy development in the domain of education.
There is a great variation in the possible functions of research, and – dependent on their particular aims - each of these can be realized through one or more research designs.

Research functions

In general, various research functions can be identified and distinguished from each other, with each reflecting certain types of research questions. Examples of research functions (with exemplary research questions illustrating the function) are:

1. *to describe:* e.g., what is the achievement of Chinese grade 8 pupils in mathematics?; what barriers to students' experience in the learning of mathematical modelling?
2. *to compare:* e.g., what are the differences and similarities between the Chinese and the Netherlands curriculum for primary education?; what is the achievement in mathematics of Chinese grade 8 pupils as compared to that in certain other countries?
3. *to evaluate:* e.g., how well does a program function in terms of competences of graduates?; what are the strengths and weaknesses of a certain approach?; etc.
4. *to explain or to predict:* e.g., what are the causes of poor performance in mathematics (i.e. in search of a 'theory' predicting a phenomenon when certain conditions or characteristics are met)?
5. *to design and develop:* e.g., what are the characteristics of an effective teaching and learning strategy aimed at acquiring certain learning outcomes?.

In many research projects, the research questions posed are often such that in fact various research functions do apply, which results in the need to identify a primary function. For example, if the research question pertains to comparing the mathematics achievement of Chinese grade 8 pupils to that in certain other countries, then as part of comparing, the researchers will evaluate the achievement of grade 8 pupils in each of the countries involved. Or, as another example, if one wants to design and develop a teaching-learning strategy for developing the competency of mathematical modelling (in grade 11 and 12), then researchers may first want to understand and carefully describe what barriers students experience with mathematical modelling, and at a later stage also evaluate whether the teaching-learning strategy that has been developed is effective. Both examples illustrate that usually a research project has a primary research function, but that other research functions need to be applied to 'serve' the primary research function.

At the level of a research project, beginning with a research problem or question, there is a logical sequence of development, namely:

Research question ⟶ *(primary) research function* ⟶ *choice of research design*.

In this chapter, the focus is on research which has design and develop as the primary research function.

Research designs

The traditional and most widely and commonly used text books on research methodology (in social science and education in particular) present and discuss a number of research approaches or designs (see e.g. Arthur, Waring, Coe, & Hedges, 2012; Cohen, Manion, & Morrison, 2007; Creswell, 2011; Denscombe, 2007; Edmonds & Kennedy, 2013). Usually each research design can be used for realizing more than one research function. Without going into detail here, examples of research designs and their possible research functions are:

Survey	to describe, to compare, to evaluate
Case studies	to describe, to compare, to explain
Experiments	to explain, to compare
Action research	to design/develop a solution to a practical problem
Ethnography	to describe, to explain
Correlational research	to describe, to compare
Evaluation research	to determine the effectiveness of a program.

Textbooks on research methodology usually do not present and discuss design research, probably due to its recently emerging status and the fact that relatively small groups across several disciplines have been responsible for its development.

Design research	to design and develop an intervention (such as programs, teaching-learning strategies and materials, products and systems) as a solution to a complex educational problem as well as to advance our knowledge about the characteristics of these interventions and the processes to design and develop them, or alternatively to design and develop educational interventions (about for example, learning processes, learning environments and the like) with the purpose to develop or validate theories.

It should be noticed that in a research project often more than one research design needs to be applied. For example, if there is a need to compare how well Chinese grade 8 pupils perform in mathematics as compared to a number of other countries, the primary research function is to compare, leading in this case to a survey as the best research design. However, as part of the development of a valid and reliable mathematics test, the researchers may do correlational research to determine whether the test being developed is valid, i.e. whether it correlates significantly with other measures of mathematics achievement.

As a final remark, it is important that design researchers, like all researchers, should keep in mind that the guiding principles for scientific research (Shavelson & Towne, 2002) also apply for their research, namely:

- pose significant questions that can be investigated
- link research to relevant theory
- use methods that permit direct investigation of the question
- provide a coherent and explicit chain of reasoning
- replicate and generalize across studies
- disclose research to encourage professional scrutiny and critique.

What is design research?

In this section, a formal definition of educational design research is firstly given, at which a distinction is made between development studies and validation studies. Then some characteristics of educational design research are presented, followed by a brief discussion of the twofold yield of all design research and of the differentiation between possible foci and functions of design research.

Definition of educational design research

As stated at the beginning of the chapter and in the previous section, two possible purposes of design research can be identified, and dependent on the purpose of the research we may distinguish between development studies and validation studies respectively.[3]

In the case of ***development studies***, the purpose of educational design research is to develop research-based solutions for complex problems in educational practice. This type of design research is defined as *the systematic analysis, design and evaluation of educational interventions with the dual aim of generating research-based solutions for complex problems in educational practice, and advancing our knowledge about the characteristics of these interventions and the processes of designing and developing them.*

On the other hand, in ***validation studies*** the purpose of design research is the development or validation of a theory, and this type design research is defined as *the study of educational interventions (such as learning processes, learning environments and the like) with the purpose to develop or validate theories about such processes and how these can be designed.*

It is important to note that design research encompasses systematic educational design processes, but that the reverse is not true: not all systematic educational design can be called research. The second aim of contributing to the body of scientific knowledge (for development studies) or to generate or validate theories (for validation studies) distinguishes design research from just systematic educational design processes.

3) McKenney and Reeves (2012) characterize this difference between the two purposes as research *on* interventions and research through interventions. Van den Akker (personal communication) characterizes this difference by labeling design research focusing on the development of solutions to complex problems as 'research-based design', and the studies aiming at developing and validating (local) theories 'design-based research'.

The twofold yield of design research, namely, research-based interventions as well as knowledge about them, or theories based on them, can also be found in definitions of design research by other authors. For example, Barab and Squire's broad definition (2004, p. 2) also encompasses these variations of educational design(-based) research when they state that

> "Design-based research is not so much an approach as it is a series of approaches, with the intent of producing new theories, artefacts, and practices that account for and potentially impact learning and teaching in naturalistic setting".

Cycles, phases and other characteristics

By its nature, design research is relevant for educational practice (and therefore also for educational policy) as it aims to develop research-based solutions for complex problems in educational practice or to develop or validate theories about processes of learning and teaching.

Whatever the purpose of design research, the research process always incorporates systematic educational design processes, as illustrated in Figure 1.

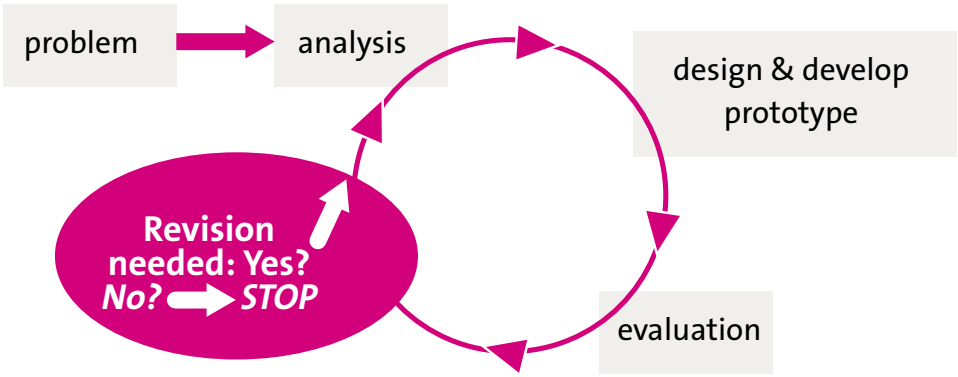


Figure 1: Iterations of systematic design cycles

It is therefore – like all systematic educational and instructional design processes - cyclical in character: analysis, design, evaluation and revision activities are iterated until an appropriate balance between ideals ('the intended') and realization has been achieved[4]. This process can be illustrated in various ways with a few examples presented here showing how different authors have visualized the design research process.

4) See for example Gustafson and Branch (2002) for a taxonomy of instructional development models based on selected characteristics; they distinguish between models with a classroom orientation, product orientation and system orientation.

Reeves (2006) depicts the design research approach as follows:

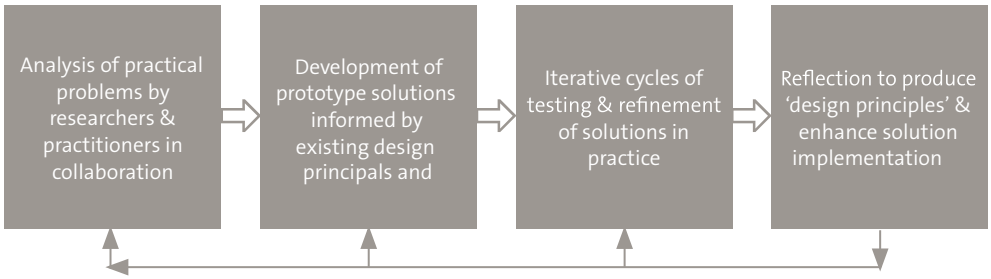


Figure 2: Refinement of problems, solutions, methods, and design principles (Reeves, 2000, 2006)

McKenney (2001) illustrates this cyclical process as follows:

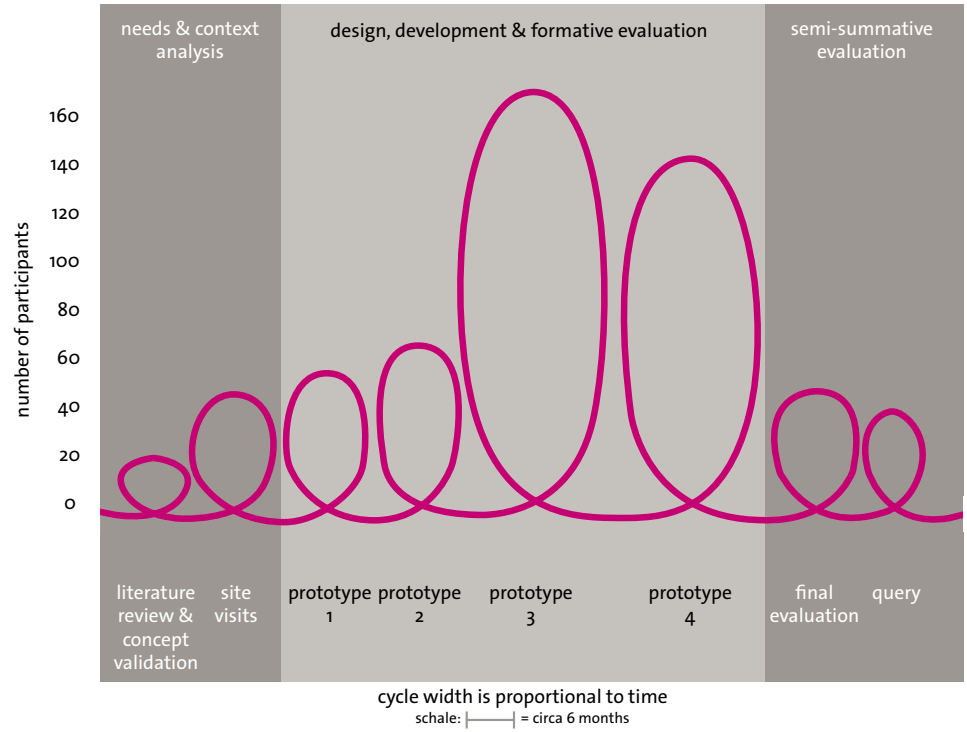


Figure 3: Display of the cascade-sea study (McKenney, 2001)

NOTE: the 'query' as the last phase in McKenney's display can be interpreted as the reflection box in the model of Reeves (Figure 2).

Another example is the Integrative Learning Design Framework that Bannan presents in chapter 4 of this book (see also Bannan-Ritland, 2003).

Authors may vary in the details of how they picture design research, but they all agree that design research consists of a number of phases. In the case of design research as development studies, the following phases are distinguished:

- *preliminary research:* needs and context analysis, review of literature, development of a conceptual or theoretical framework for the study
- *development or prototyping phase:* iterative design phase[5] consisting of iterations, each being a micro-cycle of research[6] with formative evaluation as the most important research activity aimed at improving and refining the intervention
- *assessment phase:* (semi-) summative evaluation to conclude whether the solution or intervention meets the pre-determined specifications. As also this phase often results in recommendations for improvement of the intervention, we call this phase semi-summative.

Throughout all these activities, the researcher or research group will do systematic reflection and documentation to produce the theories or design principles (a concept taken from van den Akker, 1999 – see also later on, the sub-section on development studies) as the scientific yield from the research. One may state that this systematic reflection and documentation ensures that research-based design and development of an intervention emerges as design research.

Similar research phases are found in validation studies. For example, Cobb et al. (2003) distinguish between the phases of preparing for a design experiment, conducting a design experiment and conducting a retrospective analysis.
Refer to the sub-section on validation studies and also Gravemeijer and Cobb in chapter 3 of this book.

5) it is possible that the design/development component in a such a research project will not begin from scratch but with the evaluation of an existing intervention with the aim of identifying the need for improvement, which then is followed by re-design and a number of design cycles.
6) term taken from Bannan, chapter 4

Design research authors also agree on a number of characteristics represented in this type of research. Van den Akker et al. (2006, p.5) mention the following:

Interventionist	the research aims at designing an intervention in a real world setting
Iterative	the research incorporates cycles of analysis, design and development, evaluation, and revision
Process oriented	the focus is on understanding and improving interventions (a black box model of input – output measurement is avoided)
Utility oriented	the merit of a design is measured, in part by its practicality for users in real contexts
Theory oriented	the design is (at least partly) based on a conceptual framework and upon theoretical propositions, whilst the systematic evaluation of consecutive prototypes of the intervention contributes to theory building.

With reference to several authors, such as van den Akker (1999), Kelly (2006) and Nieveen (1999), another characteristic can be added, namely

Involvement of practitioners	the research involves active participation or collaboration with practitioners in the various stages and activities of the research - this will increase the chance that the intervention will indeed become relevant and practical for the educational context which increases the probability for a successful implementation.

Shavelson et al. (2003, p.26) have suggested another characteristic of design studies, namely that they "are often multileveled in that they link classroom practices to events or structures in the school, district, and community".

The features and characteristics of design research are nicely captured by Wademan (2005) in what he calls the Generic Design Research Model (Figure 4). His model clearly illustrates that the 'successive approximation of practical products' (referred to as 'interventions') is working hand-in-hand with the 'successive approximation of theory' (which he also calls 'design principles').

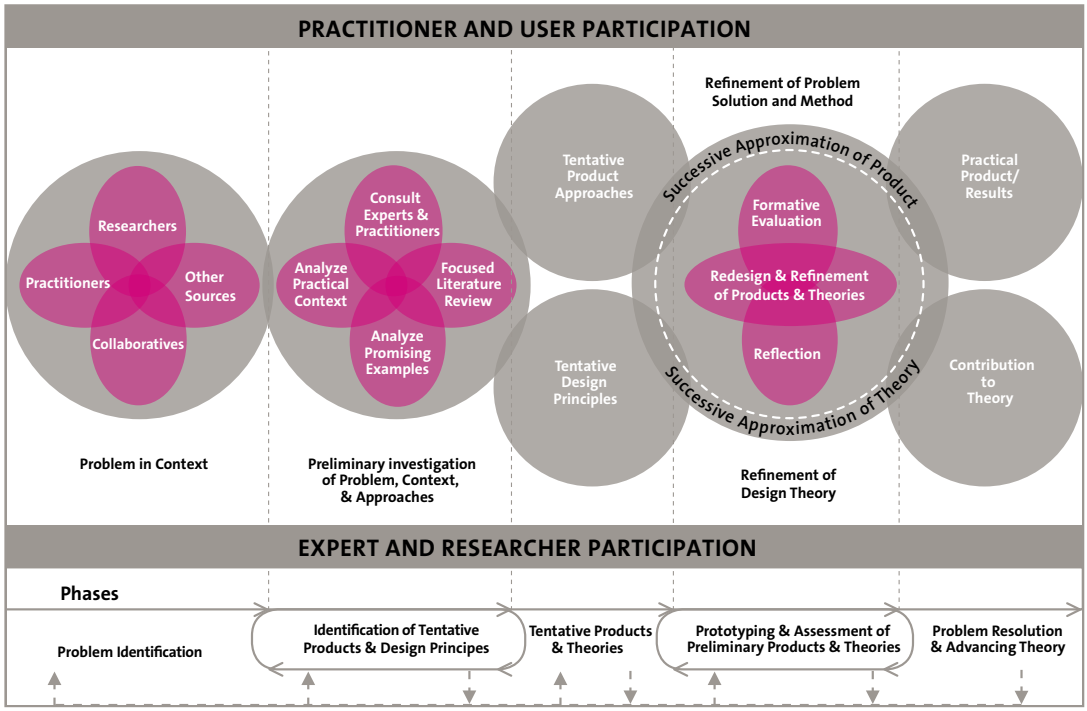


Figure 4: Generic design research model (Wademan, 2005)

Cobb et al. (2003, p.9) express a similar line of thought by stating: "Prototypically, design experiments entail both "engineering" particular forms of learning and systematically studying those forms of learning within the context defined by the means of supporting them. This designed context is subject to test and revision, and the successive iterations that result play a role similar to that of systematic variation in experiment."

It is important to note that design research follows a holistic approach, and does not emphasize isolated variables. Van den Akker et al. (2006, p. 5) point out that "While design researchers do focus on specific objects and processes (interventions) in specific contexts, they try to study those as integral and meaningful phenomena. This context bound nature of much design research also explains why it usually does not strive towards context-free generalizations". If there is a need to make a generalization, then it is an analytical generalization - in contrast to statistical generalization where the researcher strives for generalizing from sample to population. This will be further elaborated on in the section 'Generalizability in Design Research'.

Differentiation in design research

The possible variations in goals of design research have already been discussed and we have differentiated between development studies versus validation studies (see also van den Akker et al., 2006; Chapter 10). However, a further differentiation of design research is conceivable. For example, one can imagine that the implementation and/or dissemination of a particular program is supported by design research – the resulting intervention is the successfully disseminated and implemented program, whilst the systematic reflection and documentation of the process leads to a set of procedures and conditions for successful dissemination and implementation (the design principles).
In this sub-section, first a closer look at the twofold yield of design research, namely, research-based interventions and theoretical yield is given. After that, the differentiation of design research in development studies and validation studies is elaborated.

Twofold yield of design research

Design research is research and therefore an appropriate yield for design research - apart from an usable and effective intervention - is contributing to the body of knowledge in the field. In other words, the challenge for design research is to capture and make explicit the implicit decisions associated with a design process, and to transform them into guidelines for addressing educational problems (see Edelson, 2006; also Barab & Squire, 2004; and many other authors). This aspect refers to theory orientation, mentioned above as one of the characteristics of design research. Van den Akker (1999; also chapter 2 in this book), Reeves (2006; see Figure 2 above) and Wademan (2005; see Figure 4 above) use the concept of 'design principles' when they refer to the theoretical yields of design research, where others speak of new theories (e.g. Cobb et al., 2004; Barab & Squire, 2003; Edelson, 2006; Gravemeijer & Cobb in chapter 3 of this book). This variation is schematically captured in the overview in Table 1.

Table 1: The twofold yield of design research

Type of study:	Research goal:	Twofold yield:
Development studies	Development of intervention:	(i) developing a research based intervention as solution to complex problem, *and* (ii) constructing (re-usable) design principles
Validation studies	Theory development and/or validation:	(i) designing learning environments *with the purpose* (ii) to develop and validate theories about learning, learning environments, or to validate design principles
Implementation studies	Implementation or up-scaling:	(i) implementing a particular program *and* (ii) strategy and conditions under which implementation can happen (design principles).

On development studies

As stated, the starting point for the development studies type of design research is the identification of educational problems for which no or only a few validated principles ('how to do' guidelines or heuristics) are available to structure and support the design and development activities[7].

Many examples of the need for innovative interventions can be given at system level and institutional level. At system level, for example, one may want to develop a system for e-learning to serve a specific target group of students in higher education, or at the level of school or classroom, one may want, for example, to address the question of identifying effective methods for collaborative learning.

A characteristic of development studies is that - informed by prior research and review of relevant literature - researchers in collaboration with practitioners design and develop workable and effective interventions by carefully studying successive versions (or prototypes) of interventions in their target contexts, and in so doing, they reflect on their research process with the purpose of producing as a yield, design principles for developing innovative interventions that are relevant for educational practice (see also Figure 4).

7) See also the chapter of Kelly in this book where he discusses when design research is appropriate.

Two main types of design principles can be distinguished (van den Akker, 1999):
1. procedural design principles: characteristics of the design approach
2. substantive design principles: characteristics of the design (= intervention) itself.

Design principles are heuristic statements in the meaning of experience-based suggestions for addressing problems (such as the ones in design research). Heuristics are always developed in a certain context and provide therefore no guarantee for success in other contexts. However, they will be increasingly powerful when validated in various contexts (see also below, the section on Generalizability in design research).

Van den Akker (1999, p.9) developed the following format for design principles:
"If you want to design <intervention X> for the <purpose/function Y> in <context Z>, then you are best advised to give <that intervention> the <characteristics A, B, and C> [substantive emphasis], and to do that via <procedures K, L, and M> [procedural emphasis], because of <arguments P, Q, and R>."

These heuristic principles are meant to support designers in their tasks, but cannot guarantee success. They are intended to assist (in other projects) in selecting and applying the most appropriate (procedural and substantive) knowledge for specific design and development tasks. Procedural knowledge refers to the set of design activities that are considered most promising in developing an effective and workable intervention. In comparison, substantive knowledge is knowledge about essential characteristics of an intervention and can be extracted (partly) from a resulting intervention itself.
As knowledge is incorporated in interventions, it is profitable for design researchers in the preliminary phase of their research, to search for existing interventions that (although developed in another context) could be considered useful examples or sources of inspiration for the problem at hand. Careful analysis of such examples, in combination with reviewing relevant literature, will generate ideas for the new design task.
The value of knowledge resulting from a design research project will strongly increase when it is justified by theoretical arguments, well-articulated in providing directions, and convincingly backed-up with empirical evidence about the impact of those principles. It is for this reason that authors (e.g. van den Akker 1999; van den Akker et al. 2006; Reeves, 2000, 2006) state that the final stage of each (development study type of) design research project should consist of systematic reflection and documentation to produce design principles.

On validation studies

Validation studies, on the other hand, have a focus on designing for example, learning environments or trajectories with the purpose of developing and validating theories about

the process of learning and how learning environments can be designed. Gravemeijer and Cobb (in chapter 3 of this book) caution usage of the word 'validation' as it may be interpreted as 'checking' and 'confirming'. These "do not fit the exploratory character of our approach [of design research] that aims at creating innovative learning ecologies in order to develop local instruction theories on the one hand, and to study the forms of learning that those learning ecologies are intended to support on the other" (o.c., p.75).

In this section, validation studies is used in this broad sense of design research from the learning design perspective (see Gravemeijer & Cobb, chapter 3 in this book), or – more generally and quoting Cobb et al. (2003, p.10) – as design experiments "to develop a class of theories about both the process of learning and the means that are designed to support that learning, be it the learning of individual students, of a classroom community, of a professional teaching community, or of a school or school district viewed as an organization".

With reference to Cobb et al. (2003), validation studies (in their wording 'design experiments') in the learning sciences are conducted to develop theories on domain-specific learning processes. In this context, they state that "ideally, design experiments result in a greater understanding of a learning ecology – a complex, interacting system involving multiple elements of different types and levels – by designing its elements and by anticipating how these elements function together to support learning" (p.9).

Gravemeijer and Cobb (in chapter 3 of this book) build on the approach of Cobb et al. (2003) by stating that validation studies aim at the development of domain-specific instruction theories (in their case in the domain of mathematics education) at various levels:
- at the level of the instructional activities (micro theories)
- at the level of the instructional sequence (local instruction theories)
- at the level of the domain-specific instruction theory.

In validation studies, researchers do not work in controlled (laboratory or simulated) settings, but choose the natural setting of classrooms as 'test beds' (although they tend to work with above-average numbers of teaching staff). The phases Gravemeijer and Cobb (chapter 3) distinguish in validation studies are:
- preparing for the experiment: drafting a preliminary instructional design and explicating an interpretative framework
- design experiment: trying, elaborating and improving the instructional design or local instructional theory and developing an understanding of how it works

- retrospective analysis: studying the entire data set to contribute to the development of a local instructional theory and more encompassing theories.

As indicated, these phases are similar to the phases that are identified in development studies.

DiSessa and Cobb (2004, p. 83) warn that "design research will not be particularly progressive in the long run if the motivation for conducting experiments is restricted to that of producing domain specific instructional theories". The practical contribution of design research lies in developing empirically-grounded prototypical learning trajectories that may be adopted and adapted by others.

Design research often a combination of development and validation study

It is important to note that the distinction between development and validation studies is conceptually important; however, in practice design researchers may combine the two orientations in their research. For example, starting from a complex and persistent problem for example in mathematics education, the research group may decide to apply the design principles (local theories) resulting from other studies in their research. In doing so, they are not only developing an intervention, but at the same time investigating the validity of design principles (theory) developed in another context for their own problem context.

A nice example of this is the research by Fauzan (2002; see also Fauzan, Plomp,& Gravemeijer, 2013; Part B - chapter 8), whose research can be categorized as a development study type of design research as it aimed at developing a high quality geometry course based on the principles of realistic mathematics education (a constructivist approach to the teaching and learning of mathematics). However, as his research also aimed at validating whether the constructivist approach of realistic mathematics education could be successfully applied in the context of Indonesian mathematics education, this research was a validation study type of design research as well.

The remaining of this chapter discusses a number of aspects important in conducting design research without explicitly referring to the differentiation in design research as discussed in this section. Where Gravemeijer and Cobb discuss design research from the learning perspective (see chapter 3), i.e. aimed at developing and validating local instruction theories by applying design experiments, the remaining of this chapter takes the broad perspective of design research aimed at both developing innovative interventions and identifying underlying design principles or theories.

A closer look at design research

In this section, a closer look is taken at design research by discussing three specific aspects, namely, the overall research question in design research, a set of quality criteria for interventions developed in design research, and a simplified research model for design research.

The overall research question in design research

By now it is clear that designing and developing an intervention is in itself not yet design research. But one may conduct a design/development project as a research project by rigorously employing social science research methodology.
If the research goal is the development of an innovative intervention, then the researcher is striving to find design principles (or an intervention theory) that are valid in a certain context, and the overall research question can be phrased as:

What are the characteristics of an <intervention X> for the purpose/outcome Y in context Z?

Examples of research questions are:

(i) What are the characteristics of an intervention for promoting academic research writing which will best support graduates in education in the proposal stage of their research? (Dowse & Howie, 2013; Part B – chapter 40), and

(ii) What are the characteristics of micro-scale chemistry curriculum materials so that they contribute to the implementation of effective practical work in chemistry teaching in Tanzania schools? (Mafumiko, Voogt, & van den Akker, 2013; Part B – Chapter 28).

Of course, not all researchers use this type of phrasing, but the wording of the main research question in design research always implies a search for characteristics.

If design research aims to develop theory or seeks validation, the main research question also has to express this search for characteristics.
An example is:
What is an adequate learning and teaching strategy for genetics in upper secondary biology education in order to cope with the main difficulties in learning and teaching genetics, and to promote the acquisition of a meaningful and coherent understanding of hereditary phenomena? (Knippels, 2002).

One comment has to made about the need for the phrasing of the research question. Gravemeijer and Cobb (in chapter 3 of this book) argue that the goal of their approach to

design research cannot be captured in a one-sentence research question of the type 'How can we teach a certain topic effectively?' as "such a one-sentence question would have to be complemented with a series of assumptions about what requirements the answer should meet, and would also have to be embedded in an exposition of what is aimed for with the innovation." (o.c., p.75).

Indeed, the goals of design research can be phrased in terms of the outcome the researcher is aiming for. However, the example from Knippels (2002) illustrates that it is possible to phrase a research question that is sufficiently specific for design research that is aimed at theory development or validation.

Quality criteria for interventions

Design researchers aim at interventions of good quality. An obvious criterion for quality is the effectiveness of the intervention: are the intended outcomes attained? However, where effectiveness is the ultimate goal, some other criteria may assist the researchers to optimize the design and development trajectory.

Based on prior work, Nieveen (1999) proposes a number of generic criteria for high quality interventions, namely validity, practicality and effectiveness (see Table 2). She explains these criteria as follows (paraphrased by author):

The intervention should address a need, and its components should be based on state-of-the- art knowledge (content validity, also called relevance) and all components should be consistently linked to each other (construct validity, also called consistency). If the intervention meets these requirements, it is considered to be valid. A second characteristic of high-quality interventions is that teachers (or more general, representatives of the target group of users) consider the intervention to be usable and that it is easy for them to use the intervention in a way that is largely compatible with the developers' intentions. … If these conditions are met, we call these interventions practical. A third characteristic of high quality interventions is that they result in the desired outcomes, i.e. that the intervention is effective" (o.c., p.127).

In addition, Nieveen and Folmer (chapter 6) indicate the importance of the distinction between expected and actual practicality and effectiveness. Only when the target users have had practical experience with using the intervention, should one be able to get data on the actual practicality of the prototype. Similarly, only when target users have had the opportunity to use the intervention in the target setting, should the evaluator get data on the actual effectiveness. In all other instances, such as expert appraisals or a group

discussions based on the materials, the researcher should only get data on the expected practicality and/or expected effectiveness (o.c., p.161). More evaluation will then be needed to demonstrate the actual practicality and the actual effectiveness.

These (amended) criteria are presented in Table 2 – and further details can be found in chapter 6 (by Nieveen & Folmer).

Table 2. Criteria for high quality interventions (adapted from Nieveen, 1999; see also chapter 6)

Criterion	
Relevance (also referred to as **content validity**)	There is a need for the intervention and its design is based on state-of-the-art (scientific) knowledge.
Consistency (also referred to as **construct validity**)	The intervention is 'logically' designed.
Practicality	**Expected** The intervention is expected to be usable in the settings for which it has been designed and developed. **Actual** The intervention is usable in the settings for which it has been designed and developed.
Effectiveness	**Expected** Using the intervention is expected to result in desired outcomes. **Actual** Using the intervention results in desired outcomes.

Given the character of design research, these four criteria suggest a logical hierarchy as can be illustrated by 'rhetoric' questions like:

- if an intervention is *not* practical, why would it make sense to investigate its effectiveness?
- if an intervention is not well designed (i.e. not consistent), why would it make sense to investigate its practicality?

This logical hierarchy implies that the criteria may have different emphases in different phases of the research as is illustrated by Table 3.
For example, during the preliminary research (with the emphasis on analyzing the problem and reviewing the literature), the criterion of relevance (content validity) is the most dominant, with some attention for consistency (construct validity) and practicality, whilst no (or very little) attention is given to effectiveness.

On the other hand, in the initial iterations of the development phase the formative evaluation of the prototypes should focus on consistency and practicality, whilst the criterion of effectiveness will only become increasingly important in later iterations. Finally, in the assessment phase of summative evaluation, the focus will be on practicality and effectiveness.

This shift in emphasis on the quality criteria during the research is summarized in Table 3 (see also chapter 6, Table 1).

Table 3: Evaluation criteria related to phases in design research

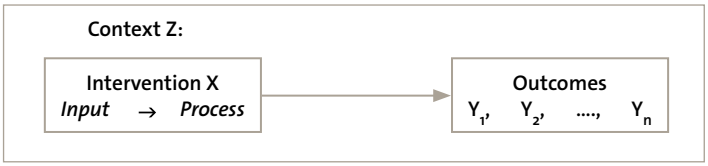
Phase	Criteria	Short description of activities
Preliminary research	Emphasis mainly on *content validity,* not much on *consistency* and *practicality*	Review of the literature and of (past and/or present) projects addressing questions similar to the ones in this study. This results in (guidelines for) a framework and first blueprint for the intervention.
Development or Prototyping phase	Initially: *consistency (construct validity)* and practicality. Later on, mainly practicality and gradually attention for effectiveness.	Development of a sequence of prototypes that will be tried out and revised on the basis of formative evaluations. Early prototypes can be just paper-based for which the formative evaluation takes place via expert judgments resulting in expected practicality (see for an example, Table 4).
Assessment phase	*practicality* and *effectiveness*	Evaluate whether target users can work with intervention (actual practicality) and are willing to apply it in their teaching (relevance & sustainability). Also whether the intervention is effective.

Simplified research model

It is not self-evident how in design research the design of interventions contribute to design principles or theory building. With reference to Wademan's generic model (1995; see also Figure 4), Reeves's exemplary schemes (2006) in Figure 2 and McKenney's display (2001) in Figure 3, we know that the researcher (or the collective of researchers and practitioners) designs and develops (in an iterative way) the intervention with the aim that, after a number of cycles, the intended outcomes are realized, or a well-grounded 'local' theory is developed. Each iteration or cycle is a micro-cycle of research, i.e. a step in the process of doing research and will include systematic reflection on the theoretical aspects or design principles in relationship to the status of the intervention, ultimately resulting in design principles or theoretical statements.
In other words and with reference to the overall research question presented above, the researcher (or research group) will conclude the following about his intervention:

Given my context Z, if I do <intervention (theory based) X > then I expect <intended outcomes Y>.

This can be displayed schematically as:



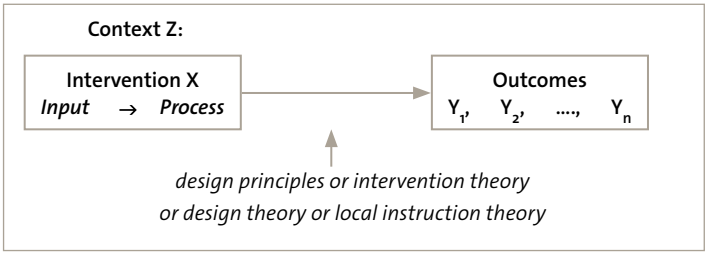
It is important to point out a few key aspects in this scheme. At first, the outcomes of the intervention are indicated as Y_1, Y_2, ..., Y_n, because often an intervention is designed to realize multiple outcomes, for example better achievement, improved student attitude, increased teacher satisfaction.

Another key aspect is that the intervention is presented as 'input → process'. It is crucial that design researchers (or the research group) realize that when designing an innovative process (for example, learning environment) one has to take into account also the inputs or conditions necessary to make the process function (for example, availability of infrastructure, a change in organization of the teaching-learning processes, teacher development). Van den Akker (chapter 2) argues and illustrates that in curriculum design research all the inter-related components of a curriculum need to be taken into account when developing an innovative curriculum. In addition, he shows the importance of a careful implementation of the intended curriculum to ascertain that there will be a good match of the intended curriculum with the implemented curriculum ('what and how it is taught') and the attained curriculum ('learning experiences and outcomes').

The teacher development part (or more general, attention for the user) can be particularly a crucial factor in a successful implementation of an innovative intervention. With reference to the display above, it is suggested that in the early stages of design research the focus is on designing the process with the active involvement of practitioners (i.e. representatives of the future users) involved in or familiar with the vision and ideas underlying the intervention.
And only after it has been proven that it is possible to design an intervention resulting in the desired outcomes, the focus of the research may shift to inputs or conditions necessary for the interventions to function in the intended context. In other words, first a 'proof of existence' should be given, before the focus of the design research shifts to dissemination and implementation, i.e. to preparing and training the intended users of the innovative intervention and assuring that other conditions be fullfilled.

So in the end, the research group has at its disposal not only the intervention resulting in the desired outcomes, but also (based on a systematic reflection and analysis of the data collected during this cyclical process) an understanding of the 'how and why' of the functioning of the intervention in the particular context within which it was developed. The design researcher will summarize this understanding of the 'how and why' of the intervention in one or more 'design principles', using the terminology developed by van den Akker (1999), Nieveen et al. (2006) and Reeves (2000, 2006). As other authors, e.g. Barak and Squire (2004) and Edelson (2006), use 'theory' as the yield of design research, one may also speak of 'intervention theory' or 'design theory' (Wademan, 2005; Figure 4) or – in the case of validation studies – of local instruction theory (see e.g. Gravemeijer & Cobb in chapter 3 in this book).

Schematically this can be represented as:



In design research, interventions are developed in a cyclical process of successive prototypes developed in a number of iterations (see Figure 1).
In the previous section, a set of criteria for good quality interventions has been presented. We argued that these criteria may need different emphases in different phases of the research (see Table 3, above).

A key idea is that when the prototype of the intervention in a certain iteration does not result in the desired outcomes for that iteration, the conclusion is that the intervention is not yet good enough – in other words, the asserted design principles (or intervention theory) for that iteration are not (yet) good enough or not yet emerging. This has to result in a re-design or refinement of the intervention, which goes hand-in-hand with the refinement of the intervention theory or design theory (as is illustrated in Figure 4, the Generic Design Research Model, Wademan, 1995).

After a number of iterations with shifting criteria (see Table 3 above), the researcher (or research group) may conclude that based on the analysis of the evaluation data, the 'realized outcomes' are close enough to the 'intended outcomes' after which he/she may conclude that the design principles appear to be effective. Or, in other words, the researcher (or research group) has developed design principles or a 'local' (intervention) theory for the context in which he/she works:

in context Z the intervention X (with certain characteristics) leads to outcomes Y1, Y2, ..., Yn.

Two examples are given to illustrate this – rather abstract – phrasing of the yield of design research. The Design-Based Learning Research Collective (2003, p. 5) states that "the design of innovations enables us to create learning conditions that learning theory suggests are productive, but that are not commonly practiced or are not well understood".
In other words included in the intervention is knowledge about how to create conditions for learning (the outcomes aimed for).

The second example is taken from science education. Lijnse (1995, p. 192) argues that design research (he calls it developmental research) is "a cyclic process of theoretical reflections, conceptual analysis, small-scale curriculum development, and classroom research of the interaction of teaching-learning processes. The final, empirically based description and justification of these interrelated processes and activities constitutes what we call a possible "didactical structure" for the topic under consideration."
In other words, the local theory consists of a didactical structure for the teaching-learning processes for a certain topic.

Generalizability in design research

Design principles and local (instruction) theories will be additionally powerful if they have been validated in the successful design of more similar interventions in various contexts. Chances for such knowledge growth will increase when design research is conducted in the framework of research programs, because then projects can build upon one another.

Here the question to what extent design principles can be generalized from one context to others is touched on. It is in this context that Edelson (2006) states that design research should result in generalizable theory. Examples have already been mentioned earlier, such as the substantive and procedural design principles in development studies and theories about, for example, learning processes and learning environments in validation studies.

In design research, just as in case studies and experimental studies, the findings cannot be generalized to a larger universe – there is no statistical generalization from sample to population as in the case of survey research. Yin (2003) points out that in case studies and experimental studies, the investigator strives to generalize a particular set of findings or results to a broader theory. This is also the case in design research where the researcher should strive to generalize 'design principles' to some broader theory or widen the domain of validity of the local instruction theory.
Yin (2003) indicates that this generalization is not automatic. Paraphrasing Yin (2003, p.37), design principles and local (instruction theories) must be tested through replications of the findings in a second, third or more cases in various contexts with the purpose of ensuring that the same results should occur. Once such replications have been made, the results might be accepted for a much larger number of similar contexts, even though further replications have not been performed. This *replication logic* is identical to that underlying the use of experiments and allows experimental scientists to generalize from one experiment to another. Yin (2003) calls this *analytical generalizability*.

But a warning should be given here. Where design principles may have been supported by a number of replications, and a new context may be similar to the ones from which design principles have emerged, each context has unique characteristics that justify the use of design principles as 'heuristic' statements: these statements provide guidance and direction, but do not give 'certainties'. The same applies for local theories. It is in this context that Reeves (2006) cites Lee Cronbach one of the most influential educational researchers of the 20th century: *"When we give proper weight to local conditions, any generalization is a working hypothesis, not a conclusion" (Cronbach, 1975, p. 125).*

Some aspects of conducting design research

This section discusses some aspects of conducting design research focusing mainly at a conceptual level. Please refer to Nieveen and Folmer (chapter 6) for a more in-depth and elaborate discussion.

Design research is conducted iteratively in collaboration with researchers and practitioners in a real-world setting. The two principle outputs (design principles or local theories, and empirically underpinned innovative interventions) can be realized if the research comprises the relevant phases of preliminary research, development or prototyping phase and assessment phase. Doing research in such a setting is challenging and demands a careful research design. It is therefore important to reflect not only on the cyclical, iterative character of the systematic design of the intervention, but also – because it is research – to follow the guiding principles for scientific research as stated by Shavelson and Towne (2002 – see above).
It is beyond the scope of this chapter to discuss how to conduct the research in these phases in detail – this is the focus of chapter 6 by Nieveen and Folmer; but an exception is made for two conceptually important topics, namely, formative evaluation as the key research activity in design research, and how cycles (or iterations) in design research can become micro-cycles of research.

Formative evaluation in design research

Formative evaluation, that is evaluation aimed at improvement, takes place in all phases and iterative cycles of design research. As illustrated in Table 3 (above), formative evaluation serves different functions, or – in other words – is aimed at different quality criteria (or combinations of these) in the various development cycles, each being a micro-cycle of research with its specific research/evaluation question and related research/evaluation design. One may say that formative evaluation has various layers in a design research project which is illustrated in Figure 5, taken from Tessmer (1993): from more informal in the early stages of a project (self-evaluation, one-to-one evaluation, expert review) to small group evaluation aimed at testing the practicality and effectiveness, to a full field test (if possible). The research/evaluation design for each cycle should reflect the specific focus and character of the cycle – see next sub-section for examples and chapter 6 by Nieveen and Folmer for more details.

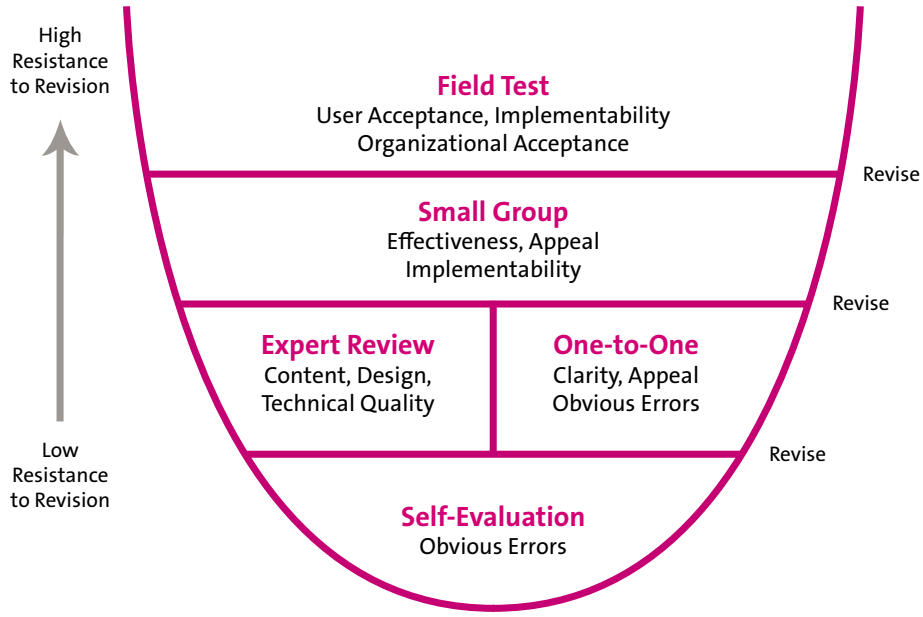


Figure 5: Layers of formative evaluation (taken from Tessmer, 1993)

Figure 5 also illustrates that many possible methods of formative evaluation can be chosen, such as[8]

- expert review and/or focus groups (important to consider 'experts in what')
- self-evaluation or screening (using check list of important characteristics or design specifications)
- one-to-one evaluation or walk through (with a representative of target audience)
- small group or microevaluation
- field test or try-out.

Of course, dependent on their specific research goals, design researchers should choose formative evaluation designs and approaches that are suitable for the purpose of that particular stage of the research for each phase and for each prototype of their intervention – as is illustrated in the first example in the next sub-section.

8) see Chapter 6 by Nieveen and Folmer for more methods and references.

Design stages as micro-cycles of research

As stated, design research has to meet criteria for good research. It is therefore important that for each cycle in the development phase the researcher (or research group) applies the methodological 'rules' for doing research. This means each cycle should begin with one or a few research/evaluation questions reflecting the quality criteria that will be emphasized in that cycle, which then leads into the design and development of an appropriate research/evaluation design. Given the layers of formative evaluation in design research (see above, Figure 5), the evaluation design can be less rigorous in the early cycles of development with more emphasis being placed in later phases or cycles.

Nieveen and Folmer (chapter 6) present a so-called 'evaluation matchboard' as a tool for design researchers for designing and planning their formative evaluations. They discuss how this tool can be used to support design researchers in each of the research cycles beginning with the relevant quality criteria (see Table 2), in proper phrasing of the research/evaluation questions, in identifying relevant respondents, in determining proper data collection methods, and so on.

In this section just three examples of how cycles in design research have been realized are presented to illustrate various aspects of designing and structuring design research studies.

NOTE: the examples do not give prescriptions; they just illustrate aspects of the research designs that the researchers drew up to address their specific research questions.

Example 1: on stages – cycles - iterations

Nieveen (1999; 2013, i.e. Part B – chapter 51) conducted design research to develop a computer assisted support system for curriculum development. Table 4 shows how she has divided the development phase in a number of stages. It also shows how in these stages various formative evaluation methods have been used with purposive samples of respondents for the respective prototypes to address the quality criteria in this project.

Table 4: Focus of design and formative evaluation of the prototypes for computer assisted support system for curriculum development (adapted from Nieveen, 1999)

		Development or prototyping phase						
	Stages in prototype development →	preliminary computer-based version	paper-based version		computer-based versions		final version	
		Users (n=5)	experts (n=3)	users (n=5)	experts (n=6)	users (n=4)	users (n=4)	users (n=17)
Validity	content *)		√ ea		√ ea			
	interface				√ ea			
Practicality	content	√ wt		√ wt	√ ea	√ me	√ to	√ ft
	interface	√ wt		√ wt	√ ea	√ me	√ to	√ ft
Effectiveness	entire system						√ to	√ ft

*): Content refers to the content of the support system

√ = primary attention of prototype and of formative evaluation

Methods of formative evaluation: me = micro evaluation; wt = walk through; ea = expert appraisal; ft = field trial; to = try-out

Relevant in the example in Table 4 is that the stages 'paper-based versions' and 'computer-based versions' have a cycle with experts as well as one with intended users. It illustrates that one may have in design research a number of cycles in a stage. Moreover, each cycle may have more than one iterations. For example, the first 'computer-based version' was evaluated using expert appraisal, then revised and again evaluated via expert appraisal resulting after a few iterations in a version with expected practicality. That version has been evaluated with a purposive sample of intended users of the intervention resulting (after a few iterations) in a version with actual practicality.

Example 2: Expected versus actual practicality or effectiveness

When introducing earlier in this chapter the quality criteria for interventions, a distinction has been made between expected and actual practicality and effectiveness. It may occur in certain studies that the researcher (or research collaborative) cannot do a final field trial of the intervention with the full (or a sample of the) target group, but has to restrict himself to expert appraisal and/or micro-evaluation of the final prototype of the intervention. It is obvious that in such a situation the actual practicality and the actual effectiveness of the intervention cannot be demonstrated, but only conclusions about the expected practicality and the expected effectiveness can be drawn. More evaluation will then be needed to demonstrate the actual practicality and the actual effectiveness.
This can be illustrated with an example adapted from Mafumiko (2006; see also Mafumiko, Voogt, & van den Akker, 2013, Part B – chapter 28) who conducted design research to investigate whether micro–scale experimentation can contribute to improving the chemistry curriculum in Tanzania. His research model is summarized in Figure 6.

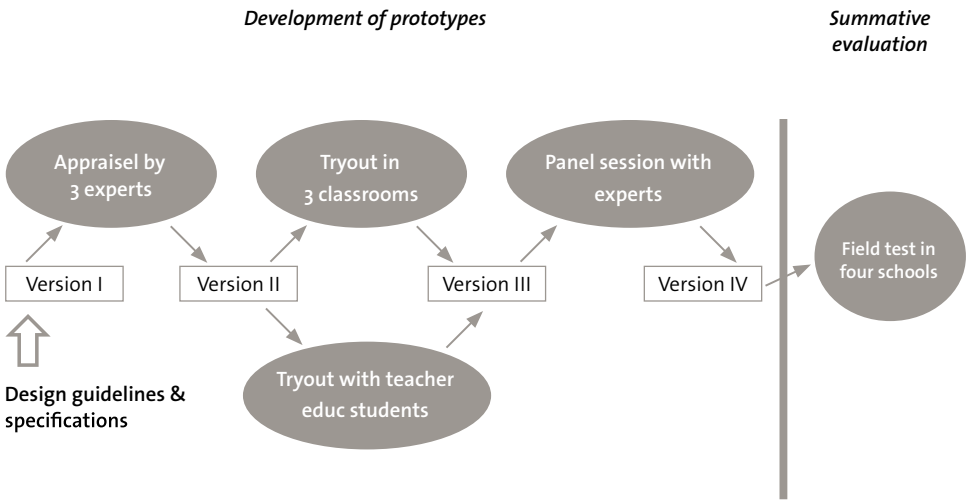


Figure 6: Example of a design research research model (adapted from Mafumiko, 2006)

Suppose a researcher would restrict himself to the development of prototypes of the intervention as illustrated in Figure 6, and does not plan to investigate whether Version IV works in the target context. In such a situation the most he can conclude is whether his intervention is *expected* to be practical and effective for the target context. Only when he would conduct a field test, he will be in the position to decide upon *actual* practicality and actual effectiveness (which is what Mafumiko did).

Example 3: Cycles as micro-cycles of research

Figure 7 shows an example of how cycles in design research can be designed as micro-cycles of research: the Design Research Model developed by Dowse (in press; see also Dowse and Howie, 2013, i.e. Part B - chapter 40). The research question in this study was 'What are the characteristics of an intervention for promoting academic research writing which will best support master's students in education in the proposal stage of their research?' This research model has been developed drawing on guidelines presented in Nieveen (2009) and Plomp (2009)[9].

Figure 7 shows the main research question, the three phases in the research, and eight operational cycles within these phases[10]. The criteria for good quality interventions (presented in Table 2) are emphasised in the respective phases and are operationalized in specific research questions to be addressed in the various cycles of research. The research model also indicates how this research has been operationalized (see Dowse and Howie, 2013, i.e. Part B - chapter 40 for further explanation and details).

Note: as stated above, the model of Dowse is an *example* of how in a study the micro-cycles of research can be drawn up. The exemplary character can be illustrated by pointing to the way Dowse applies the quality criteria for interventions: her model presents one criteria per cycle in her research, whilst it is suggested in Table 3 that the criteria may have different emphases in different phases of the research (as is illustrated in Table 3 and chapter 6, Table 2).

9) The revised version of these chapters are the Chapters 1 and 6 in this book.

10) Note that 'cycle' is used in the meaning of a distinguisable part of the research project, and not in the meaning of full cycle in a systemic design/development/evaluation process.

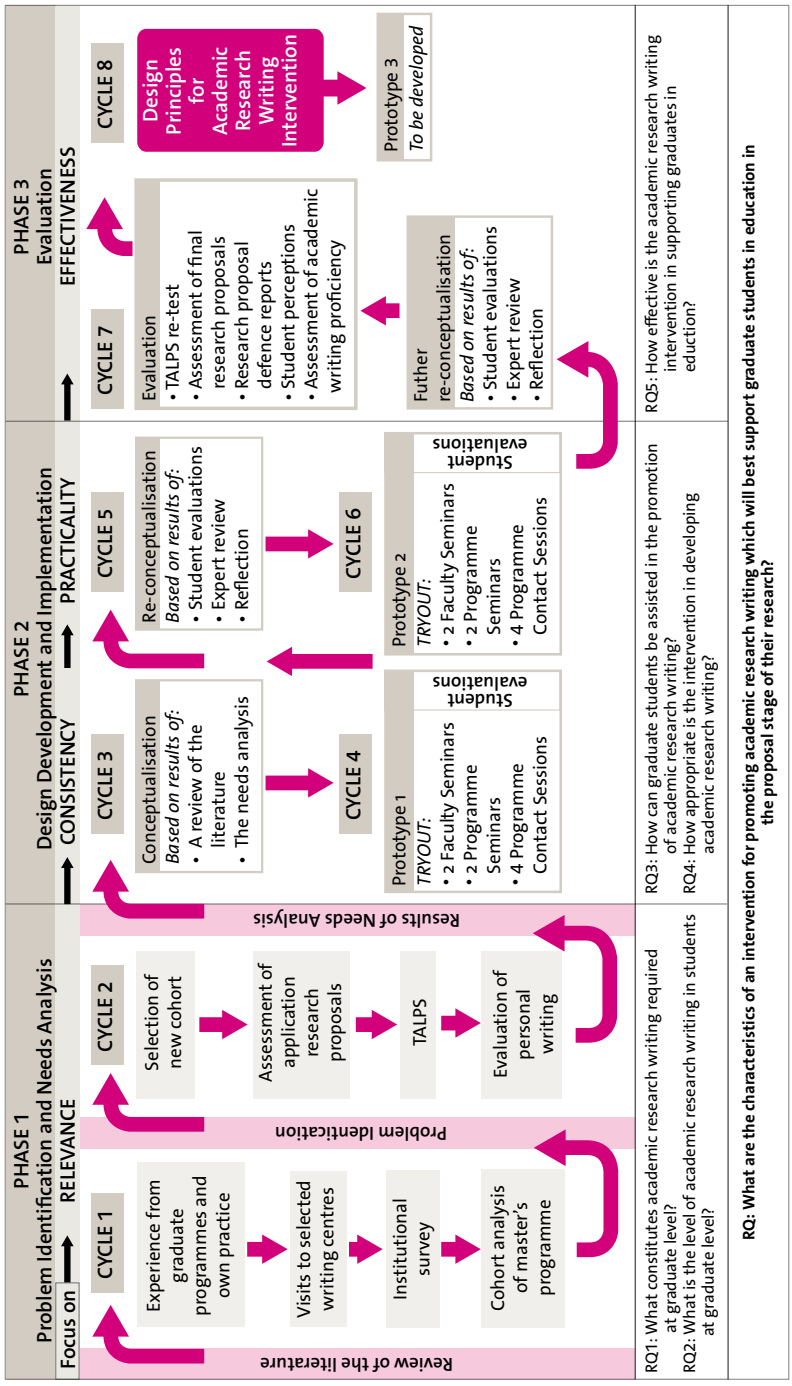


Figure 7: Design Research Model for the development of an academic research writing intervention (Dowse, in press; also Dowse and Howie (2013; see chapter 40, Part B of this book))

Design research challenges

Design research is conducted in close collaboration with educational practice. Not only is the problem, situated in educational practice, addressed , but a key feature of this research is that educational practitioners are actively involved, often as members of the research team.

This leads to a number of challenges that are typical for this type of research. McKenney, Nieveen, and van den Akker (2006, p. 83, 84) have discussed some of these and provide suggestions for how to address them. Their points are briefly summarized here.

The researcher is designer and often also evaluator and implementer

Several measures can be taken to compensate for these potential conflicts of interest:

- make research open to professional scrutiny and critique by people outside the project
- the researcher applies the following rule of thumb: shift from a dominance of 'creative designer' perspective in the early stage, towards the 'critical researcher' perspective in later stages (this is reflected in Tessmer's layers of formative evaluation, Figure 5)
- have a good quality research design, for example:
  - *strong chain of reasoning* (Krathwohl, 1998) - the metaphor expresses the idea that each part of the research design should be equally strong
  - *triangulation* – to increase the quality of data and of analysis triangulation of data sources and data collection methods should be applied, as well as investigator triangulation to avoid the influence of any specific researcher (see e.g. Denscombe, 2007, p. 136)
  - *empirical testing* of both the practicality and the effectiveness of the intervention
  - *systematic documentation*, *analysis* and *reflection* of the design, development, evaluation and implementation process and their results
- pay attention to validity and reliability of data and instruments
- apply a variety of methods and tactics: for example, use practitioners and other researchers as 'critical friends'; use multiple observers/raters and calculate inter-observer/rater reliability, etc.

Real-world settings bring real-world complications

Design research is conducted in real-world settings because it addresses complex problems in educational practice. One of the problems is that the researcher can be a 'cultural stranger' (Thijs, 1999) in the setting of the research and that participants (for example, principals or teachers not involved in the research, etc) are hesitant to be completely open to a researcher coming from the outside.

McKenney et al. (2006) point to the importance of collaboration and mutual beneficial activities to gain participants' trust and thorough understanding of the context (i.e. insider perspective). On the other hand, they also point to the advantages of being an outsider as this may allow the researcher to develop a degree of objectivity and "freedom (or forgiveness) for honesty that is not permitted to those within a particular group" (p. 85).

Adaptability of the research design

Design research is cyclical and takes place in real-world settings. Each cycle has to take the findings of the previous ones into account. On the one hand, the research design has to change (or develop) from one cycle to the other, but on the other hand, an ever-changing research design can be weak. In this context, McKenney et al. (2006, p. 84) refer to the notion of evolutionary planning, i.e. "a planning framework that is responsive to field data and experiences at acceptable moments during the course of the study". This is already alluded to in the discussion of formative evaluation (see Tessmer,1993, Figure 5).

The need for adaptability pertains also to the role of the researcher. According to van den Akker (2005, in McKenney et al., 2006), the synergy between research and practice can be maximized when researchers demonstrate adaptability by:

- being prepared, where desirable, to take on the additional role of designer, advisor, and facilitator, without losing sight of their primary role as researcher
- being tolerant with regard to the often unavoidably blurred role distinctions and remaining open to adjustments in the research design if project process so dictates
- allowing the study to be influenced, in part, by the needs and wishes of the partners, during what is usually a long-term collaborative relationship.

"Such adaptability requires strong organizational and communicative capabilities on behalf of the researcher. Adaptability also requires sound understanding of research rigor so that prudent changes and choices that maximize value and minimize threats to quality are made." (McKenney et al., 2006, p. 84).

To address the challenges mentioned, McKenney et al. (2006, p. 85, 86) present a few guidelines for conducting design research that may help researchers in monitoring the scientific character of his/her research:

- have an explicit conceptual framework (based on review of literature, interviews of experts, studying other interventions)
- develop a congruent study design, i.e. apply a strong chain of reasoning with each cycle having its research design
- use triangulation (of data source, data type, method, evaluator and theory) to enhance the reliability and internal validity of the findings
- apply both inductive and deductive data analysis

- use full, context-rich descriptions of the context, design decisions and research results
- member check, i.e. take data and interpretations back to the source to increase the internal validity of findings.

As it is beyond the scope of this chapter to elaborate on these guidelines, please refer to McKenney et al. (2006) and other chapters in van den Akker et al. (2006).

Concluding remarks and outline of the book

In the field of education there is much need for research relevant to educational practice. We have argued that for complex practical problems and for research question(s) calling for the design and development of an intervention design research is an appropriate research approach.
In this final section, first the key difference between design research and action research will be presented, followed by the suggestion or a plea for a programmatic approach. Then some suggestions for further readings are presented, followed by an outline of this book

Design research versus action research

Given its focus on practical problems and its nature of conducting the research in a real-world setting with active involvement of practitioners, design research may look like action research. So one may wonder how design research is related to action research. Indeed, action research also deals with real-world problems, is aimed at improving practice, is cyclical in nature and participative (Denscombe, 2007), but the essential difference is that action research is not aimed at generating design principles – it has a particular niche among professionals who want to use research to improve their own practices (p.122).

Strength of a programmatic approach

We discussed how design researchers should strive for generalizable design principles in the meaning of generalizing to a broader theory. When design research is conducted within the framework of a program of research addressing fundamental problems in educational practice, it will result in a specific body of knowledge, namely substantive and procedural design principles that may contribute to improve education.

There is another reason for applying a programmatic approach. Much design research takes place within the framework of graduate research studies or projects. In many graduate studies the primary goal is student's graduation and not - due to the limitation

in time - the design and develop of a comprehensive, well-tested solution to the design and development problem. However, by applying a programmatic approach design research by graduate students may build upon what other students have accomplished and concluded.

Further readings

There is a number of books for further reading about educational design research, such as Kelly, Lesh and Baek (2008), Reinking and Bradley (2007), McKenney and Reeves (2012), Richey and Klein (2007), and van den Akker, Gravemeijer, McKenney, and Nieveen (2006), whilst soon Fishman, Penual, Allen, and Cheng (forthcoming) will appear. Since the first edition of this book was published in 2009, many design research projects have been undertaken and reported. Part B of our book (Plomp & Nieveen, 2013) presents 51 cases of successful design research. In addition, chapter 7 of this book presents a number of references and sources relevant for design research.

In preparing this chapter, van den Akker, Gravemeijer, McKenney, and Nieveen (2006) has been an important source of inspiration and ideas. Their book presents not only a number of approaches to design research in the chapters by Gravemeijer and Cobb (2006)[11], Reeves (2006) and McKenney et al. (2006), but also addresses issues like assessing the quality of design research proposals (chapters by Phillips, 2006, and by Edelson, 2006) and the quality of design research (chapter by Kelly, 2006), forming a rich source for further reflection and elaboration.

Outline of this book

This chapter concludes with a brief outline of the remainder of this book. The book comprises of two parts, namely Part A –Educational design research: An introduction and Part B: Educational design research: Illustrative cases.

Part A presents – after this introduction chapter - design research from various angles: the curriculum perspective by van den Akker in chapter 2, the learning design perspective by Gravemeijer and Cobb in chapter 3, whilst Bannan presents in chapter 4 an illustrated example from the domain of instructional technology, and Kelly discusses in chapter 5 when design research is an appropriate research design. Reference has already been made to chapter 6 by Nieveen and Folmer in which they discuss how formative evaluation in educational design research can be designed and conducted. Finally, chapter 7 presents – without striving for completeness - a number of references and sources on educational design research.

11) An update of this chapter in included as chapter 3 in this book

Part B of the book (Plomp & Nieveen, 2013) presents a collection of 51 cases of successful design research from all over the world and covering all education levels, including a few cases about workplace learning. Each case chapter in Part B presents references to a number of key sources for the research reported in that chapter.
Please refer to **http://international.slo.nl/edr** for more information and access to this book. An additional special characteristic of Part B is that it is a 'supra-book', that is users can make their own book by selecting the case chapters that suit their intended use.

Finally, we hope that this book will inspire many researchers who are not yet familiar with design research as a research design appropriate to develop research-based solutions to complex problems or to develop or validate theories about learning processes, learning environments and the like, to embark on this type of research and contribute to the further development of educational design research.

Acknowledgements: *in preparing this chapter, much use has been made of van den Akker, Gravemeijer, McKenney and Nieveen (2006). I am very grateful to co-editor Nienke Nieveen and to Sarah Howie and Cilla Dowes for their constructive comments in preparing the revised version of the chapter.*

References

Anderson, T., & Shattuck, J. (2010). Design-based research: A decade of progress in education research? *Educational Researcher, 41(1)*, 16-25.

Arthur, J., Waring, M., Coe, R., & Hedges, L.V. (2012). *Research methods & methodologies in education*. Los Angeles: Sage.

Bannan-Ritland, B. (2003). The role of design in research: The integrative learning design framework. *Educational Researcher, 32(1*), 21-24.

Barab, S., & Squire, K. (2004). Design-based research: Putting a stake in the ground. *Journal of the Learning Sciences,* 13(1), 1-14.

Brown, A.L. (1992). Design experiments: Theoretical and methodological challenges in creating complex interventions in classroom settings. *The Journal of the Learning Sciences, 2(2)*, 141-178.

Burkhardt, H. (2006). From design research to large-scale impact: Engineering research in education. In J. van den Akker, K. Gravemeijer, S. McKenney, & N. Nieveen (Eds.), *Educational design research* (pp. 121-150). London: Routledge.

Cobb, P., Confrey, J., diSessa, A., Lehrer, R., & Schauble, L. (2003). Design experiments in educational research. *Educational Researcher,* 32(1), 9-13.

Cohen, L., Manion, L., & Morrison, K. (2013, 7th ed.). *Research methods in education*. London: Taylor & Francis Group.

Creswell, J.W. (2011, 4th ed.). *Educational research: Planning, conducting, and evaluating quantitative and qualitative research.* Upper Saddle River (NJ, USA): Pearson Education International.

Cronbach, L.J. (1975). Beyond the two disciplines of scientific psychology. *American Psychologist,* 30(2), 116-27.

Denscombe, M. (2007, 3rd ed.). *The good research guide for small-scale social research projects.* Buckingham, UK: Open University Press.

Design-Based Research Collective. (2003). Design-based research: An emerging paradigm for educational inquiry. *Educational Researcher, 32(1)*, 5-8.

diSessa, A.A., & Cobb, P. (2004). Ontological innovation and the role of theory in design experiments. *Educational Researcher, 32 (1*), 77-103.

Dowse, C. (in press). *Learning to write by writing to learn: A postgraduate intervention to develop academic research writing*. PhD Assessment and Quality Assurance, University of Pretoria, Pretoria, South Africa.

Dowse, C., & Howie, S. (2013). Promoting academic research writing with South African master's students in the field of education. In T. Plomp, & N. Nieveen (Eds.), *Educational design research – Part B: Illustrative cases* (pp. 851-879). Enschede, the Netherlands: SLO.

Edelson, D.C. (2006). Balancing innovation and risk: Assessing design research proposals. In J. van den Akker, K. Gravemeijer, S. McKenney, & N. Nieveen (Eds.). (2006). *Educational design research* (pp. 100-106). London: Routledge.

Edmonds, W.A., & Kennedy, T.D. (2013). *An applied reference guide to research designs.* Los Angeles: Sage.

Eilks, I., & Ralle, B. (2002). Participatory action research in chemical education. In B. Ralle & I. Eilks (Eds.), *Research in chemical education - What does this mean?* (pp. 87-98). Aachen, Shaker. (free access at www.chemiedidaktik.uni-bremen.de_eng/publikationen.php)

Fauzan, A. (2002). *Applying realistic mathematics education (RME) for teaching geometry in Indonesian primary schools.* Doctoral Dissertation. Enschede: University of Twente. Available from: http://purl.org/utwente/58707

Fauzan, A., Plomp, T., & Gravemeijer, K. (2013). The development of an rme-based geometry course for Indonesian primary schools. In T. Plomp, & N. Nieveen (Eds.), *Educational design research – Part B: Illustrative cases* (pp. 159-178). Enschede, the Netherlands: SLO.

Fishman, B. J., Penuel, W. R., Allen, A.-R., & Cheng, B. H. (Eds.). (In press). Design-based implementation research: Theories, methods, and exemplars. *National Society for the Study of Education Yearbook.* New York, NY: Teachers College Record.

Gravemeijer, K., & Cobb, P. (2006). Design research from a learning design perspective. In J. van den Akker, K. Gravemeijer, S. McKenney, & N. Nieveen (Eds.). (2006). *Educational design research* (pp. 17-51). London: Routledge.

Gustafson, K.L., & Branch, R.M. (2002). *Survey of instructional development models* (4th ed). Syracuse, NY: ERIC Clearinghouse on Information & Technology.

Kelly, A.E. (2006). Quality criteria for design research: Evidence and commitments. In J. van den Akker, K. Gravemeijer, S. McKenney, & N. Nieveen (Eds.), *Educational design research* (pp.107-118). London: Routledge.

Kelly, A.E., Lesh, R.A., & Baek, J.Y. (Eds.). (2008). *Handbook of design research methods in education innovations in science, technology, engineering, and mathematics learning and teaching.* New York: Lawrence Erlbaum Associates.

Knippels, M.C.P.J. (2002). *Coping with the abstract and complex nature of genetics in biology education.* Doctoral dissertation. Utrecht, The Netherlands: University of Utrecht – CD-Bèta Press.

Krathwohl, D.R. (1998, 2nd ed.). *Methods of educational and social science research: An integrated approach.* New York: Longman.

Lijnse, P.L. (1995). "Developmental research" as a way to an empirically based "didactical structure" of science. *Science Education, 29(2)*, 189-199.

Mafumiko, F. S. M. (2006). *Micro-scale experimentation as a catalyst for improving the chemistry curriculum in Tanzania.* Doctoral thesis. Enschede, the Netherlands: University of Twente. Available from: http://purl.org/utwente/55448

Mafumiko, F., Voogt, J., & Van den Akker, J. (2013). Design and evaluation of micro-scale chemistry experimentation in Tanzanian schools. In T. Plomp, & N. Nieveen (Eds.), *Educational design research – Part B: Illustrative cases* (pp. 581-600). Enschede, the Netherlands: SLO.

Marks, R., & Eilks, I. (2010). Research-based development of a lesson plan on shower gels and musk fragrances following a socio-critical and problem-oriented approach to chemistry teaching. *Chemistry Education Research and Practice*, 11, 129-141. (free access at http://pubs.rsc.org/en/journals/journalissues/rp)

McKenney, S. (2001). *Computer-based support for science education materials developers in Africa: Exploring potentials*. Doctoral dissertation. Enschede, the Netherlands: University of Twente.

McKenney, S., Nieveen, N., & van den Akker, J. (2006). Design research from a curriculum perspective. In J. van den Akker, K. Gravemeijer, S. McKenney, & N. Nieveen (Eds.), *Educational design research* (pp.62-90). London: Routledge.

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

Nieveen, N. (1999). Prototyping to reach product quality. In J. van den Akker, R.M. Branch, K. Gustafson, N. Nieveen, & T. Plomp (Eds), *Design approaches and tools in education and training* (pp. 125-136). Boston: Kluwer Academic.

Nieveen, N., McKenney, S., & van den Akker, J. (2006). *Educational design research:* The value of variety. In J. van den Akker, K. Gravemeijer, S. McKenney, & N. Nieveen (Eds.), *Educational design research* (pp. 151-158). London: Routledge.

Nieveen, N. (2009). Formative evaluation in educational design. In T. Plomp, & N. Nieveen (Eds.), *An introduction to educational design research* (pp. 89-101). Enschede, the Netherlands: SLO.

Penuel, W.R., Fishman, B.J., Cheng, B.H., & Sabelli, N. (2011). Organizing research and development at the intersection of learning, implementation, and design. *Educational Researcher,* 40(7), 331-337.

Phillips, D.C. (2006). Assessing the quality of design research proposals: Some philosophical perspectives. In J. Van den Akker, K. Gravemeijer, S. McKenney, & N. Nieveen (Eds.), *Educational design research* (pp. 93-99). London: Routledge.

Plomp, T. (2009). Educational design research: An introduction. In T. Plomp, & N. Nieveen (Eds.), *An introduction to educational design research* (pp. 9-35). Enschede, the Netherlands: SLO.

Plomp, T., & Nieveen, N. (Eds.). (2013). *Educational design research: Illustrative cases*. Enschede, the Netherlands: SLO. (free access at www.international.slo.nl)

Reeves, T. (2000). *Enhancing the worth of instructional technology research through "design experiments" and other developmental strategies*. Paper presented at the AERA. Retrieved Oct. 17, 2013 from http://it.coe.uga.edu/~treeves/AERA2000Reeves.pdf

Reeves, T.C. (2006). Design research from a technology perspective. In J. van den Akker, K. Gravemeijer, S. McKenney, & N. Nieveen (Eds.), *Educational design research* (pp. 52-66). London: Routledge.

Reinking, D., & Bradley, B.A. (2008). *On formative and design experiments: Approaches to language and literacy research*. New York & London: Teachers College, Columbia University.

Richey, R., & Klein, J.D. (2007). *Design and development research: methods, strategies, and issues*. London: Routledge.

Shavelson, R.J., & Towne, L. (2002). *Scientific research in education*. Washington DC: National Academic Press.

Shavelson, R.J., Phillips, D.C., Towne, L., & Feuer, M.J. (2003). On the science of education design studies. *Educational Researcher,* 32(1), 25-28.

Tessmer, M. (1993). P*lanning and conducting formative evaluations*. London: Kogan Page.

Thijs, A. (1999). *Supporting science curriculum reform in Botswana: The potential of peer coaching. Doctoral dissertation*. Enschede, the Netherlands: University of Twente.

Van den Akker, J. (1999). Principles and methods of development research. In J. van den Akker, R.M. Branch, K. Gustafson, N. Nieveen, & T. Plomp (Eds.), *Design approaches and tools in education and training* (pp. 1-14). Boston: Kluwer Academic.

Van den Akker, J. (2005). Hoe kan samenwerking leiden tot meer succes en wijsheid in onderwijsontwikkeling? [How can collaboration lead to more success and wisdom in education development?]. *Pedagogische Studiën, 82(4),* 343-7.

Van den Akker, J., Gravemeijer, K, McKenney, S. & Nieveen, N. (Eds.). (2006). *Educational design research*. London: Routledge.

Wademan, M.R. (2005). *Utilizing development research to guide people-capability maturity model adoption considerations.* Doctoral dissertation. Syracuse: Syracuse University. Dissertation Abstracts International, 67-01A, 434. (UMI No. 3205587)

Walker, D. (2006). Toward productive design studies. In J. van den Akker, K. Gravemeijer, S. McKenney, & N. Nieveen (Eds). (2006). *Educational design research* (pp.107-118). London: Routledge.

Yin, R.K. (2003, 3ed ed.). *Case study research: Design and methods*. Newbury Park, CA: Sage.

2. Curricular Development Research as a Specimen of Educational Design Research

Jan van den Akker

Introduction

As noted in the introductory chapter of Tjeerd Plomp (2013) in this book, 'educational design research' is a sort of umbrella term for a number of related research approaches that all (ultimately) aim at contributing to the knowledge base about improving learning and teaching in real life contexts (cf. van den Akker, Gravemeijer, McKenney, & Nieveen, 2006). In the title of this chapter I try to convey multiple conceptual messages:

- The overall perspective and focus is on addressing *curricular* problems, essentially dealing with (changing) aims and contents of learning.
- The main function of the research is to inform and support decision making in the process of curriculum development.
- While 'design' is a critical activity within curriculum development, we use the somewhat broader term '*development* research' to underline the multi-stage character of curriculum development trajectories, where preliminary, analytical studies often preceed design and formative evaluation iterations, afterwards followed by implementation studies (cf. van den Akker, 1999; van den Akker & Kuiper, 2008; van den Akker, Kuiper, & Nieveen, 2012).

Moreover, the emphasis in our approach is better reflected by the wording 'research-based development' than the often used 'design-based research'. The primary function of research is thus supportive.

Having said that, it is realistic to note the worldwide pattern that curriculum development is almost notorious for its weak relationship with research. Socio-political and practical arguments usually dominate curriculum decision-making. Priorities for curriculum projects seldom arise from systematic monitoring and analysis of practices and outcomes. Available research-based knowledge is often insufficiently used during the development process. And empirical information about actual uptake, implementation and large scale impact of curriculum innovations is often lacking. Altogether, one may conclude that curriculum development is hardly an evidence-based enterprise, in contrast to much policy rhetoric. However, probably, few people would argue for an overly strong evidence-based strategy of curriculum development - that would be in contrast to the often value-driven nature of most curricular decision-making. But many would like to experience a stronger research-informed approach of curriculum problems.

This chapter (actually a slightly revised version of my 2009 chapter in the original issue of

this book) will explore how a better cross-fertilisation between educational research and curriculum development may strengthen the information base for curriculum policies and classroom practices. After an effort to articulate our conceptualisation of curriculum and curriculum development, the emphasis of the exploration (building on previous publications, in particular van den Akker, 1999) will be on the potential of curricular development research, an research approach that combines three related goals:

- optimisation of (curricular) interventions/products (e.g. curriculum frameworks, educative materials)
- (curriculum) design principles (as contribution to the knowledge base)
- professional development (of all participants).

The role of research will be outlined for various stages of curriculum development, with particular attention to:

- quality criteria for curriculum interventions
- adequate research methods and procedures
- knowledge growth and generalisation issues.

The problematic fate of most curriculum reforms; may research help?

Curriculum reform has a dubious reputation, with more sobering than real and lasting success stories. One might even say that large-scale curriculum reform has a tendency to fail, as a universal experience (Cuban, 1992; Fullan, 2007; Leyendecker, 2008). Hargreaves and Fink (2006, p.6) put this succinctly: '*Change in education is easy to propose, hard to implement, and extraordinarily difficult to sustain*'. And curriculum changes, it can be argued, belong to the hardest category...

Notwithstanding big investments in research and development and in-service education, the target group of teachers often appears poorly informed about the intended innovation, while its practical application remains limited and its impact on student learning is unclear. Simple explanations for those innovation failures are inadequate, but a few gaps are often visible:

- weak connections between various system levels (national, local, school, classroom)
- lack of internal consistency within the curriculum design
- insufficient cooperation between various actors in educational development.

The general pattern is that the worlds of policy, practice and research are widely separated. A crucial challenge for more successful innovation in education is to build bridges between many levels, factors and actors.

How may research help in addressing educational challenges? The kind of help usually varies over different types of research. Plomp (2013) distinguishes various questions,

aims and functions of research, such as: to describe, to compare, to evaluate, to explain, to predict, to design and to develop. One may also discern various primary orientations of research: theory, practice or policy. Much policy-oriented research on education occurs through surveys, monitoring and assessment, focusing on (descriptive) measures about actual practices and outcomes. However, the central orientation in this chapter will be on research that focuses on curriculum improvement and innovation.

The next section (building on: van den Akker, 2003; Thijs & van den Akker, 2009) focuses on summarizing a set of concepts and perspectives that help to increase the transparency and balance of curriculum analysis, development and discourse.

Curriculum, what's in a name?

When there is a myriad of definitions of a concept in the literature (as with curriculum), it is often difficult to keep a clear focus on its essence. In these cases it often helps to search for the etymological origin of the concept. The Latin word *'curriculum'* (related to the verb *currere* i.e. running) refers to a *'course'* or *'track'* to be followed. In the context of education, where learning is the central activity, the most obvious interpretation of the word curriculum is then to view it as a course, trajectory, or *'plan for learning'* (cf. Taba, 1962). This very short definition (reflected in related terms in many languages) limits itself to the core of all other definitions, permitting all sorts of elaborations for specific educational levels, contexts, and representations. Obviously, contextual specification is always needed in curriculum conversations to clarify the perspective.

Given this simple definition, a differentiation between various levels of the curriculum has proven to be very useful when talking about curricular activities (policy-making; design and development; evaluation and implementation). The following distinctions appear to be helpful:

- international/comparative (or *supra* level)
- system/society/nation/state (or *macro*) level (e.g. national syllabi or core objectives)
- school/institution (or *meso*) level (e.g. school-specific curriculum)
- classroom (or *micro*) level (e.g. textbooks, instructional materials)
- individual/personal (or *nano*) level.

The supra level usually refers to international debates or agreements on aims and quality of education, and is sometimes fuelled by the outcomes of internationally comparative studies. Curriculum development at the supra and macro levels is usually of a *'generic'* nature, while *'site-specific'* approaches are more applicable for the levels closer to school (meso) and classroom (micro) practice. Moreover, the process of curriculum development

can be seen as narrow (developing a specific curricular product) or broad (a long term, ongoing process of curriculum improvement, often including many related aspects of educational change, e.g. teacher education, school development, and examinations). In order to understand problems of curriculum decision-making and enactment, a broader description of curriculum development is often most appropriate: usually a long and cyclical process with many stakeholders and participants; in which motives and needs for changing the curriculum are formulated; ideas are specified in programmes and materials; and efforts are made to realise the intended changes in practice.

Moreover, curricula can be represented in various forms. Clarification of these forms is especially useful when trying to understand the problematic efforts to change the curriculum. A common broad distinction is between the three levels of the *'intended'*, *'implemented'*, and *'attained'* curriculum. A more refined typology (van den Akker, 2003) is outlined in box I.

Box 1: Typology of curriculum representations

INTENDED	Ideal	Vision (rationale or basic philosophy underlying a curriculum)
	Formal/Written	Intentions as specified in curriculum documents and/or materials
IMPLEMENTED	Perceived	Curriculum as interpreted by its users (especially teachers)
	Operational	Actual process of teaching and learning (also: curriculum-in-action)
ATTAINED	Experiential	Learning experiences as perceived by learners
	Learned	Resulting learning outcomes of learners

Traditionally, the intended domain refers predominantly to the influence of curriculum policy-makers and curriculum developers (in various roles), the implemented curriculum relates especially to the world of schools and teachers, and the attained curriculum has to do with the students.

Besides this differentiation in representations, curriculum problems can be approached from various analytical angles. For example, Goodlad (1994) distinguishes the following three different perspectives:

- *'substantive'*, focusing on the classical curriculum question about what knowledge is of most worth for inclusion in teaching and learning
- *'technical-professional'*, referring to how to address concrete tasks of curriculum development
- *'socio-political'*, referring to curriculum decision-making processes, where values and interests of many different individuals and agencies are at stake.

Some might argue that this list is too limited as it refers especially to curriculum issues for *'traditional'* planning for learning in schools, and does not include the more *'critical'* perspectives that are amply present in curriculum theory literature (e.g. Pinar, Reynolds, Slattery, & Taubman, 1995). However, from a primary interest in curriculum improvement, the three perspectives seem useful and appropriate.

The vulnerable curricular spider's web

One of the major challenges for curriculum improvement is creating balance and consistency between the various components of a curriculum (i.e. plan for learning). What are those components? The relatively simple curriculum definition by Walker (1990) includes three major planning elements: content, purpose and organisation of learning. However, curriculum design and implementation problems have taught us that it is wise to pay explicit attention to a more elaborated list of components. Elaborating on various typologies, we have come to adhere to a framework (see Box 2) of ten components that address ten specific questions about the planning of student learning.

Box 2: Curriculum components

Rationale or Vision	Why are they learning?
Aims & Objectves	Towards which goals are they learning?
Content	What are they learning?
Learning Activities	How are they learning?
Teacher Role	How is the teacher facilitating learning?
Materials & Resources	With what are they learning?
Grouping	With whom are they learning?
Location	Where are they learning?
Time	When are they learning?
Assessment	How to measure how far learning has progressed?

The *'rationale'* (referring to overall principles or central mission of the plan) serves as a major orientation point, and the nine other components are ideally linked to that rationale, and preferably also consistent with each other. For each of the components many sub-questions are possible, not only on substantive issues (see the next section), but, for example, also on organisational aspects as:

- Grouping:
  - How are students allocated to various learning trajectories?
  - Are students learning individually, in small groups, or whole-class?
- Location:
  - Are students learning in class, in the library, at home, or elsewhere?
  - What are the social/physical characteristics of the learning environment?
- Time:
  - How much time is available for various subject matter domains?
  - How much time can be spent on specific learning tasks?

Our preferential visualisation of the ten components is to arrange them as a spider's web (Figure 1), not only illustrating its many inter-connections, but also underlining its vulnerability. Thus, although the emphasis of curriculum design on specific components may vary over time, eventually some kind of alignment has to occur to maintain coherence. A striking example is the trend toward integration of ICT in the curriculum, with usually initial attention to changes in materials, resources and location. Many implementation studies have exemplified the need for a more comprehensive approach and systematic attention to the other components before one can expect robust changes. The spider's web also illustrates a familiar expression: every chain is as strong as its

weakest link. This seems another very appropriate metaphor for a curriculum, pointing to the complexity of efforts to improve the curriculum in a balanced, consistent and sustainable manner.

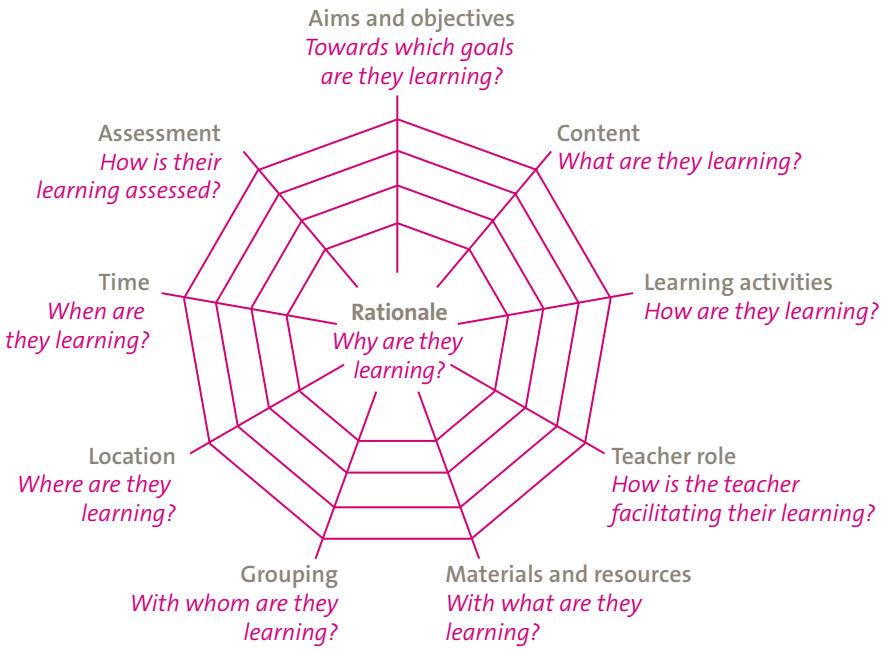


Figure 1: Curricular spider's web

Perspectives on substantive choices

A classic approach to the eternal curriculum question of what to include in the curriculum (or even more difficult as well as urgent: what to exclude from it?) is to search for a balance between three major sources or orientations for selection and priority setting:

- Knowledge: what is the academic and cultural heritage that seems essential for learning and future development?
- Society: which problems and issues seem relevant for inclusion from the perspective of societal trends and needs?
- Learner: which elements seem of vital importance for learning from the personal and educational needs and interests of the learners themselves?

Answers to these questions usually constitute the rationale of a curriculum. Inevitably, choices have to be made, involving compromises between the various orientations (and their respective proponents and pressure groups). Often, efforts fail to arrive at generally acceptable, clear and practical solutions. The result of adding up all kinds of wishes is that curricula tend to get overloaded and fragmented. Implementation of such incoherent curricula tends to lead eventually to student frustrations, failure, and dropout.

How to create a better curriculum balance? Easy answers are not available, but a few alternatives seem promising. First, in view of the multitude of (academic) knowledge claims, it sometimes helps to reduce the large number of separate subject domains to a more limited number of broader learning areas, combined with sharper priorities in aims for learning (focusing on basic concepts and skills).
Second, referring to the avalanche of societal claims, more interaction between learning inside and outside the school may reduce the burden. However, the most effective response is probably to be more selective in reacting to all sorts of societal problems. As Cuban (1992) phrased it clearly: "schools should not feel obliged to scratch the back of society every time society has an itch".
And third, about the learners' perspective: worldwide, many interesting efforts are going on to make learning more challenging and intrinsically motivating by moving from traditional, teacher- and textbook- dominated instruction towards more meaningful and activity-based learning approaches. Obviously, ICT creates new challenges, but also offers new opportunities for addressing the substantive dilemmas described.

Development strategies

To sketch curriculum development as a problematic domain is actually an understatement. From a socio-political stance, it often seems more appropriate to describe it as a war zone, full of conflicts and battlefields between stakeholders with different values and interests. Problems manifest themselves in the (sometimes spectacular and persistent) gaps between the intended curriculum (as expressed in policy rhetoric), the implemented curriculum (real life in school and classroom practices), and the attained curriculum (as manifested in learner experiences and outcomes). A typical consequence of these tensions is that various frustrated groups of participants blame each other for the failure of reform or improvement activities. Although such blaming games often seem rather unproductive, there are some serious critical remarks to be made on many curriculum development approaches worldwide. First of all, many curriculum reform efforts can be characterized by overly big innovation ambitions (especially of politicians) within unrealistically short timelines and with very limited investment in people, especially teachers. Second, often there is a lack of coherence between the intended

curriculum changes with other system components (especially teacher education and assessment/examination programs). And last, but not least, timely and authentic involvement of all relevant stakeholders is often neglected.

From a strategic point of view, the literature has offered us many (technical-professional) models and strategies for curriculum development. Three prominent approaches are Tyler's (1949) *'rational-linear'* approach, Walker's (1990) *'deliberative'* approach, and Eisner's (1979) *'artistic'* approach. As it does not fit within the purpose of this essay to explain these models in particular, the reader is referred to educative texts of Marsh and Willis (2003) or to the overview of approaches in Thijs and van den Akker (2009).
Obviously, the context and nature of the curriculum development task at hand will determine to a large extent what kind of strategy is indicated. It is noteworthy that we are beginning to see more *'blended'* approaches that integrate various trends and characteristics of recent design and development approaches in the field of education (for an overview and a series of examples: see van den Akker, Branch, Gustafson, Nieveen, & Plomp, 1999, or van den Akker & Kuiper, 2008). Some key characteristics of these are:

- **Pragmatism**: Recognition that there is not a single perspective, overarching rationale or higher authority that can resolve all dilemmas for the curriculum choices to be made. The practical context and its users are in the forefront of curriculum design and enactment.
- **Prototyping**: Evolutionary prototyping of curricular products and their subsequent representations in practice is viewed as more productive than quasi-rational and linear development approaches. Gradual, iterative approximation of curricular dreams into realities may prevent paralysis and frustrations. Formative evaluation of tentative, subsequent curriculum versions is essential to such curriculum improvement approaches.
- **Communication**: A communicative-relational style is desirable in order to arrive at the inevitable compromises between stakeholders with various roles and interests and to create external consistency between all parties involved.
- **Professional development**: In order to improve chances on successful implementation, there is a trend towards more integration of curriculum change and professional learning and development of all individuals and organisations involved.

Curricular development research (CDR) is an approach that incorporates some of these characteristics, and it becomes even more promising by adding the element of knowledge growth to it. CDR can strengthen the knowledge base in the form of design principles that offer heuristic advice to curriculum development teams, if (more than in common development practices) deliberate attention is paid to theoretical embedding of design issues and empirical evidence is offered about the practicality and effectiveness of the curricular interventions in real user settings.

However, there are several persistent dilemmas for curriculum development that cannot easily be resolved, let alone through generic strategies. For example: how to combine aspirations for large-scale curriculum change and system accountability with the need for local variations and ownership? The tension between these conflicting wishes can be somewhat reduced when one avoids the all too common *'one size fits all'* approach. More adaptive and flexible strategies will avoid detailed elaboration and over-specification of central curriculum frameworks. Instead, they offer substantial options and flexibility to schools, teachers, and learners. Although struggles about priorities in aims and content will remain inevitable, the principle of *'less is more'* should be pursued. However, what is incorporated in a core curriculum should be clearly reflected in examination and assessment approaches.

The *'enactment'* perspective (teachers and learners together create their own curriculum realities) is increasingly replacing the *'fidelity'* perspective on implementation (teachers faithfully follow curricular prescriptions from external sources). That trend puts even more emphasis on teachers as key people in curriculum change. Both individual as well as team learning is essential (Fullan, 2007). Teachers need to get out of their customary isolation. Collaborative design and piloting of curricular alternatives can be very productive (e.g. Handelzalts, 2009), especially when experiences are exchanged and reflected upon in a structured curriculum discourse. Interaction with external facilitators can contribute to careful explorations of the *'zone of proximal development'* of teachers and their schools. Cross-fertilization between curriculum, teacher, and school development is a *conditio sine qua non* for effective and sustainable curriculum improvement. The increasingly popular mission statements of schools to become attractive and inspiring environments for students and teachers can only be realised when such integrated scenarios are practised.

The potential of curricular development research

Various motives for initiating and conducting educational design research can be mentioned. A basic motive stems from the experience that many research approaches (e.g. experiments, surveys, correlational analyses), with their focus on descriptive knowledge, hardly provide prescriptions with useful solutions for a variety of design and development problems in education. Probably the greatest challenge for professional designers is how to cope with the manifold uncertainties in their complex tasks in very dynamic contexts. If they do seek support from research to reduce those uncertainties, several frustrations often arise: answers are too narrow to be meaningful, too superficial to be instrumental, too artificial to be relevant, and, on top of that, they usually come too late to be of any use. Curriculum designers do appreciate more adequate information to create a solid ground for their choices and more timely feedback to improve their products. Moreover, the professional

community of developers as a whole would be helped by a growing body of knowledge of theoretically underpinned and empirically tested design principles and methods.
Another reason for CDR stems from the highly ambitious and complex nature of many curriculum reform policies in education worldwide. These reform endeavors are sometimes related to specific ideologies, usually affect many system components, are often multi-layered, including both large-scale policies and small-scale realization, and are very comprehensive in terms of factors included and people involved. These radical *'revolutions'*, if promising at all, cannot be realized on the drawing table.
The scope of diverse needs is often very wide, the problems to be addressed are usually ill-specified, the effectiveness of proposed interventions is mostly unknown beforehand, and the eventual success is highly dependent on implementation processes in a broad variety of contexts. Therefore, such curriculum reform efforts would profit from more evolutionary (interactive, cyclical, spiral) approaches, with integrated research activities to feed the process (both forward and backward). Such an approach would provide more opportunities for *'successive approximation'* to the ideals and for more strategic learning in general. In conclusion: CDR seems a wise and productive approach for curriculum development, offering a sort of middle ground between overly ideological or overly technocratic perspectives.

Features of curricular development research

Curricular development research is often initiated for complex, innovative interventions for which only very few validated principles are available to structure and support the design and development activities. Since in these situations the image and impact of the curricular intervention to be developed is often still unclear, the research focuses on realizing limited, but promising, examples of these interventions. The aim is not to elaborate and implement complete interventions, but to come to (successive) prototypes that increasingly meet the innovative aspirations and requirements. The process is often iterative, cyclic or spiral: analysis, design, evaluation and revision activities are iterated until a satisfying balance between ideals and realization has been achieved.

To what extent do these CDR activities differ from what is typical for design and development approaches in professional practices? What are the implications of the accountability of researchers to the *scientific forum*? At the risk of exaggerating the differences, let us outline some of them, based on what is known about routinized, standard patterns in curriculum development practices. Of course, a lot of activities are more or less common for both approaches, so the focus will be on those additional elements that are more prominent in design research than in common design and development practices.

(1) Preliminary investigation
A more intensive and systematic preliminary investigation of curriculum tasks, problems, and context is made, including searching for more accurate and explicit connections of that analysis with state-of-the-art knowledge from literature. Some typical activities include: literature review; consultation of experts; analysis of available promising curriculum examples for related purposes; case studies of current practices to specify and better understand needs and problems in intended user contexts.

(2) Theoretical embedding
More systematic efforts are made to apply state-of-the-art knowledge in articulating the theoretical rationale for curriculum design choices. Moreover, explicit feedback to assertions in the design rationale about essential characteristics of the intervention (substantive design principles) is made after empirical testing of its quality. This theoretical articulation can increase the *'transparency'* and *'plausibility'* of the rationale. Because of their specific focus, these theoretical notions are usually referred to as *'mini'* or *'local'* theories, although sometimes connections can also be made to *'middle-range'* theories with a somewhat broader scope.

(3) Empirical testing
Clear empirical evidence is delivered about the practicality and effectiveness of the curriculum for the intended target group in real user settings. In view of the wide variation of possible interventions and contexts, a broad range of (direct/indirect; intermediate/ultimate) indicators for *'success'* should be considered.

(4) Documentation, analysis and reflection on process and outcomes
Much attention is paid to systematic documentation, analysis and reflection on the entire design, development, evaluation and implementation process and on its outcomes in order to contribute to the expansion and specification of the methodology of curriculum design and development.

Typical questions for CDR are:
- What are essential features of successful curricular interventions?
- How do those interventions operate in real life?
- How can they be designed and implemented?

In view of its aim, CDR is never a *'quick fix'* operation, but it usually takes a long trajectory, where the research is intertwined with continuous development activities - from problem formulation up to and including implementation. More than most other research approaches, CDR aims at making both practical and scientific contributions. In the search for innovative *'solutions'* for curriculum problems, interaction with

practitioners (in various professional roles: teachers, policy-makers, developers, and the like) is essential. The ultimate aim is not to test whether theory, when applied to practice, is a good predictor of events. The inter-relation between theory and practice is more complex and dynamic: is it possible to create a practical and effective curriculum for an existing problem or intended change in the real world? The innovative challenge is usually quite substantial, otherwise the research would not be initiated at all. Interaction with practitioners is needed to gradually clarify both the problem at stake and the characteristics of its potential solution. An iterative process of *'successive approximation'* or *'evolutionary prototyping'* of the *'ideal'* intervention is desirable.

Direct application of theory is not sufficient to solve those complicated problems. One might state that a more *'constructivist'* development approach is preferable: researchers and practitioners cooperatively construct and test workable interventions and articulate principles that underpin the effects of those interventions. Another reason for cooperation is that without the involvement of practitioners it is impossible to gain clear insight in potential curriculum implementation problems and to generate measures to reduce those problems. New interventions, however imaginative their design, require continuous anticipation at implementation issues. Not only for *'social'* reasons (to build commitment and ownership of users) but also for *'technical'* benefits: to improve their fitness for survival in real life contexts. Therefore, rigorous testing of practicality is a *conditio sine qua non* in D&DR.

Emphasis on formative evaluation

As has become clear in the previous sections, formative evaluation holds a prominent place in curricular design research. The main reason for this central role is that formative evaluation provides the information that feeds the optimization of the intervention and the cyclical learning process of curriculum developers during the subsequent loops of a design and development trajectory. It is most useful when fully integrated in a cycle of analysis, design, evaluation, revision, *et cetera*, and when contributing to improvement of the curriculum.

Thus, the basic contribution of formative evaluation is quality improvement of the curriculum under development. *'Quality'*, however, is an abstract concept that requires specification. During development processes, the emphasis in criteria for quality usually shifts from relevance, to consistency, to practicality, to effectiveness. *'Relevance'* refers to the extent that the intended curriculum is perceived to be a relevant improvement to practice, as seen from the varied perspectives of policy makers, practitioners and researchers. *'Consistency'* refers to the extent that the design of the curriculum is based

on state-of-the-art knowledge and that the various components of the intervention are consistently linked to each other (cf. the curricular spider's web). *'Practicality'* refers to the extent that users (and other experts) consider the intervention as clear, usable and cost-effective in 'normal' conditions. *'Effectiveness'* refers to the extent that the experiences and outcomes with the intervention are congruent with the intended aims. Also, issues of scalability and sustainability may be included in a broad interpretation of effectiveness.

The methods and techniques for evaluation will usually be attuned to that shift in criteria. For example, adequate evaluation of consistency can start with comments of critical friends on initial drafts and then move over to more systematic expert appraisal. Practicality is often tested via micro-evaluations and try-outs in real classroom practices. Evaluation of effectiveness usually requires (more large-scale) field tests. In later stages of formative evaluation, methods of data collection will usually be less intensive, but with an increasing number of respondents (e.g. achievement test for many students at the end compared to in-depth interview with a few experts in the beginning). See Nieveen and Folmer (2013) for more elaborate explanations and suggestions for these shifts in formative evaluation.

Formative evaluation within CDR should not only concentrate on locating shortcomings of the intervention in its current (draft) version, but especially generate suggestions on how to improve these weak points. Richness of information, notably salience and meaningfulness of suggestions on how to make an intervention stronger, is therefore more productive than standardization of methods to collect and analyze data. Also, efficiency of procedures is crucial. The lower the costs in time and energy for data collection, processing, analysis and communication will be, the bigger the chances of actual use and impact on the development process. For example, samples of respondents and situations for data collection will usually be relatively small and purposive compared to sampling procedures for other research purposes. The added value of getting *'productive'* information from more sources tends to decrease, because the opportunities for *'rich'* data collection methods (such as interviews and observations) are limited with big numbers. To avoid an overdose of uncertainty in data interpretation, often triangulation (of methods, instruments, sources, and sites) is applied. These arguments especially hold true for early stages of formative evaluation, when the intervention is still poorly crystallizSed.

Generalization of curricular design research findings

The practically most relevant outcome of CDR is its contribution towards optimization of the curricular product and its actual use, leading to better instructional processes and learning results. However, a major contribution to knowledge to be gained from CDR is in the form of (both substantive and methodological) *'design principles'* to support developers in their task. These principles may be captured in (a growing set of) heuristic statements of a format such as:

- *If you want to design intervention X [for purpose/function Y in context Z]*
- *then you are best advised to give that intervention the characteristics C1, C2,..., Cm* [substantive emphasis]
- *and to do that via procedures P1, P2, ..., Pn* [methodological emphasis]
- *because of theoretical arguments T1, T2, ..., Tp*
- *and empirical arguments E1, E2, ..., Eq*

Thus, the design principles include not only statements about the (substantive) what and (methodological) how of the intended interventions, but also offer theoretical explanations and empirical underpinning to justify these knowledge claims. Obviously these heuristic principles cannot guarantee success, but they are intended to select and apply the most appropriate (substantive and methodological) knowledge for specific design and development tasks.
It is not uncommon in CDR that such knowledge, especially the substantive knowledge about essential curriculum characteristics, can partly be extracted from a resulting prototype itself. That is one of the reasons that make it so profitable to search for and carefully analyze already available curricula to generate ideas for new design tasks. However, the value of that knowledge will strongly increase when justified by theoretical arguments, well-articulated in providing directions, and convincingly backed-up with empirical evidence about the impact of these principles. Moreover, these heuristic principles will be additionally powerful if they have been validated in successful design of more interventions in more contexts. Chances for such knowledge growth will increase when CDR is not conducted through isolated studies, but conducted within the framework of research programs, because projects can then build upon one another (*'standing on shoulders'*). Such a strategy also increases the chances for a sharper insight into the essence of successful interventions versus variations in other features.

Since data collection in CDR, especially during formative curriculum evaluation, is often limited to small (and purposive) samples, efforts to generalize findings cannot be based on statistical techniques, focusing on generalizations from sample to population. Instead one has to invest in *'analytical'* forms of generalization (cf. Yin, 2003): readers/users need to

be supported to make their own attempts to explore the potential transfer of the research findings to theoretical propositions in relation to their own context.
Reports on CDR can facilitate the task of *'analogy reasoning'* by a clear theoretical articulation of the design principles applied, and by reflection on the results afterwards. Moreover, it is helpful to offer a careful description of both the evaluation procedures as well as the implementation context. Especially since a detailed description of the process-in-context may increase the *'ecological'* validity of the findings, so that others can estimate in what respects and to what extent transfer from the reported situation to their own is possible. Another option that may stimulate exploration of possibilities for (virtual) generalization and transfer to various settings is to organize interactive (network) meetings with experts and practitioners from related contexts to discuss the plausibility of the research findings and recommendations for related tasks and contexts.
In view of the (rapidly growing, but still relatively) modest familiarity of CDR to the wider audience, it is wise to invest a lot in the *'transparency'*, *'plausibility'*, *'trustworthiness'* and *'reconstructiveness'* of its arguments and findings. Also at stake is the *'credibility'* (expertise in depth and breadth; track record) of the research team and its partners.

Last but not least, CDR may offer drafts of various relevant curriculum versions (with proven consistency and practicality) that can be compared in more quantitative, large-scale, (quasi-) experimental studies. Obviously, such more summative evaluations are better done by other, more independent researchers who were not previously involved in the design stage.

References

Cuban, L. (1992). Curriculum stability and change. In P. Jackson (Ed.), *Handbook of research on curriculum* (pp. 216-247). New York: Macmillan.

Eisner, E. (1979). *The educational imagination: On the design and evaluation of educational programs*. New York: Macmillan.

Fullan, M. (2007). *The new meaning of educational change*. Fourth Edition. New York: Teachers College Press.

Goodlad, J. (1994). Curriculum as a field of study. In T. Husén, & T. Postlethwaite (Eds.), *The international encyclopedia of education* (pp. 1262-1276). Oxford: Pergamon Press.

Hargreaves, A., & Fink, D. (2006). *Sustainable leadership*. San Francisco, CA: Jossey-Bass.

Leyendecker, R. (2008). *Curriculum reform in Sub-saharan Africa: Opportunities and obstacles*. Doctoral dissertation. Enschede: University of Twente.

Marsh, C., & Willis, P. (2003). *Curriculum: Alternative approaches, ongoing issues* (third edition*)*. Upper Saddle River, NJ: Merrill/PrenticeHall.

Nieveen, N., & Folmer, E. (2013). Formative evaluation in educational design research. In T. Plomp, & N. Nieveen (Eds.), *Educational design research – part A: An introduction* (pp. 152-169). Enschede, the Netherlands: SLO.

Pinar, W., Reynolds, W., Slattery, P., & Taubman, P. (1995). *Understanding curriculum. An introduction to the study of historical and contemporary curriculum discourses*. New York: Peter Lang.

Plomp, T. (2013). Educational design research: An introduction. In T. Plomp, & N. Nieveen (Eds*.), Educational design research – part A: An introduction* (pp. 10-51). Enschede, the Netherlands: SLO.

Taba, H. (1962). *Curriculum development: Theory and practice*. New York: Harcourt, Brace & World.

Thijs, A., & van den Akker, J. (Eds.). (2009). *Curriculum in development*. Enschede, the Netherlands: SLO.

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

Van den Akker, J. (1999). Principles and methods of development research. In J. van den Akker R.M. Branch, K. Gustafson, N. Nieveen, & T. Plomp (Eds.), *Design approaches and tools in education and training* (pp. 1-14). Dordrecht: Kluwer Academic Publishers.

Van den Akker, J. (2003). Curriculum perspectives: An introduction. In J. van den Akker, W. Kuiper, & U. Hameyer (Eds.), *Curriculum landscapes and trends* (pp. 1-10). Dordrecht: Kluwer Academic Publishers.

Van den Akker, J., Branch, R., Gustafson, K., Nieveen, N., & Plomp, T. (Eds.). (1999). *Design approaches and tools in education and training*. Dordrecht: Kluwer Academic Publishers.

Van den Akker, J., Gravemeijer, K., McKenney, S., & Nieveen, N. (Eds.). (2006). *Educational design research*. London: Routledge.

Van den Akker, J., & Kuiper, W. (2008). Research on models for instructional design: Towards more productive approaches. In J. Spector, D. Merrill, J. van Merrienboer, & M. Driscoll (Eds.), *Handbook of research on educational technology and communications* (pp. 739-748). Third edition. Mahwah, NJ: Lawrence Erlbaum Associates.

Van den Akker, J., Kuiper, W., & Nieveen, N. (2012). Bruggen slaan tussen beleid, praktijk en wetenschap in curriculumontwikkeling en –onderzoek [Building bridges between policy, practice and science in curriculum development and research]. *Pedagogische Studiën, 89*, 399-410.

Walker, D. (1990). *Fundamentals of curriculum*. Fort Worth, TX: Harcourt, Brace, College.

Yin, R. K. (2003). *Case study research: Design and methods*. London: Sage.

3. Design Research from the Learning Design Perspective[1]

Koeno Gravemeijer and Paul Cobb

In this contribution, we want to elaborate an approach to design research that has been used and refined in a series of design research projects in which the two authors collaborated over a ten-year period. To locate our contribution in this book, we may categorize our approach as falling in the broader category of what Plomp (2013) calls 'validation studies'. We want to caution, however, that connotations that the word 'validation' may call up, such as 'checking' and 'confirming', do not fit the exploratory character of our approach that aims at creating innovative learning ecologies in order to develop local instruction theories on the one hand, and to study the forms of learning that those learning ecologies are intended to support on the other hand. The research projects on which we will focus involve a research team taking responsibility for a group of students' learning for a period of time. And all concern the domain of mathematics education (including statistics education). We use the metaphor of learning ecologies to emphasize that learning environments are conceptualized as interacting systems rather than as either a collection of activities or a list of separate factors that influence learning.

> "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, Confrey, diSessa, Lehrer, & Schauble, 2003, pp. 9–13).

The approach to design research, which we developed over the years, has its roots in the history of the two authors. One is that of socio-constructivist analysis of instruction. The other is that of the work on realistic mathematics education (RME) that is carried out in the Netherlands. The underlying philosophy of design research is that you have to understand the innovative forms of education that you might want to bring about in order to be able to produce them. This fits with the adagio, that 'if you want to change something, you have to understand it, and if you want to understand something, you have to change it'. The two sides of this adagio mirror the authors' two histories. The RME approach was inspired by a need for educational change, the socio-constructivist approach by a desire for understanding.

1) An earlier version of this chapter has been published as: Gravemeijer, K. & Cobb, P. (2006). Design research from the learning design perpective. In J. van den Akker, K. Gravemeijer, S. McKenney, & N. Nieveen (Eds.) *Educational Design research* (pp. 45-85). London: Routledge.

If we take the first of these perspectives, we may observe that the notion of design research has been around for a long time. Various forms of professional instructional design may be perceived as informal predecessors of design research. The recognition that instructional design often had an innovative character, while the available scientific knowledge base was far too limited to ground the design work sparked the idea for a type of instructional design that integrated design and research. This idea was strengthened by the experience that conscious and thorough instructional design work brought a learning process in which the designers developed valuable and well-grounded knowledge in what retrospectively might be called design experiments.

Over time a number of proposals have been made to define design research in mathematics education, of which Brown's (1992) article on design experiments is one of the most notable. In the Netherlands, Freudenthal (Freudenthal, Janssen, & Sweers, 1976) was perhaps the first to propose an approach of this type with his concept of 'developmental research', an idea that was further elaborated by Streefland (1990) and Gravemeijer (1994, 1998).

Freudenthal's ideas were put to practice in the Dutch Institute for the Development of Mathematics Education, IOWO (the later OW&OC, now called Freudenthal Institute), which developed numerous local instruction theories. Each local instruction theory (Gravemeijer, 1994) consists of a theory about a possible learning process, together with theories about possible means of supporting that learning process. By generalizing over those local instruction theories, Treffers (1987) developed the so-called domain specific instruction theory of realistic mathematics education (RME).

The second part of the adagio, 'if you want to understand something, you have to change it', points to the other predecessor of our collaborative work on design research, the constructivist 'teaching experiment methodology' (Cobb & Steffe, 1983; Steffe, 1983). Within this methodology, one-on-one teaching experiments aimed primarily at understanding how students learn rather than at educational change. These one-on-one teaching experiments were later expanded into classroom teaching experiments. The need for classroom teaching experiments arose when analysis of traditional instruction within the same (socio-constructivist) research program, produced only negative advice for the teachers; advice of the type: 'Don't do this, don't do that'. To create more productive classroom environments, the researchers had to take the responsibility for the design of the instruction of a classroom for an extended period of time. In doing so, the one-on-one teaching experiment methodology was expanded to classroom teaching experiments (see Yackel, Gravemeijer, & Sfard, 2011, for an extensive reconstruction of this process).

The focus on understanding is a salient characteristic of design research. In this respect, the distinction Bruner (1994) makes between research that aims at (statistical) explanation, and research that aims at understanding comes to mind. We may use this

distinction to emphasize that the goal of design research is very different from research along the lines of an experimental or quasi-experimental research design. And different goals implies different methods and different forms of justification. In relation to this we may quote the NCTM Research Advisory Committee (1996) that observes 'a shift in norms of justification' in mathematics education research. This is a shift, the Committee argues, from research that proves that treatment A works better than treatment B, towards research that has as its goal to provide *an empirically grounded theory on how the intervention works*. This latter goal also differentiates design research from action research. The intent of action research is typically limited to effecting change in the local setting where the research is carried out, whereas design research aims to learn about what the innovation may entail and how it could be achieved. The goal of design experiments is to develop theories about both the process of learning and the means that are designed to support that learning. One may work towards this goal in two ways, either by developing local instruction theories, or by developing theoretical frameworks that addresses more encompassing issues—such as the classroom culture, or the role of symbols. In our approach to design research we try to combine the two.

In the following, we make the issue of what design research is for us concrete by discussing the three phases of conducting a design experiment , which are 1) preparing for the experiment, 2) experimenting in the classroom, and 3) conducting retrospective analyses. In doing so, we will address a range of methodological considerations. To ground the discussion in a concrete design experiment, we will use an experiment on statistics to illustrate the various phases. Although some may not consider statistics a part of mathematics, we contend this illustrative case of statistics education is compatible with the kind of mathematics education we seek to bring about.

As a point of clarification, we may add that because the goals of design research differ from those of (quasi-)experimental research, the way the research is framed also differs. Typically, the goal of design research is not translated into a single research question. One might of course formulate a research question of the type, 'How can we teach such and such a topic effectively?' But such a one-sentence question would have to be complemented with a series of assumptions about what requirements the answer should meet, and would also have to be embedded in an exposition of what is aimed for with the innovation. Moreover, new questions and new conjectures may arise during the execution of the research project. Goals of design research are therefore typically cast in terms of the innovative learning ecologies and the kind of theories one is aiming for.

Phase one, the preparing for the experiment

From a design perspective, the goal of the preliminary phase of a design research experiment is to formulate a local instruction theory that can be elaborated and refined while conducting the intended design experiment. From a research perspective, a crucial issue is that of clarifying its theoretical intent. In elaborating these points, we will start by clarifying how one goes about establishing the learning goals, or instructional end points to which one is aiming, and the instructional starting points. Next we will discuss the conjectured local instruction theory that the research team has to develop. This local instruction theory encompasses both provisional instructional activities, and a conjectured learning process that anticipates how students' thinking and understanding might evolve when the instructional activities are employed in the classroom. We will close this section by elaborating on the theoretical intent of an experiment.

End points

The preparation for a classroom design experiment typically begins with the clarification of the mathematical learning goals. Such a clarification is needed, as one cannot adopt the educational goals that are current in some domain. These goals will in practice largely be determined by history, tradition, and assessment practices. Design researchers therefore cannot just take these goals as a given when starting a design experiment. Instead, they will have to problematize the topic under consideration from a disciplinary perspective, and ask themselves: What are the core ideas in this domain?
We may illustrate this activity of problematizing with our work in the domain of early statistics.

The conventional content of statistics at the US Middle School level (12-14 year old students) is rather meager. It is basically a collection of separate topics - such as mean, median, and mode - and standardized graphical representations. Reviewing the literature that was available at that time, however, did not offer much help, there appeared to be no consensus on what the central ideas should be. By analyzing what doing statistics entails, we came to the conclusion that the notion of distribution plays a central role. We concluded that distribution could function as an overarching idea that could go through elementary school, middle school, and up to high school and college. From this perspective, notions like 'center', 'skewness', 'spread', and 'relative frequency' are ways of characterizing how the data are distributed, rather than separate topics or concepts on themselves. In addition, different types of statistical representations come to the fore as different ways of structuring and organizing data sets in order to detect relevant patterns and trends.

This elaboration serves to emphasize that the goal of design research is not to take the currently instituted or institutionalized school curriculum as a given, and to try to find

better ways to achieve the given goals. Instead, the research team has to scrutinize those goals from a disciplinary point of view, in order to establish what the most relevant or useful goals are. Consequently, the design research we describe here is interventionist in character. In our example, part of our agenda was to attempt to influence what statistics should be in school, at least at a middle school level in the US.

Starting points

In order to be able to develop a conjectured local instruction theory, one also has to consider the instructional starting points. Mark that the focus in doing so is to understand the consequences of earlier instruction, not merely to document the typical level of reasoning of 12 or 14 year old students in a given domain. Here, the existing research literature can be useful. Psychological studies on students' levels of reasoning on a given topic can usually be interpreted as documenting the effects of prior instructional history. To complement such a literature study, the researchers will also have to carry out their own assessments, before starting a design experiment. In some cases, they may be able to use available items and instruments. In addition to written tests, there will also be a need for other forms of assessment, such as interviews, or whole class performance assessments to assess the actual levels of understanding of the students in the experimental classroom, and to identify potential starting points. We have found performance assessments to be particularly useful in documenting instructional starting points. We may illustrate this with the example of the statistics design experiment.

In preparation for the design experiment in data analysis, we gave a number of tasks to two classes. Then, rather than attempting to support the students' learning in the whole class discussion, the role of the teacher was to probe the students' understanding and reasoning, and to find out why they used particular approaches. These performance assessments clearly revealed the consequences of the students' prior instruction. For the students, data analysis was trying to remember what you're supposed to do with numbers. Data were not numbers plus context for them, to use a phrase from David Moore (1997). In his view, statisticians are always dealing with data plus context. In other words, data for these students were not measures of an attribute of a situation that was relevant with regard to the problem or issue under investigation. So, our initial challenge in the design experiment was to support a change in what statistics was about for these students, so that they were actually analyzing data.

Local instruction theory

Given the potential end points on the one hand, and the instructional starting points on the other hand, the research team has to formulate a local instruction theory. Such a conjectured local instruction theory consists of conjectures about a possible leaning

process, together with conjectures about possible means of supporting that learning process. The means of support encompass potentially productive instructional activities and (computer) tools as well as an envisioned classroom culture and the proactive role of the teacher. The research team tries to anticipate how students' thinking and understanding might evolve when the planned but revisable instructional activities are used in the classroom. In this manner, the research team tries to reconcile the need to plan in advance, and the need to be flexible when building on the students' current understandings when the design experiment is underway.

In many domains, the available research literature provides only limited guidance. In the case of statistics we had to work hard to find five relevant articles. The sort of articles that are relevant for construing local instruction theories are reports of the process of students' learning in a particular domain together with descriptions of the instructional settings, the tasks, and the tools that enabled or supported that learning.

To compensate for the lack of guidance that the literature offers, design researchers have to turn to other resources, such as curricula, texts on mathematics education, and the like. Actually, the design researcher may take ideas from whatever sources to construe an instructional sequence. Mark, however, that adopting often means adapting. In this respect, the way of working of a design researcher resembles the manner of working of what the French call a 'bricoleur.' A bricoleur is an experienced tinker/handy person, who uses as much as possible those materials that happen to be available. To do so, many materials will have to be adapted; the bricoleur may even have to invent new applications, which differ from what the materials were designed for. The design researcher follows a similar approach, labeled 'theory-guided bricolage' (Gravemeijer, 1994), to indicate that the way in which selections and adaptations are made will be guided by a (possibly still emergent) domain specific instruction theory.

The classroom culture and the proactive role of the teacher

Instructional designers typically focus on instructional tasks and tools as potential means of support. We would argue, however, that one also has to consider the characteristics of the envisioned classroom culture and proactive role of the teacher. One cannot plan instructional activities without considering how these activities are going to be enacted in the classroom. Design researchers therefore also have to consider the nature of classroom norms and the nature of classroom discourse. We know from experience that the norms of argumentation can differ radically from one classroom to another, and that they can make a profound difference in the nature and the quality of the students' mathematical learning (Cobb, Yackel, & Wood, 1989). Considerations on classroom norms and classroom discourse should therefore, be included in the design.

Thus one of the tasks of the teacher will be to establish the desired classroom culture. Further, the proactive role of the teacher will include introducing of the instructional

activities, or more specifically in the case of statistics, guiding the process of talking though the process of data creation. Further, the teacher will have to select possible topics for discussion, and orchestrate whole-class discussions on these topics. In most cases, this will require significant learning on the part of the teacher. In this regard, Van Eerde coined the term 'dual design research' to describe a methodological approach in which the goal of supporting the teacher's learning is addressed in a parallel design research project (Gravemeijer & van Eerde, 2009).

Theoretical intent

In addition to elaborating a preliminary instructional design, the research group also has to formulate the theoretical intent of the design experiment. For the goal of a design experiment is not just to describe what happened in a particular classroom. Analyses will have to be cases of a more general phenomenon that can inform design or teaching in other situations. One of the primary aims of a design experiment is to support the constitution of an empirically grounded local instruction theory.

Another aim of a design experiment might be to place classroom events in a broader context by framing them as instances of more encompassing issues. For example, analyses might be conducted that focus on the proactive role of the teacher, teacher's and students' negotiation of general classroom norms, or the teacher's learning. Also the role of symbolizing and modeling, or more generally of semiotic processes, in supporting students' learning can become an explicit focus of investigation. As a final example, we may mention that the statistics design experiment became a case of cultivating students' mathematical interests in that in the course of these experiments students became very interested in conducting data analysis to investigate issues. They came to view this as an activity worthy of their engagement. This relates to issues such as motivation and persistence. Ultimately, this might influence their decision whether to continue to study mathematics or not. For us, the cultivation of students' domain specific interests is an important aspect of mathematical literacy in its own right. As this latter example illustrates, that aims of a design experiment may be elaborated during the teaching experiment, or even afterwards.

In addition to these more encompassing issues, we may point to a third type of theory that may emerge during a series of design experiments; that of an ontological innovation. A series of design experiments can serve as the context for the development of theories or theoretical frameworks that entail new scientific categories that can do useful work in generating, selecting, and assessing design alternatives. The development of a conceptual framework to describe the phenomena under study is an essential part of a scientific endeavor. New categories, however, do not come readymade, and cannot simply be

captured by writing down a definition. New categories have to be invented and embedded in a supporting theoretical framework. Defining scientific terms is more like finding and validating a new category of existence in the world, for which we may use the term 'ontological innovation' (diSessa & Cobb, 2004).
Examples of such ontological innovations include the interpretative framework for interpreting classroom discourse and communication, which we will discuss later (Cobb & Yackel, 1996), the 'discovery' of meta-representational competence (diSessa, 1992, 2002), the theory of quantitative reasoning (Thompson, 1994, 1996), the design heuristic of emergent modeling (Gravemeijer, 1999), and RME theory in general (Treffers, 1987; Gravemeijer, 1994). The new frameworks and categories may be sought for, but often they emerge from design experiments in answer to the need to get a handle on surprising observations. The initial conceptualization, however, will typically be crude and in need of further elaboration and improvement. Ontological innovations therefore become a topic of a research program that spans a series of design experiments, within which the theoretical frameworks will be revised and refined to adjust them to a range of design contexts.

Mark 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. The social norms and the socio-mathematical norms that we will discuss in more detail later, 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. On the other hand, the same framework reveals what norms to aim for to make the design experiment successful. RME theory may play a similar dual role; the theory not only guides the design, but also offers a framework for interpreting the learning process of the students. One point of attention, for instance, will be the variety of solution procedures that the students produce. This can be seen as an indication of the extent in which these solution procedures are student inventions rather than unreflected copies of examples given by the teacher or other students. Moreover, according to the reinvention principle, one expects the variation in solution procedures to correspond with the conjectured reinvention route.

Phase two, the design experiment

The second phase consists of actually conducting the design experiment. When all the preparation work has been done, the overall endpoints are specified, the starting points are defined, and a conjectured local instruction theory is formulated, the design experiment can start. The research group—which may consist of one or more researchers

and preferably also includes the teacher—will take the responsibility for the learning process of a group of students, whether for five weeks, for three months, or even for a whole school year. However, before describing this second phase, it is important to clarify the intent or purpose for actually experimenting in the classroom.

Although, for some, the term 'experiment' may evoke associations with experimental, or quasi-experimental, research, the objective of the design experiment is not to try and demonstrate that the initial design or the initial local instruction theory works. The overall goal is not even to assess whether it works, although of course the researchers will necessarily do so. Instead the purpose of the design experiment is both to test and improve the conjectured local instruction theory that was developed in the preliminary phase, and to develop an understanding of how it works.

We will start our discussion of the design experiment with the iterative sequence of tightly integrated cycles of design and analysis, which is key to the process of testing, improving, and understanding. Next we will briefly touch upon the kind of data that are generated. Then we address the need for explicating the interpretative framework(s) one uses, on the one hand for interpreting classroom discourse and communication, and on the other hand for interpreting students' mathematical reasoning and learning.

Micro cycles of design and analysis

At the heart of the design experiment lays a cyclic process of (re)designing, and testing instructional activities and other aspects of the design. In each lesson cycle, the research team conducts an anticipatory thought experiment by envisioning how the proposed instructional activities might be realized interactively by the teacher and the students, and what students might learn as they participate in them. During the enactment of the instructional activities in the classroom, and in retrospect, the research team tries to analyze the actual process of the students' participation and learning. And, on basis of this analysis, the research team makes decisions about the validity of the conjectures that are embodied in the instructional activity, the establishment of particular norms and so forth, and about the revision of those specific aspects of the design. The design experiment therefore consists of a cyclic processes of thought experiments and instruction experiments (Freudenthal, 1991) (see Figure 1).

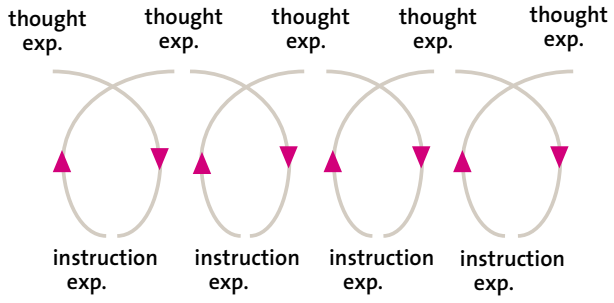


Figure 1: Developmental research, a cumulative cyclic process

We may associate these micro cycles of design and analysis with Simon's (1995) 'mathematical teaching cycle'. According to this idea of a mathematical teaching cycle, a mathematics teacher will first try to anticipate in advance what the mental activities of the students will be when they will participate in some envisioned instructional activities, and next will try to find out to what extend the actual thinking processes of the students correspond with the hypothesized ones during the enactment of those activities, to finally reconsider potential or revised follow-up activities. To characterize the teacher's thinking, Simon coins the term, 'hypothetical learning trajectory,' which he describes as: 'The consideration of the learning goal, the learning activities, and the thinking and learning in which the students might engage (...)' (Simon, 1995, p. 133). The mathematical teaching cycle, then, may be described as conjecturing, enacting, and revising hypothetical learning trajectories.

We may compare the micro cycles of design and analysis with the concept of an empirical cycle of hypotheses testing. A fundamental difference, however, is that the evaluation of the former concerns inferences about the mental activities of the students, not merely observable behavior of the students. Since, for the design researcher, the goal is not just to find out whether the participation of the students in those particular activities results in certain anticipated behaviors, but to understand the relation between the student's participation and the conjectured mental activities.
To give an example of such more encompassing conjectures we may return to our example of statistics.

Earlier we stated that one of our initial goals was that the students would actually be analyzing data, not just numbers without context. With that in mind, we instituted a process that we called 'talking through the process of data creation'. On the basis of pragmatic considerations, and since our focus was on data analysis, we did not involve the students in activities of data gathering. We did not, however, want the data to drop out of thin air for the students. Moreover, following Cobb and Tzou (2009), we would argue that data are not ready available; data are created. Data are the result of measuring, and often specific measures are construed to find an answer to a certain question. We conjectured that it would be essential for students to experience this process of creating data to answer a question if data were to be measures rather than mere numbers for them. We may illustrate this with an example.
In one of the initial instructional activities, we wanted the students to compare data on a life span of two brands of batteries. However, it was important that they do so for a reason that they considered legitimate. The teacher therefore began by asking the students if they used batteries, and what do they used them for. They told that they used them in portable CD-players, tape recorders, and so forth. So, for them the quality of batteries appeared to

be a significant issue. Next the teacher asked about the things that they focus on when buying batteries. The students came up with life span and costs. So together teacher and students identified life span as a relevant dimension. Then the discussion turned to how to figure out which of two different brands of batteries would have the better life span. And the students were asked to come up with ideas about how to make measurements. They offered various proposals, often the idea came up of putting a number of batteries in 'identical' appliances, everything form torch flash lights or torches, to clocks, to whatever. It was only against that background of actually having talked through the data creation process that the data the students were to analyze were introduced. In doing so, we conjectured that as a consequence of engaging in this process the data that were introduced would have a history for the students. As shown in Figure 2, the data on the life-span data of two brands of batteries, which are presented by 'magnitude-value bars' in the first computer minitool.

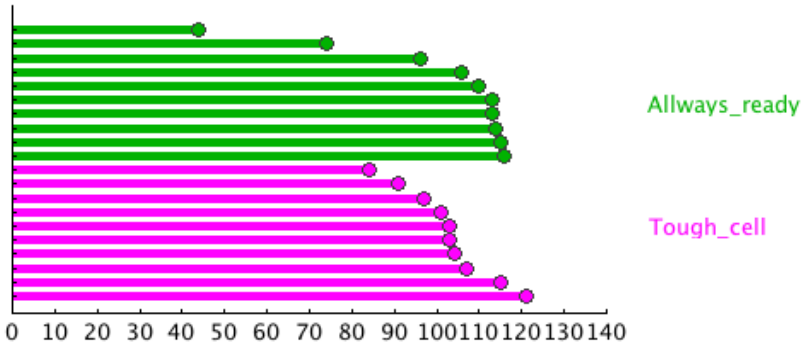


Figure 2: Two data sets in Minitool 1
Lifespans of Always Ready and Tough Cell batteries in hours

Each bar signifies the life span of a single battery. This computer tool has a number of options; the students can for example sort the bars by size or by the colors that correspond with different sub sets. When we introduced this type of visual representation, we purposely chose situations with linearity, such as time, that in our view would fit with this representation. We conjectured that this representation would be relatively transparent for the students thanks to their experience with scale lines and the like. We further conjectured that the students would focus on the position of the end points of the bars when comparing the data sets, and that the combination of a significant number of high values of the Always Ready batteries in combination with a few short life spans would create opportunities for a productive discussion.

In this illustration we focused on various conjectures, such as the conjecture that by engaging the students in the task of comparing two sets of data, which differed markedly in distribution of data values—while using the first minitool—would lead to a discussion about how the data values are distributed. We would be negligent if we did not clarify that the actual conjectures were in fact more complex, in that they also encompassed choices about organization of the classroom activities and classroom norms, as well as the nature of instructional activities and tools. These are relatively detailed conjectures about the means of supporting shifts in students' reasoning that we anticipated would be important.

As a clarifying note, it is helpful to distinguish between two complementary ways of identifying causal relations, the regularity conception of causality that is connected to observed regularities, and a process oriented conception of causal explanation, 'that sees causality as fundamentally referring to the actual causal mechanisms and processes that are involved in particular events and situations' (Maxwell, 2004, 4). Within the latter, 'causal explanation' refers to 'the mechanisms through which and the conditions under which that causal relationship holds' (Shadish, Cook, & Campbell, 2002, cited in Maxwell, 2004, p. 4). In contrast to regularity conception of causality that is connected to observed regularities, causal explanation can in principle be identified in a single case (Maxwell, 2004, p. 6). These mechanisms are exactly the kind of causal explanation that the design researchers seek to develop when atempting to understand how certain learning ecologies foster given forms of learning. In this sense, the micro cycles of thought- and instruction experiments correspond to a process-oriented conception of causal explanation, while the empirical cycle corresponds with regularity conception of causality. Note, however, that in the context of design research, it will not be sufficient to come to understand one student's thinking. Instead, to be of value, the researchers must document that a significant proportion of students reason in a comparable manner. In addition, regularities in the variation in student thinking will be essential for productive classroom discussions.

In a design experiment, the mini cycles of thought and instruction experiments serve the development of the local instruction theory. In fact there is a reflexive relation between the thought and instruction experiments, and the local instruction theory that is being developed. At one hand, the conjectured local instruction theory guides the thought and instruction experiments, and at the other hand, the micro cycles of design and analysis shape the local instruction theory (Figure 3).

Figure 3: Reflexive relation between theory and experiments

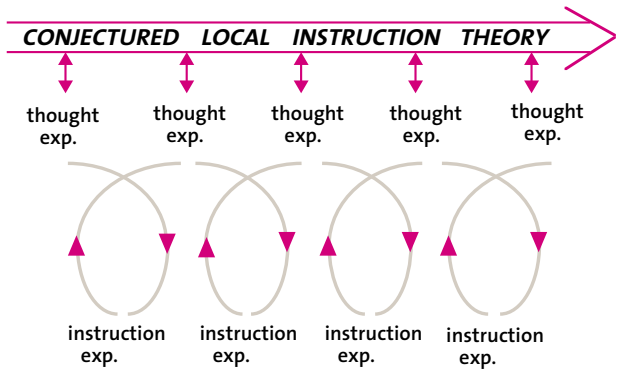
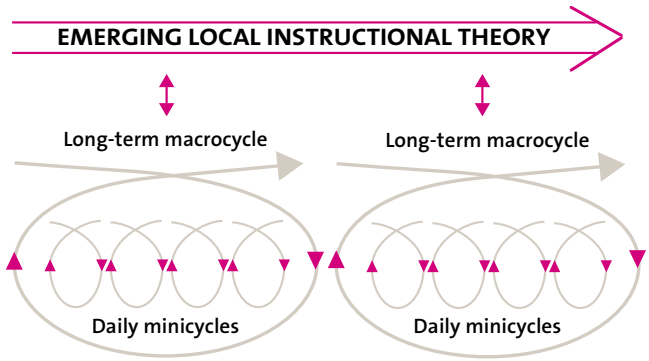


Figure 4: Micro and macro design cycles



These micro cycles require that the research team engages in an ongoing analysis of individual students' activity and of classroom social processes to inform new anticipatory thought experiments, the design or revision of instructional activities, and sometimes the modification of learning goals. In service of such an analysis, it is critical in our experience that the researchers are present in the classroom when the design experiment is in progress, and conduct short debriefing sessions with the collaborating teacher immediately after each classroom session in order to develop shared interpretations of what might be going on in the classroom.

We also find it vital to have longer periodic meetings. The focus of these meetings is primarily on the conjectured local instruction theory as a whole. A local instruction theory encompasses both the overall process of learning and the instructional activities that are designed to foster the mental activities that constitute the long-term process. So we may also observe a process of conjecturing and revising on two levels, on the level of the individual classroom sessions, and on the level of the instructional sequence as a whole. In

addition to the adaptation of the overall learning process during a design experiment, we may also discern macro design cycles, which span entire experiments, in the sense that the retrospective analysis of a design experiment can feed forward to inform a subsequent experiment (see Figure 4).
These new macrocycles may also involve other groups of learners or different settings—in other countries for instance. From this process emerges a more robust local instructional theory that, we would add, still is potentially revisable.

Data generation

Decisions about the types of data that need to be generated in the course of an experiment depend on the theoretical intent of the design experiment. These are in a sense pragmatic decisions in that the data have to make it possible for the researchers to address the issues that were identified as the theoretical intent at the start of the design experiment. If the design experiment focuses on the development of a local instruction theory, for instance, it makes sense to video record all classroom sessions, to conduct pre- and post-interviews with the students, to make copies of all of the students' work, and to assemble field notes. In addition, appropriate benchmark assessment items that have been used by other researchers might be incorporated if they are available. Usually, a wide collection of data will be needed because the data have to enable the research team to document the collective mathematical development of the classroom community, the developing mathematical reasoning of individual students, and the emerging learning ecology.

We also find it crucial to audio-record the regular research group meetings because these meetings offer one of the best opportunities to document the learning process of the research team. Data generation therefore involves keeping a log of the ongoing interpretations, conjectures, decisions, and so forth.
The specific foci of a design experiment may require additional types of data. To give an illustration, we return to the statistics experiment again, which also became a case of cultivating students' mathematical interests (Cobb & Hodge, 2003). We were therefore interested in how the students perceived their obligations in the classroom and in how they evaluated those obligations. As a consequence, a member of the research team conducted student interviews that focused on these issues while the experiment was in progress. It turned out to be more productive to conduct these interviews with pairs or groups of three students. So, this specific research interest necessitated another form of data collection.

Interpretative framework(s)

A key element in the ongoing process of experimentation is the interpretation of both the students' reasoning and learning and the means by which that learning is supported and organized. We contend that it is important to be explicit about how one is going about interpreting what is going on in the classroom.

In (quasi-)experimental research, the relation between the empirical reality and the scientific interpretation is made explicit by operationalizing the variables that are taken into account. Likewise, design researchers have to explicate how they translate observations of events in the classroom into scientific interpretations. The researchers will necessarily employ an interpretive framework to make sense of the complexity and messiness of classroom events both while a design experiment is in progress and when conducting a retrospective analysis of the data generated during an experiment. It is essential in our view that researchers explicate the basic constructs of their interpretive framework if inquiry is to be disciplined and systematic. Key elements of such a (potentially revisable) interpretative framework include (a) a framework for interpreting the evolving classroom learning environment, and (b) a framework for interpreting student mathematical reasoning and learning mathematics. In the following we will first discuss the framework we use to interpret classroom discourse and communication, and next turn to the domain specific instruction theory for realistic mathematics education that is used as a conceptual framework for interpreting student learning. In doing so, we clarify that for us socio-constructivism functions as a background theory.

Emergent perspective

The framework that we currently use for interpreting classroom discourse and communication is the 'emergent perspective' (Cobb & Yackel, 1996; Yackel & Cobb, 1996) (see Figure 5). We mentioned aspects of this framework earlier as examples of an ontological innovation.

Social Perspective	Psychological Perspective
Classroom social norms	Beliefs about our own role, others' roles, and the general nature of mathematical activity
Socio-mathematical norms	Specifically mathematical beliefs and values
Classroom mathematical practices	Mathematical conceptions and activity

Figure 5: An interpretive framework for analyzing individual and collective activity at the classroom level

The framework can be viewed as a response to the issue of attempting to understand mathematical learning as it occurs in the social context of the classroom. With regard to the specifics of the framework, the column headings 'Social Perspective' and 'Psychological Perspective' involve a focus on the classroom community and on individual students' reasoning respectively. In the following paragraphs, we first discuss social norms, then socio-mathematical norms, and finally classroom mathematical practices.

Social norms refer to expected ways of acting and explaining that become instantiated through a process of mutual negotiation between the teacher and students. The social norms will differ significantly between classrooms that pursue traditional school mathematics, or reform mathematics. In traditional mathematics classrooms, the role of the teacher is to explain and evaluate, while the social norms include the obligation of the students to try to figure out what the teacher has in mind, and to act accordingly. Examples of norms for whole-class discussions in reform math classrooms include obligations for the students to explain and justify solutions, to attempt to make sense of explanations given by others, to indicate agreement and disagreement, and to question alternatives in situations where a conflict in interpretations or solutions has become apparent.
The psychological correlate to social norms concerns the teacher's and students' individual beliefs about their own and others' roles. The reflexivity between social norms and individual beliefs is better understood when analyzing the negotiation process of classroom communities. On the one hand, individuals' beliefs about ways to act contribute to the negotiation of social norms. On the other hand, an individual's beliefs are enabled and constrained as he or she participates in this negotiation process.

The *socio-mathematical norms* can be distinguished from social norms as ways of explicating and acting in whole-class discussions that are specific to mathematics. Examples of such socio-mathematical norms include what counts as a different mathematical solution, a sophisticated mathematical solution, an efficient mathematical solution, and an acceptable mathematical explanation and justification. The students' personal beliefs about what makes a contribution acceptable, different, sophisticated or efficient encompasses the psychological correlate of the socio-mathematical norms. Students develop personal ways of judging whether a solution is efficient or different, and these beliefs are mutually negotiated as the classroom microculture is continually being structured. That is, the teacher cannot merely state specific guidelines for what types of solutions are acceptable and expect the guidelines to be understood and enacted by students. Instead, socio-mathematical norms are continually negotiated and redefined as the teacher and students participate in discussions.
The analysis of socio-mathematical norms has proven to be pragmatically significant when conducting design experiments in that it clarifies the process by which teachers

may foster the development of intellectual autonomy in their classrooms. To create the opportunity for the students to take over the teacher's responsibility as validators, socio-mathematical norms have to be in place that enable students to make independent judgments that contribute to the teacher's instructional agenda.

The last social aspect of the theoretical framework concerns the *mathematical practices* that are established in the classroom (see also Cobb, Stephan, McClain, & Gravemeijer, 2001). A mathematical practice can be described as the normative ways of acting, communicating and symbolizing mathematically at a given moment in time. In contrast to the socio-mathematical norms that are specific to mathematics, the mathematical practices are specific to particular mathematical ideas or concepts. In addition, mathematical practices necessarily evolve in the course of an experiment whereas socio-mathematical norms tend to be more stable. An indication that a certain mathematical practice has been established is that explanations pertaining to the particular practice have become beyond justification. Individual students' mathematical interpretations and actions constitute the psychological correlates of the classroom mathematical practices. Their interpretations and the mathematical practices are reflexively related in that students' mathematical development occurs as they contribute to the constitution of the mathematical practices. Conversely, the evolution of mathematical practices does not occur apart from students' reorganization of their individual activity.
We may conclude by noting that in the context of a design experiment, a detailed analysis of evolving classroom practices offers a way of describing the actual learning process of the classroom community as a whole. This offers a viable alternative for describing the learning process of the classroom rather than implying either that all students are learning in unison, or of attempting to describe the learning processes of each individual student.

RME theory

When discussing theoretical intent of design experiments, we noted that ontological innovations, such as interpretative frameworks, serve a dual role, both as lenses for making sense of what is happening in a real world instructional setting, and as guidelines or heuristics for instructional design. On the one hand, we may observe that although the emergent framework was initially developed to interpret classroom discourse and communication, it also offers guidelines on the classroom culture characteristics that fit the intended learning ecology. On the other hand, it may be observed that the RME theory not only offers design heuristics, but also may function as an interpretative framework for interpreting student activity in terms of learning mathematics.
In the following we elaborate this dual role of RME theory. Given its origin, we focus first on the instructional design perspective.

RME emerged at least in part in resistance to instructional and design approaches that treated mathematics as a ready-made product. Freudenthal (1971, 1973) argued that mathematics should primarily have the character of an activity for the students. A process of guided reinvention then would have to ensure that this mathematical activity fosters the construal of mathematics as a body of knowledge by the students. This requires that the instructional starting points are experientially real for the students, which means that one has to present the students problem situations in which they can reason and act in a personally meaningful manner. The objective of guided reinvention is that the mathematics that the students develop will also be experientially real for them. Learning mathematics should ideally be experienced as expanding one's mathematical reality.

We may further elaborate this point by clarifying the way in which Freudenthal conceives reality: 'I prefer to apply the term reality to what common sense experiences as real at a certain stage' (Freudenthal, 1991, 17). He goes on to say that reality is to be understood as a mixture of interpretation and sensual experience, which implies that mathematics, too, can become part of a person's reality. Reality and what a person perceives as common sense is not static but grows, and is affected by the individual's learning process. The goal of realistic mathematics education then is to support students in creating some new mathematical reality. This is to be realized by guided reinvention, or, 'mathematizing'—if we take a student perspective. Mathematizing literally means organizing subject matter to make it more mathematical[2]. The idea is that students should be given the opportunity to reinvent conventional mathematics by mathematizing both subject matter from reality and mathematical matter—the latter especially being their own mathematical activity—under guidance of the teacher (Freudenthal, 1971). An alternation of these two forms of mathematizing, which Treffers (1987) denotes mathematizing horizontally and vertically[3] should enable students to reach a higher level of mathematical thinking. On the one hand, mathematizing horizontally by translating contextual problems into mathematical problems may help students in grounding their mathematical thinking in their own experiential reality. On the other hand, mathematizing vertically, which may involve inventing new symbolizations, new conceptions or new solution procedures, allows for the constitution of some new mathematical reality, as '[t]he activity on one level is subjected to analysis on the next, the operational matter on one level becomes subject matter on the next level' (Freudenthal, 1971, p. 417). This shift from 'activity' to 'subject matter' relates to the shift from procedures to objects, which Sfard (1991) observed in the history of mathematics.

2) We may characterize mathematizing by pointing to the mathematical ambition to be general, exact, concise, and sure. Mathematizing thus involves activities such as generalizing, formalizing, curtailing and proving.
3) The grouping by Treffers (1987) suggests that all mathematical activity falls into one of these categories. However, an additional mathematical activity is indispensable, which concerns carrying out known mathematical operations (Gravemeijer, 2005).

If we look at the history of mathematics, we may observe that mathematics emerged from solving problems, or as Freudenthal puts it, from organizing subject matter. According to Freudenthal (1983), mathematical 'thought-things', such as concepts, tools and procedures, are invented to organize certain phenomena. The reinvention heuristic then suggests that the instructional designer should try to find situations that create the need for the students to invent the mathematical thought things the students are supposed to construct. To find such situations, the instructional designer should analyze the relation between those mathematical 'thought-things', and the phenomena they organize. This phenomenological analysis lays the basis for a didactical phenomenology (ibid), which also incorporates a discussion of what phenomenological analysis means from an educational perspective. For example, to construct distribution as a mathematical object, students should be confronted with situations where it is reasonable and sensible for them to achieve a goal by organizing phenomena in terms of distributions.

Freudenthal's level-theory also shaped the RME-view on educational models. Instead of ready-made models, RME looks for models that may emerge first as *models of* situated activity, and then gradually evolve into entities of their own to function as *models for* more sophisticated mathematical reasoning (Gravemeijer, 1999). According to this 'emergent-modeling' heuristic, the model and the new mathematical reality co-evolve; the emergence of the model is reflexively related to the development of some new mathematical reality. The teacher may support this process by supporting a shift in the students' attention from the context situation that the model refers to, towards the mathematical relations involved. In this manner, the students may develop a network of mathematical relations. Then the model can begin to function as a model for more sophisticated mathematical reasoning, in that the model derives its meaning from this network of mathematical relations. At the same time, junctions in this network of mathematical relations may become mathematical objects that constitute a new mathematical reality. As a further elucidation, we may note that, the term model should not be taken too literally in that it can also concern a model situation, or a model procedure. Moreover, what is taken as 'the model' from a more overarching design perspective will be constituted as a series of sub-models in the instructional activities.

As we argued before, the RME domain specific instruction theory also offers a framework for interpreting student activity in terms of learning mathematics (see also Gravemeijer, 1994). It orients the researcher to focus, for instance, on the various learning processes that might take place, with a special attention to the question of whether the students are inventing their own solution procedures or are merely imitating the teacher or some leading students. In such a case, one might look at the variety of students' solution procedures. On the basis of the reinvention principle, one would further expect to recognize the reinvention

route in the students' solutions. In addition, one would expect that the students would spontaneously drop back in their collective learning history when they are faced with new problems that represent difficulties for them. If they instead choose informal procedures that do not correspond with the reinvention route that has been followed, this would be an indication that that route is not experienced as a natural reinvention process.
In a similar manner, the researcher may investigate whether the models that are used fit with the informal solution procedures demonstrated by the students: Do the students use similar procedures with the model, as they did (or would do) without the model? In other words, the model must not dictate to the students how to proceed, but must be a resources that fits with their thought processes (Gravemeijer, 1993). Along these lines, the RME framework might generate additional points of focus, such as the following:

- Do the students rely on their own domain-specific knowledge?
- Do the instructional activities provide the expected traction for the students' informal solution procedures?
- Do the solutions that the students develop offer possibilities for vertical mathematization?
- Do the students mathematize their own informal mathematical activities?

And so forth. We will not, however, try to be exhaustive here. (For a more elaborated discussion see Gravemeijer, 1994, chapter 6).

We want to close this section on the second phase of the design experiment methodology by presenting a short sketch of the instructional sequence that was developed in the statistics design experiment.

We clarify the set-up of the statistics sequence by first describing how the didactical phenomenological analysis plays out in this case. The first step in this analysis was to analyze the notion of distribution as a mathematical (or statistical) thought thing. This led to the conclusion that distribution can be thought of as a density function, indicating that density can be conceived of as that which is organized by distribution as a thought thing. Density—as a thought thing in and of itself—in turn organizes collections of data points in a space of possible data values. This insight can be made concrete as a dot plot, showing data points on an axis (these data points on an axis can be viewed as thought things that organize data values, while the height of the accumulation of data in a given point can be interpreted as a measure for density). The measures can in turn be thought of as a means for getting a handle on some real world phenomena; the notion of data creation can also be construed as a form of organizing.
This phenomenological analysis reveals a possible reinvention route in which a cumulative process of organizing would lead the students through the above steps in reverse order. This lays the basis for the following instructional sequence.

Point of departure is a bottom-up approach in which the computer minitools are experienced by the students as sensible tools to use given their current conceptions of analyzing data. So for the students, the primary function of the minitools is to help them structure and describe data sets in order to make a decision or judgment. In this process, notions such as mean, mode, median, skewness, spreadoutness, and relative frequency may emerge as ways of describing how specific datasets are distributed within this space of values. Further, in this approach, various statistical representations or inscriptions may emerge as different ways of structuring distributions. In fact, the minitools are so designed, that they can support a process of progressive mathematization by which these conventional statistical tools are reinvented.
At the same time, the activity of structuring data sets by using the minitools fosters a process by which the students come to view data sets as entities that are distributed within a space of possible values. The intent is to support a process in which the means of symbolizing, and the meaning of what these symbolizations signify for the students co-evolve, similar to that which Meira (1995) describes when he speaks of a 'dialectical relation between notations-in-use and mathematical sense making' (Meira, 1995, p. 270; see also Roth & McGinn, 1998; Cobb, 2002; and Gravemeijer, 2002).

The backbone of the sequence consists of a series of symbolic representations that are embedded in the computer tools. The idea is that the activities with the computer tools succeed each other in such a manner that the activity with the newer tool is experienced as a natural extension of the activity with the earlier tool. The starting point is in the measures, or magnitudes, that constitute a data set. With the first minitool, magnitude-value bars (Figure 2) are introduced where each value bar signifies a single measure. (Initially, the measures under investigation are of a linear type, like 'length', and 'time'. Later, this is generalized to other types of measures.) We conjectured that as a consequence of participating in discussions about various data sets represented by value bars, the students would begin to focus on the end points of a value bars. As a consequence, these end points come to signify the corresponding value bars. This allows for the introduction of a line plot as a more condense inscription that omits the value bars and preserves only the end points (Figure 8). The second minitool offers students a range of options for structuring data sets represented as line plots that include creating equal intervals, creating two equal groups, and creating four equal groups of data points. We conjectured that as a result of analyzing data sets by using these options, the students would begin to reason about data in terms of density, and come to see the shape of the line plot as signifying the distribution of data values in terms of density.

In retrospect, we may recognize the emergent-models design heuristic with 'a graphical representation of the shape of a distribution' as the overarching model. This overarching

model is instantiated by various sub-models that change over time. The graph was initially introduced in an informal manner, as a way of inscribing a set of measures by representing each measure by a bar (Figure 2). We can see this as a pre-stage of the model, where the set of measures is still very much tied to the situation. Nonetheless, from a statistical perspective, shape of the distribution is visible in the way the endpoints are distributed in regard to the axis. In this phase, we can speak of the graphical representation as a *model of* a set of measures. Next we introduced activities that were designed to draw the students' attention to distribution of the end points of the bars. This supported the introduction of the line plot, where the second minitool was used to structure data sets in various ways to answer the questions at hand. Analyses that involved structuring the data into four equal groups with the corresponding tool option (which anticipates the box plot) were particularly important in drawing the students' attention to distribution of density. This then supported a gradual shift from seeing the graph as signifying as a set of measures to seeing it as signifying a distribution. If once this latter shift occurred, the graph could be used to reason about distributions. Students could, for instance, discern various types of distributions (with the normal distributions as one of them), and could reason about characteristics of (univariate) distributions, like skewness (Figure 6). The model had then become a model for reasoning about distributions.

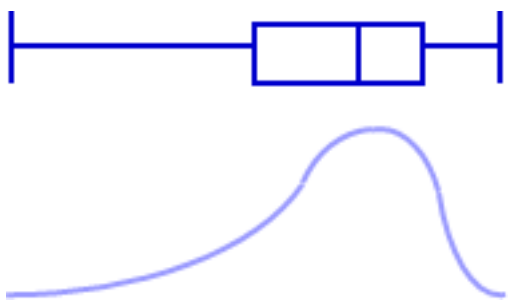


Figure 6: Box plot as a model for reasoning about distributions

Phase three, the retrospective analysis

Thus far, we have discussed the planning of a design experiment and the prolonged experimentation in the classroom that is central to the methodology. A further aspect of the methodology concerns the retrospective analyses that are conducted of the entire data set collected during the experiment. The goal of the retrospective analyses will of course depend on the theoretical intent of the design experiment. However, one of the primary aims is typically to contribute to the development of a local instruction theory. Other goals may concern more encompassing issues, or ontological innovations. Although differences in theoretical objectives are reflected in differences in the retrospective analyses, the form of the analysis will necessarily involve an iterative process of analyzing the entire data set. We will, therefore, first describe the retrospective analyses in general, and then discuss analyses to develop a local instruction theory, and next analyses conducted to address more general research topics.

Since the primary aim of the design research we are discussing here, is to come to understand how the learning ecology accounts for the learning process of the students, a broad array of data will be assembled to capture both the evolving learning ecology and the progress in thinking and reasoning of the students during the design experiment. The data sets typically include (but are not limited to) video-recordings of all classroom lessons, video-recorded individual interviews conducted with all students before and after the experiment, or pre- and posttests, to assess their mathematical learning, copies of all the students' written work, field notes, and audio-recordings of both the daily debriefing session and weekly project meetings. The challenge then is to analyze this comprehensive data set systematically while simultaneously documenting the grounds for particular inferences. Claims will be based on a retrospective, systematic and thorough analysis of the entire data set collected during the experiment. To ascertain the credibility of the analysis, all phases of the analysis process have to be documented, including the refining and refuting of conjectures. Final claims and assertions can then be justified by backtracking through the various levels of the analysis, if necessary to the original video-recordings and transcripts. It is this documentation of the research team's data analysis process that provides an empirical grounding for the analysis. Further, it provides a means of differentiating systematic analyses in which sample episodes are used to illustrate general assertions from questionable analyses in which a few possibly atypical episodes are used to support unsubstantiated claims. Additional criteria that enhance the trustworthiness of an analysis include both the extent to which it has been critiqued by other researchers who do not have a stake in the success of the experiment and the extent to which it derives from a prolonged engagement with students and teachers (Taylor & Bogdan, 1984). This latter criterion is typically satisfied in the case of classroom design experiments and constitutes a strength of the methodology.

The specific approach we use is a variant of Glaser and Strauss's (1967) constant comparative method (see also Cobb & Whitenack, 1996). This is an iterative method that aims to ascertain what the data tell us about the experiment. Here we may start with a first round of analysis to develop an initial image of *what happened* during the teaching experiment. This initial image will be cast in terms of conjectures, which are tested against the whole data set. The results of this analysis are taken as the basis for a next round of analysis which may aim at identifying patterns or *explanations*, followed by a second round of conjecture testing.

One of the aims of the analysis might, for example, be to describe the actual learning process of the classroom community as a whole. In order to describe this learning process we perform a detailed analysis of the consecutive classroom practices. To determine how the mathematical practices developed, we work through the data chronologically. An indication that a certain mathematical practice has been established is that explanations pertaining to the particular practice have become beyond justification. When checking whether a particular mathematical practice has been established, one of the key criteria is that a student who appears to violate a norm of argumentation that corresponds with that practice will be challenged by their peers. If we find instances where such challenges do not occur, we either have to revise our conjecture about the mathematical practices that have been established. This may also signify that the practices have evolved, and a new practice has replaced the mathematical practice under consideration.

As a result of this first round of data analysis, we end up with a sequence of conjectures and refutations that are tied to specific episodes. In the second phase of a retrospective analysis, this sequence of conjectures and refutations are in effect treated as a new data set that has to be analyzed. It is while 'meta-analyzing' these episode-specific conjectures, confirmations and refutations, that particular episodes become to be seen as pivotal. And they are pivotal in the context of the analysis, because they allow us to decide between two or more competing conjectures. These are the episodes that are typically included in research reports. As an illustration, we present some typical episodes from the statistics design experiment.

We already described the battery lifespan problem in which the data were represented as magnitude bars in the first computer tool. The students first worked on this problem in groups, and then the teacher initiated a whole class discussion of the students' analyses. The computer tool was projected on an overhead screen, the data were sorted the data by size, and the so-called 'range tool' option was used to highlight the ten highest data values (see Figure 7).

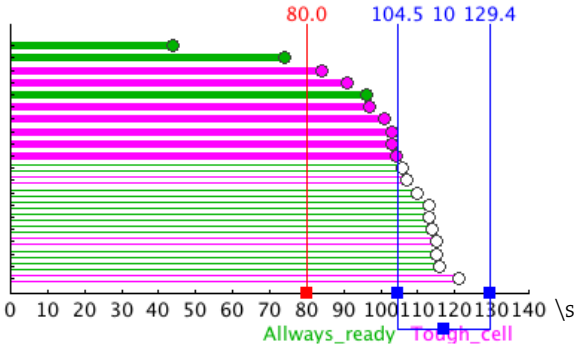


Figure 7: Battery life span data: Always ready and tough cell batteries

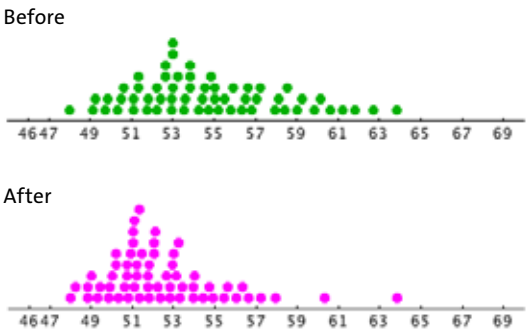


Figure 8: Speed data, before and after a speed campaign

One of the students, Casey, argued that the green batteries were better because seven of the top ten were green (Always Ready), and her argument is supported by another student.

Janice: She's saying that out of ten of the batteries that lasted the longest, seven of them are green, and that's the most number, so the Always Ready batteries are better because more of those batteries lasted longer.

However, this argument was challenged by another student, James, who argued that four of the pink bars (Tough Cell) were 'almost in that area and then if you put all those in you would have seven (rather than three pinks).'
Later in the discussion, Brad asked for the value tool (the single vertical line) to be placed at 80, in order to substantiate his claim that the Tough Cell brand is better.

Brad: See, there's still green ones (Always Ready) behind 80, but all of the Tough Cell is above 80. I would rather have a consistent battery that I know will get me over 80 hours than one that you just try to guess.

One of the issues of interest in this episode is the use of the word 'consistent', which the students introduce as an informal way of describing the extent to which data sets were bunched up or spread out. This episode also proved to be pivotal in documenting that a norm of argumentation was being established, namely that students were obliged to explain why the way in which they had partitioned or organized or structured the data gave insight into the problem or issue under investigation. We were able to demonstrate that this norm remained stable throughout the experiment.

A second illustrative episode concerns a comparison of two sets of data that showed the speeds of cars before, respectively after, a campaign against speeding (Figure 8).

In this case, one of the students had focused on the shape of the data sets to compare how they were distributed.

Janice: If you look at the graphs and look at them like hills, then for the before group the speeds are spread out and more than 55, and if you look at the after graph, then more people are bunched up close to the speed limit which means that the majority of the people slowed down close to the speed limit.

What is of interest here is that this student did not use the word 'hill' to refer to the figural image, but instead used it as a metaphor to describe the distribution of the density of the data ('bunched up, close') as giving her insight into the effectiveness of the campaign against speeding. The students continued to use this metaphor throughout the design experiment to indicate that the 'majority' of the data points were 'bunched-up'. In a follow up experiment, we found that the students could even identify where the hill was in the value-bar representation of the first computer minitool (Bakker, 2004), which underscores the metaphorical character of this term. This circumstance is especially noteworthy since the students had to construe a meaning for the vertical axis of the line plot by themselves. They had to construe the height of the dots at a specific position as a measure for the density of the data points at that point.
Later on the students started to use the hill metaphor in conjunction with the four-equal-groups representation, in which five vertical bars split a data set in four groups with the same number of data points (Figure 9a). They found out that the data were bunched up where two bars were closest to each other, thereby indicating the position of the hill.[4]

4) While implicitly assuming an uni-model distribution.

Eventually they could 'see the hill' in a four-equal-groups representation with the data points hidden (Figure 9b).

Figure 9a and b: Speed data, before and after a speed campaign

As a third example we may describe an episode in which the students had to compare data on T-cell counts for two different treatments for AIDS-patients, an experimental treatment with 46 patients, and a standard treatment with 186, where the goal is to raise the patients' T-cell counts. Various groups of students analyzed these data in a range of different ways. One group of students identified the intervals where the 'hill' was located in each data set, where the data were bunched up. And on this basis they argued that the new, experimental treatment was effective, because the 'hill' was in a higher interval than the hill in the standard treatment data. Another group of students had used the four-equal-groups option (Figure 10).

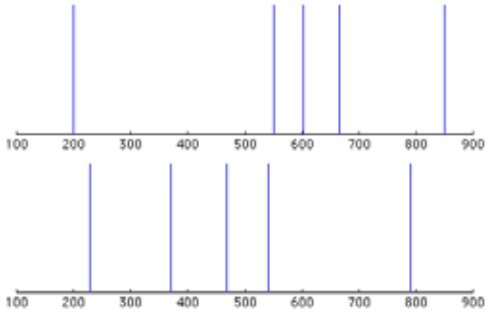


Figure 10: T-cell data, four-equal-groups inscription, with data points hidden

This is a precursor of the box plot in that each interval contains 25% of the data. They had used another available option to hide the dots. Their argument was: the new treatment is better because 75% of the data is above 550, whereas in the traditional treatment 75% is below. Note that we could picture the shape of the hill in this representation, if we knew this was a uni-modal distribution.

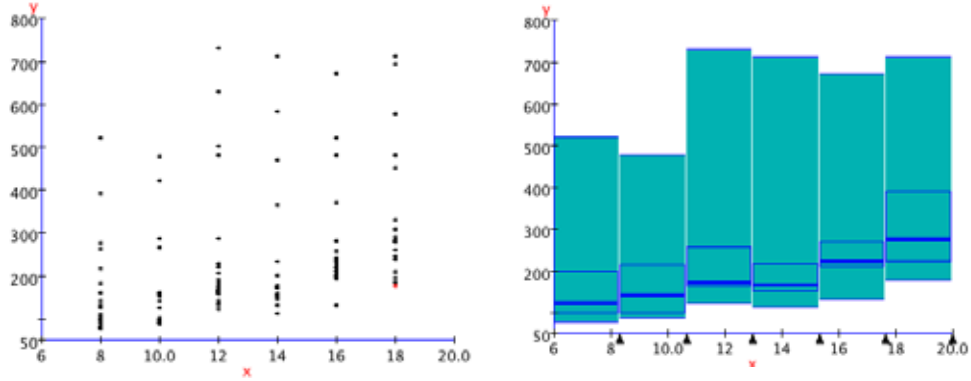


Figure 11a and b: Salary against years of education

We may briefly show why this notion of shape of a univariate distribution became important for analyzing bi-variate data in a subsequent design experiment conducted the following school year with some of the same students. In this follow-up experiment, we asked the students to compare, for instance, data on years of education, against salary levels for men, and women. The students analyzed data of this type by using a third computer minitool (figure 11a). With this tool, vertical 'slices' of the shown data points could be treated as a series of univariate data sets with vertical axes, and the four-equal-groups option (rotated 90 degrees) could be applied on each of these sub-sets (Figure 11b).

Here in doing so, the students typically talked about where the 'hill' was located or where the 'clutter' was in the data. As the students discovered, the ranges were similar for the men's and women's salary levels. The big difference was that the data for females was skewed much more heavily towards the bottom end of the distribution for each level of education. As this example clarifies, analyzing bi-variate data is not so much about drawing a line through a cloud of dots, but about investigating how the distribution of the dependent variable changes as the independent variable changes.

Reconstructing the local instruction theory

One of the primary aims of a retrospective analysis is to support the constitution of a revised local instruction theory. However, it is important to emphasize that the results of design experiments cannot be linked to pre and posttests results in the same direct manner as is common in a standard formative evaluation. This is because the revised local instruction theory and prototypical instructional sequence will differ from what was tried out in the classroom. Because of the testing and revising of instructional activities (and the corresponding conjectures) while the design experiment was in progress, a revised, potentially superior instructional sequence has to be construed afterwards. It does not

make sense, for example, to include instructional activities that did not fulfill expectations, but the fact that these activities were enacted in the experiment will nonetheless have affected the students' learning. Adaptations will therefore have to be made when the non-, or less-functional activities are left out. Consequently, the instructional sequence will be put together by focusing on and reconstructing the instructional activities that proved to constitute the effective elements of the sequence—assessed by the types of mathematical activity in which the students engaged when these instructional activities were used in the classroom. This reconstruction of an improved sequence will be based on the observations and inferences made during the design experiment, complemented by the insights gained by conducting retrospective analyses. In this manner, it can be claimed that the results of a design experiment are empirically grounded even though the pre and posttests results do not cover the reconstructed instructional sequence.

As a point of clarification we may add that, although the constitution of a revised local instruction theory is primarily a reconstruction activity, the retrospective analysis may spark design ideas that go beyond those that were tried out in the classroom. These insights might in turn create the need for a new experiment, starting with a new conjectured local instruction theory. Here, the cyclic nature of the methodology that we noted at the level of instructional design micro-cycles reappears at a broader level. An entire design experiment and the subsequent retrospective analysis together constitute a larger, macro-cycle of design and analysis (Figure 4).

In this cycle, the conjectures and assumptions formulated at the outset when planning a design experiment are scrutinized in the retrospective analysis. An example of such an analysis can be found in Cobb, Gravemeijer, Yackel, McClain, and Whitenack (1997). Here, the retrospective analysis indicated that several key assumptions that underpinned an instructional sequence were ill founded. As a consequence, the instructional sequence was radically revised and a further design experiment was conducted. An extensive report of this largely successful follow-up experiment can be found in Stephan, Bowers, Cobb, and Gravemeijer (2003).

Encompassing issues and ontological innovations

In addition to retrospective analyses that directly aim at the reconstruction and revision of a local instructional theory, a retrospective analysis might be conducted to place classroom events in a broader context by framing them as instances of more encompassing issues. Earlier, we mentioned as examples analyses that focus on the role of the teacher, the teacher's learning, the role of semiotic processes, or on the process of cultivating the students' mathematical interests. In addition we mentioned ontological innovations, which might include issues such as the interpretative framework for interpreting classroom discourse and communication, meta-representational competence, quantitative reasoning or emergent modeling.

In such cases, the aim of the analysis is to frame events that occurred in the design experiment classroom as instances, or paradigm cases, of a broader class of phenomena. The goal is to come to understand (the role of) the specific characteristics of the investigated learning ecology in order to develop theoretical tools that make it possible to come to grips with the same phenomenon in other learning ecologies. Data analysis that aims at understanding a paradigm case differs significantly from data analyses that aim at establishing causal relations within a regularity conception of causality. Claims are not based on statistical analysis, but on a systematic and thorough analysis of the data set.

Virtual replicability

Metaphorically speaking, the course of a design experiment can be characterized in terms of the learning process of the research team. We would argue that this learning process has to justify the products of the research project. This characterization is especially fitting for the construal of the local instruction theory, which encompasses two processes, (a) the learning process that is inherent to the cyclic process of (re)designing and testing instructional activities and other aspects of the initial design, and (b) the retrospective analysis that scrutinizes, and builds on, this primary process, and looks for patterns that may explain the progress of the students. In relation to this learning process, we can refer to the methodological norm of 'trackability' that is used as a criterion in ethnographic research. Smaling (1990, 1992) connects trackability with the well-known criterion of 'reliability'. He notes that reliability refers to the absence of accidental errors and is often defined as reproducibility. He goes on to say, that for qualitative research this means virtual replicability. Here the emphasis is on virtual. It is important that the research is reported in such a manner that it can be retraced, or virtually replicated by other researchers. This ethnographic norm of trackability fits with Freudenthal's conception of developmental or design research:

> Developmental research means: 'experiencing the cyclic process of development and research so consciously, and reporting on it so candidly that it justifies itself, and this experience can be transmitted to others to become like their own experience.' (Freudenthal, 1991, p. 161)

Likewise, Smaling (1990, p. 6) states that trackability can be established by reporting on, 'failures and successes, on the procedures followed, on the conceptual framework and on the reasons for the choices made'. Note that this norm of trackability does not necessarily require that everyone has to subscribe the conclusions of the researchers. Eventually, outsiders, who have virtually replicated the learning process of the researchers, may interpret their experiences differently or come to different conclusions on the same experiential basis. The power of this approach is that it creates an experiential basis for discussion.

Ecological validity

A central assumption that underpins our work is that instructional innovations developed in the course of a design research experiment can be used productively to support students' learning in other classrooms. However, as we know only too well, the history of research in education in general, and in mathematics education in particular, is replete with more than its share of disparate and often irreconcilable findings. A primary source of difficulty is that the independent variables of traditional experimental research are often relatively superficial and have little to do with either context or meaning. As a consequence, it has frequently been impossible to account for the differences in findings when different groups of students receive supposedly the same instructional treatment. In contrast to traditional experimental research, the challenge when conducting design experiments is not that of replicating instructional innovations by ensuring that they are realized in precisely the same way in different classrooms. The conception of teachers as professionals who continually adjust their plans on the basis of ongoing assessments of their students' mathematical understanding in fact suggests that complete replicability is neither desirably nor, perhaps, possible (cf. Ball, 1993; Simon, 1995). Design research aims for ecological validity, that is to say, (the description of) the results should provide a basis for adaptation to other situations. The premise is that an empirically grounded theory of how the intervention works accommodates this requirement. Therefore, one of the primary aims of this type of research is not to develop the instructional sequence as such, but to support the constitution of an empirically grounded local instruction theory that underpins that instructional sequence. The intent is to develop a local instruction theory that can function as frame of reference for teachers who want to adapt the corresponding instructional sequence to their own classrooms, and their personal objectives. One element that can be helpful in this respect, is offering, what is called, a 'thick description' of what happened in the design experiment. By describing details of the participating students, of the teaching-learning process, and so forth, together with an analysis of how these elements may have influenced the whole process, outsiders will have a basis for deliberating adjustments to other situations. Conversely, feedback from teachers on how the instructional sequence was adjusted to accommodate various classrooms can strengthen the ecological validity significantly. We therefore find it critical to have repeated trials in a variety of settings.

In the case of the statistics sequence, for example, we worked with middle school students, with 'at risk' high school students, perspective elementary teachers, practicing teachers, and there have also been follow-up groups, including a series of design experiments by Arthur Bakker (2004), in The Netherlands. We have been surprised by the extent to which he have been able to document regularities in the development of the participants' thinking across these various settings. That is to say, there is diversity in how a group of participants reasoned at any point in time. But we were able predict with some confidence

the primary types of analyses or forms of reasoning within a group at any point in the experiment. We think that is useful knowledge from a teacher's point of view in that enables teachers to anticipate the types of reasoning that they can build on or work with.

Developing domain specific instruction theories

Design research provides a means of developing local instruction theories that can serve to support for teachers who adapt instructional sequences as part of their teaching practice. In addition, design research also contributes to the development of a domain specific instruction theory, in our case the RME theory. This theory emerges in an iterative, cumulative process that embraces a series of design research projects. In this regard, we can speak of theory development at various levels:

- At the level of the instructional activities (micro theories)
- At the level of the instructional sequence (local instruction theories)
- At the level of the domain-specific instruction theory.

The relations between these levels can be clarified by drawing on the distinction that Kessels and Korthagen (1996) make between 'episteme' and 'phronesis'. Following Aristotle, they use the Greek word episteme to refer to scientific knowledge, and the word phronesis to refer to 'practical wisdom'. They argue that the incompatibility of the products of scientific research with the needs of teachers can be traced to the contrast between these two realms. Teachers rely on practical wisdom, which they share with one another in the form of narratives. They experience scientific knowledge that is produced by research as too abstract and too general to directly inform their practice (see also Hiebert & Stigler, 1999). In design research, scientific knowledge is grounded in practical wisdom while simultaneously providing heuristics and practical theories that can strengthen the practical wisdom of teachers. In this respect, we would argue that design research has the potential to bridge the gap between theory and practice, as domain-specific instruction theory can be categorized as episteme and micro-didactical theories as phronesis. We stress the word "potential" here because the types of innovative instructional practice developed in our teaching experiments differ significantly from the current practice in most classrooms. Consequently, most teachers will need sustained support to develop these types of instructional practice (Stephan, Underwood-Gregg, & Yackel, in press).

In recent years, an increasing number of design experiments have been conducted with teachers to investigate professional development that aims to support them in reorganizing their instructional practices. The findings of this work indicate the importance of organizing professional development around specific high-leverage instructional practices (Ball, Sleep, Boerst, & Bass, 2009) and of creating opportunities for teachers to both investigate and enact those practices (Grossman, Compton, Igra,

et al., 2009). As Grossman and colleagues and have noted, teacher education tends to emphasize *pedagogies of investigation* at the expense of *pedagogies of enactment*. Pedagogies of investigation involve analyzing and critiquing representations of practice such as student work and video-cases of teaching (Borko, Jacobs, Eiteljorg, & Pittman, 2009; Sherin & Han, 2004). Pedagogies of enactment involve planning for, rehearsing, and enacting aspects of practice in a graduated sequence of increasingly complex settings (e.g., teaching other pre-service teachers who play the role of students, working with a small groups of students, teaching an entire class). Grossman et al. argue convincingly that pedagogies of investigation and enactment are both necessary if teachers are to develop classroom practices that focus on student reasoning. This claim is supported by studies of professional learning, which emphasize the value of co-participating in activities that approximate the targeted practices with colleagues who have already developed those practices (Bruner, 1996; Forman, 2003; Lave & Wenger, 1991).

Developing ways of analyzing innovations

A related challenge is that of developing ways of analyzing innovations that make their realization in different classrooms commensurable. An analysis of classroom events structured in terms of constructs such as social norms, socio-mathematical norms, and classroom mathematical practices serves to relate the students' mathematical learning in a particular classroom to their participation in sequences of instructional activities as they were realized in that classroom. As we noted earlier, classroom social norms, and socio-math norms can make a profound difference in the nature and the quality of the students' mathematical reasoning.

This part of the retrospective analysis raises its own methodological issues. A theoretical analysis is the result of a complex, purposeful problem-solving process. And because frameworks of reference will differ, one would not expect that different researchers would necessarily develop identical theoretical constructs when analyzing the same set of design experiment data. This implies that the notion of replicability is not relevant in this context. Following Atkinson, Delamont, and Hammersley (1988), we suggest that the relevant criteria are instead those of the generalizability and the trustworthiness of the constructs developed. We touched on the issue of generalizability when discussing the importance of viewing classroom events as paradigm cases of more encompassing issues. It is this framing of classroom activities and events as exemplars or prototypes that gives rise to generalizability. This, of course, is not generalization in the sense that the characteristics of particular cases are ignored and they are treated as interchangeable instances of the set to which assertions are claimed to apply. Instead, the theoretical analysis developed when coming to understand one case is deemed to be relevant when interpreting other cases. Thus, what is generalized is a way of interpreting and understanding specific cases that preserves their individual characteristics. For example, we conjectured that

much of what we learned when investigating symbolizing and modeling in a first-grade design experiment that focused on arithmetical reasoning would inform analyses of other students' mathematical learning in a wide range of classroom situations including those that involve the intensive use of technology. This in fact proved to be the case in sequences of design experiments that focused on the students' development of statistical reasoning (Cobb, 1999; Cobb, McClain, & Gravemeijer, 2003). It is this quest for generalizability that distinguishes analyses whose primary goal is to assess a particular instructional innovation from those whose goal is the development of theory that can feed forward to guide future research and instructional design.

Whereas generalizability is closely associated with the notion of a paradigm case, trustworthiness is concerned with the reasonableness and justifiability of inferences and assertions. This notion of trustworthiness acknowledges that a range of plausible analyses might be made of a given data set for a variety of different purposes. The issue at hand is that of the credibility of an analysis. As we have indicated, the most important consideration in this regard is the extent to which the analysis of the longitudinal data set is both systematic and thorough.

Different conceptions of design research

As a final methodological commentary, we want to position our approach to design research in the broader context of research perspectives. We may begin by noting that the process-oriented conception of causality (Maxwell, 2004) is a legitimate way of establishing causal relations. We may further point to the fact that we are aiming at explaining what happens in a specific teaching experiment in a given classroom. We are not aiming at statistical generalization by means of a representative sample that is based on the regularity type of causal descriptions. Instead, we aim at gaining greater understanding of learning ecologies, which are conceptualized as interacting systems rather than as a list of separate factors that influence learning (Cobb, Confrey, diSessa, Lehrer, & Schauble, 2003). In our conception of design research, generalizability is related to supporting others in using the products of a design experiment to inform their efforts to support learning in other settings. In this conception, generalizability is linked with framing an experiment as a paradigmatic case of a broader class of phenomena. We stress this point in order to highlight differences in perspectives on design research. In contrast to methodologies that aim at describing phenomena by identifying variables and the relations between them, we adhere a holistic approach, in which the aim is not try to establish the influence of individual variables but to develop a more encompassing explanation. Against this background, we would argue that the demand that design researchers work on ways to make testable claims generalizable in order to make design research more scientific - as for instance made by Collins, Joseph, Bielaczyc (2004) and Kelly (2004) - is at odds with our conception of design research.

Design and research

Although our emphasis in the above paragraphs has been on ways of justifying the results of design experiments, we do not want to lose sight of the fact that design and research is about researching and designing. We have discussed issues such as validity and trustworthiness at some length much of the current debate about design research has focused on justification.[5] However, the design aspect of the methodology is equally important. Design research presupposes that there is an adequately grounded basis for designing the instructional sequence/innovative learning ecology. The description 'learning ecology' introduced by Cobb et al. (2003) might be more adequate as it accentuates that we are dealing with a complex, interacting system involving multiple elements of different types and levels by designing these elements and by anticipating how these elements function together to support learning. Taking into account the complexity of a learning ecology, this implies the need for a very broad framework. The research of Doerr and Zangor (2000) serves to illustrate the complexity of a learning ecology. The authors found that productive use of graphic calculators requires coherence between the following elements of a learning ecology:

- the beliefs of the teacher
- the ability of the teacher to work with the graphic calculator
- the classroom culture (social norms en socio-math norms), and social practices
- the design of the instructional sequence
- the characteristics of the instructional tasks
- the manner in which the graphic calculator is construed as a tool
- and last but not least, the pedagogical-didactical skills of the teacher in making this whole system work.

In light of this list, it can be argued that the theoretical base for the design should incorporate general background theories such as socio-constructivism, or socio-cultural theory, domain-specific theory and theories on specific elements of the learning ecology, such as theories on tool use.
In addition to this the research team should be well informed about the state-of-the-art professional knowledge of the domain under consideration.

Acknowledgement

The analysis reported in this paper was supported by the National Science Foundation under grant No. REC 0231037. The opinions expressed do not necessarily reflect the view of the Foundation.

5) As seemed to be the case at the Symposium Design-Based Research: Grounding a New Methodology, at the AERA 2004.

Literature

Atkinson, P., Delamont, S., & Hammersley, M. (1988). Qualitative research traditions: A British response to Jacob. *Review of Educational Research, 58*, 231-250.

Ball, D. (1993). With an eye on the mathematical horizon: Dilemmas of teaching elementary school mathematics. *Elementary School Journal, 93*, 373-397.

Ball, D. L., Sleep, L., Boerst, T., & Bass, H. (2009). Combining the development of practice and the practice of development in teacher education. *Elementary School Journal, 109*(5), 458-474.

Bakker, A. (2004). *Design research in statistics education: On symbolizing and computer tools*. Utrecht: CDß-Press.

Borko, H., Jacobs, J., Eiteljorg, E., & Pittman, M. E. (2009). Video as a tool for fostering productive discussions in mathematics professional development. *Teaching and Teacher Education, 24*, 417-436.

Brown, A.L. (1992). Design experiments: Theoretical and methodological challenges in creating complex interventions in classroom settings. *The Journal of the Learning Sciences, 2*(2), 141-178.

Bruner, J. (1994). *Four ways to make meaning*. Invited address at the AERA Conference 1994, New Orleans.

Bruner, J. (1996). *The culture of education*. Cambridge, MA: Harvard University Press.

Cobb, P. (1999). Individual and collective mathematical development: The case of statistical data analysis. *Mathematical Thinking and Learning, 1*(1), 5-44.

Cobb, P. (2002). Reasoning with tools and inscriptions. *Journal of the Learning Sciences, 11*(2-3), 187-215.

Cobb, P., & Steffe, L.P. (1983). The constructivist researcher as teacher and model builder. *Journal for Research in Mathematics Education, 14*(2), 83-95.

Cobb, P., Yackel, E., & Wood, T. (1989). Young children's emotional acts while doing mathematical problem solving. In D. McLeod & V.M. Adams (Eds.), *Affect and mathematical problem solving: A new perspective* (pp. 117-148). New York: Springer-Verlag.

Cobb, P., & Whitenack, J. (1996). A method for conducting longitudinal analyses of classroom video recordings and transcripts. *Educational Studies in Mathematics, 30*, 213-228.

Cobb, P., & Yackel, E. (1996). Constructivist, emergent, and sociocultural perspectives in the context of developmental research. *Educational Psychologist, 31*, 175-190.

Cobb, P., Gravemeijer, K., Yackel, E., McClain, K., & Whitenack, J. (1997). Symbolizing and mathematizing: The emergence of chains of signification in one first-grade classroom. In D. Kirshner & J. A. Whitson (Eds.), *Situated cognition theory: Social, semiotic, and neurological perspectives* (pp. 151-233). Mahwah, NJ: Lawrence Erlbaum Associates.

Cobb, P., Stephan, M., McClain, K., & Gravemeijer, K. (2001). Participating in classroom mathematical practices. *Journal of the Learning Sciences, 10* (1&2), 113–164.

Cobb, P., McClain, K., & Gravemeijer, K.P.E. (2003). Learning about statistical covariation. *Cognition and Instruction, 21*(1), 1-78.

Cobb, P., & Hodge, L. L. (2003, April). *An interpretive scheme for analyzing the identities that students develop in mathematics classrooms*. Paper presented at the annual meeting of the American Educational Research Association, Chicago.

Cobb, P., Confrey, J., diSessa, A., Lehrer, R., & Schauble, L. (2003). Design experiments in education research. *Educational Researcher, 32*(1), 9–13.

Cobb, P., & Tzou, C. (2009). Supporting students' learning about data creation. In W-M. Roth (Ed.). *Mathematical representations at the interface of body and culture* (pp. 135-170). Charlotte, NC: Information Age Publishing.

Collins, A., Joseph, D., & Bielaczyc, K. (2004): Design research: Theoretical and methodological issues, *Journal of the Learning Sciences,13*(1), 15-42.

Doerr, H.M. & Zangor, R. (2000). 'Creating meaning for and with the graphing calculator', *Educational Studies in Mathematics: 41*, 143-163.

diSessa, A.A. (1992). Images of learning. In E. De Corte, M.C. Linn, H. Mandl, & L. Verschaffel (Eds.), *Computer-based learning environments and problem solving* (pp. 19–40). Berlin: Springer.

diSessa, A.A. (2002). Students' criteria for representational adequacy. In K. Gravemeijer, R. Lehrer, B. van Oers, & L.Verschaffel (Eds.), *Symbolizing, modeling and tool use in mathematics education* (pp.105–129). Dordrecht: Kluwer.

diSessa, A. A., & Cobb, P. (2004). Ontological innovation and the role of theory in design experiments. *The Journal of the Learning Sciences, 13*(1), 77-103.

Forman, E. A. (2003). A sociocultural approach to mathematics reform: Speaking, inscribing, and doing mathematics within communities of practice. In J. Kilpatrick, W. G. Martin & D. Schifter (Eds.), *A research companion to principles and standards for school mathematics*, (pp. 333-352). Reston, VA: National Council of Teachers of Mathematics.

Freudenthal, H. (1971). Geometry between the devil and the deep sea. *Educational Studies in Mathematics, 3*, 413-435.

Freudenthal, H. (1973). *Mathematics as an educational task*. Dordecht: Reidel.

Freudenthal, H. (1983). *Didactical phenomenology of mathematical structures*. Dordrecht: Reidel.

Freudenthal, H. (1991). *Revisiting mathematics education*. Dordecht: Kluwer Academic Publishers.

Freudenthal, H., Janssen, G.M., & Sweers, W.J. (1976). *Five years IOWO on H. Freudenthal's retirement from the directorship of IOWO: IOWO snapshots*. Dordrecht, Reidel.

Glaser, B. G., & Strauss, A.L. (1967). *The discovery of grounded theory: Strategies for qualitative research*. New York: Aldine.

Gravemeijer, K.P.E. (1993). The empty number line as an alternative means of representation for addition and subtraction. In J. de Lange Jzn, I. Huntley, C. Keitel, & M. Niss (Eds.), *Innovation in mathematics education by modelling and applications* (pp. 141-159).Chichester: Ellis Horwood.

Gravemeijer, K.P.E. (1994). *Developing realistic mathematics education*. Utrecht: Cdß Press.
Gravemeijer, K. (1998). Developmental research as a research method. In J. Kilpatrick & A. Sierpinska (Eds.), *What is research in mathematics education and what are its results? ICMI Study Publication; book 2* (pp. 277-295). Dordrecht: Kluwer Academic Publishers.
Gravemeijer, K. (1999). How emergent models may foster the constitution of formal mathematics. *Mathematical Thinking and Learning. 1*(2), 155-177.
Gravemeijer, K. (2002) Preamble: From Models to Modeling. In K. Gravemeijer, R. Lehrer, B. van Oers, & L. Verschaffel (Eds.), *Symbolizing, modeling and tool use in mathematics education* (pp. 7 – 22). Dordrecht, the Netherlands: Kluwer Academic Publishers.
Gravemeijer, K.P.E. (2005). Revisiting 'Mathematics education revisited'. In H. ter Heege, T. Goris, R. Keijzer, & L. Wesker (Eds.). *Freudenthal 100* (pp. 106-113). Utrecht: Freudenthal Instituut.
Gravemeijer, K., & Van Eerde, D. (2009). Design research as a means for building a knowledge base for teachers and teaching in mathematics education. *Elementary School Journal, 109*(5), 510-524.
Grossman, P., Compton, C., Igra, D., Ronfeldt, M., Shahan, E., & Williamson, P. W. (2009). Teaching practice: A cross-professional perspective. *Teachers College Record, 111*(9), 2055-2100.
Hiebert, J., & Stigler, J.W. (1999). *The teaching gap*. New York: Free press.
Kelly, A. (2004). Design research in education: Yes, but is it methodological? *Journal of the Learning Sciences, 13*(1), 115-128.
Kessels, J.P.A.M., & Korthagen, F.A.J. (1996). The relation between theory and practice: Back to the classics. *Educational Researcher, 25*, 17-22.
Lave, J., & Wenger, E. (1991). *Situated learning: Legitimate peripheral participation*. London: Cambridge University Press.
Lijnse, P.L. (1987). Onderzoek, ontwikkeling en ontwikkelingsonderzoek: In en na het PLON. *Panama-Post. Tijdschrift voor nascholing en onderzoek van het reken-wiskundeonderwijs, 6*(1), 35-43.
Meira, L. (1995). The microevolution of mathematical representations in children's activity. *Cognition and Instruction, 13*, 269-313.
Moore, D. S. (1997). New pedagogy and new content: The case of statistics. *International Statistical Review, 65*(2), 123-165.
Maxwell, J.A. (2004). Causal explanation, qualitative research, and scientific inquiry in education. *Educational Researcher, 33*(2), 3-11.
NCTM Research Advisory Committee (1996). Justification and reform. *Journal for Research in Mathematics Education 27*(5), 516-520.
Plomp, T. (2013). An introduction to educational design research. In T. Plomp & N. Nieveen (Eds.), *Educational design research – part A: An introduction* (pp. 10-51). Enschede, the Netherlands: SLO.

Roth, W-M., & McGinn, M. K. (1998). Inscriptions: Toward a theory of representing as social practice. *Review of Educational Research, 68*(1), 35-59.

Sfard, A. (1991). On the dual nature of mathematical conceptions: Reflections on processes and objects as different sides of the same coin. *Educational Studies in Mathematics, 22*, 1-36.

Shadish, W. R., Cook, T. D., & Campbell, D. T. (2002). *Experimental and quasi-experimental designs for generalized causal inference.* Boston: Houghton Mifflin.

Sherin, M. G., & Han, S. Y. (2004). Teacher learning in the context of video club. *Teaching and Teacher Education, 20*, 163-183.

Simon, M. A. (1995). Reconstructing mathematics pedagogy from a constructivist perspective. *Journal for Research in Mathematics Education, 26*, 114-145.

Smaling, A. (1990). Enige aspecten van kwalitatief onderzoek en het klinisch interview [Some aspects of qualitative research and the clinical interview]. *Tijdschrift voor nascholing en onderzoek van het reken-wiskundeonderwijs, 8*(3), 4-10.

Smaling, A. (1992). Varieties of methodological intersubjectivity: The relations with qualitative and quantitative research, and with objectivity. *Quality & Quantity, 26,* 169-180.

Streefland, L. (1990). *Fractions in realistic mathematics education: A paradigm of developmental research.* Dordrecht: Kluwer Academic Publishers.

Steffe, L.P. (1983). The teaching experiment methodology in a constructivist research program. In M. Zweng, T. Green, J. Kilpatrick, H. Pollak, & M. Suydam (Eds.), *Proceedings of the Fourth International Congress on Mathematical Education* (pp. 469-471). Boston: Birkhäuser.

Stephan, M., Bowers, J., Cobb, P., & Gravemeijer, K. (Eds.). (2003). Supporting students' development of measuring conceptions: Analyzing students' learning in social context. *Journal for Research in Mathematics Education Monograph No. 12.* Reston, VA: National Council of Teachers of Mathematics.

Stephan, M., Underwood-Gregg, D., & Yackel, E. (in press). Guided reinvention: What is it and how do teachers learn this teaching approach? In Y. Li (Ed.), *Transforming mathematics instruction: Multiple approaches and practices.*

Taylor, S.J., & Bogdan, R. (1984). *Introduction to qualitative research methods: The search for meanings.* (2nd ed.). New York: Wiley.

Thompson, P.W., & Thompson, A.G. (1994). Talking about rates conceptually, part I: A teacher's struggle. *Journal for Research in Mathematics Education, 25*, 279–303.

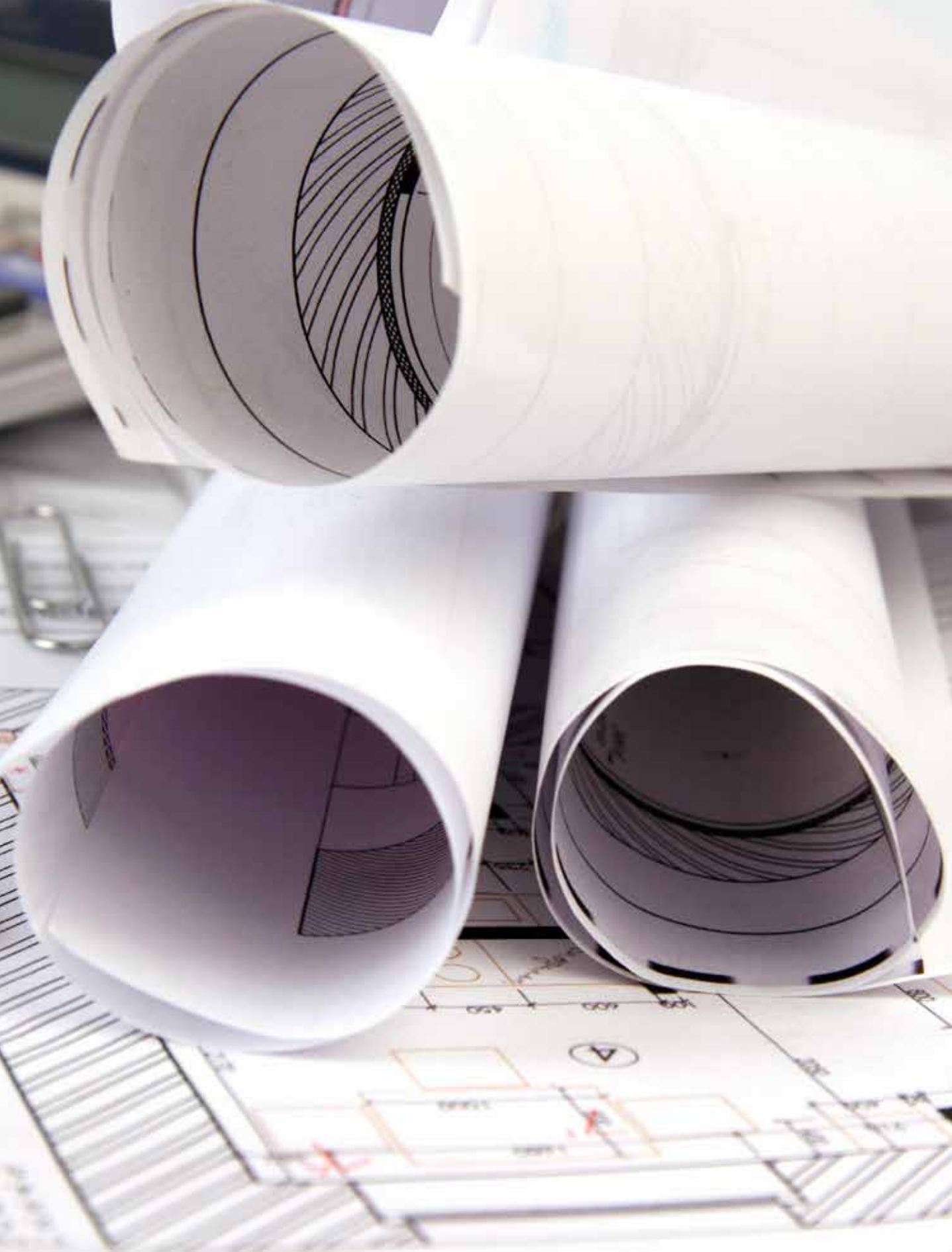
Thompson, A.G., & Thompson, P.W. (1996). Talking about rates conceptually, part II: Mathematical knowledge for teaching. *Journal for Research in Mathematics Education, 27*, 2–24.

Treffers, A. (1987). *Three dimensions: A model of goal and theory description in mathematics instruction—The Wiskobas Project*. Dordrecht, The Netherlands: Reidel.

Van den Akker, J. (1999). Principles and methods of development research. In J. van den Akker, R.M. Branch, K. Gustafson, N. Nieveen & T. Plomp (Eds.), *Design approaches and tools in education and training* (pp. 1-14). Boston: Kluwer Academic Publishers.

Yackel, E., & Cobb, P. (1996). Sociomathematical norms, argumentation, and autonomy in mathematics. *Journal for Research in Mathematics Education, 27*, 458-477.

Yackel, E., Gravemeijer, K., & Sfard, A. (Eds.). (2011). *A journey in mathematics education research: Insights from the work of Paul Cobb.* Dordrecht: Springer.



4. The Integrative Learning Design Framework: An Illustrated Example from the Domain of Instructional Technology

Brenda Bannan

Articulating a clear definition and process of design research is a prominent and progressive topic among educational researchers (Andersen & Shattuck, 2012; McKenney & Reeves, 2012; Kelly, Lesh, & Baek, 2008; van den Akker, Gravemeijer, McKenny, & Nieveen, 2006). Design research studies involve complex interactions and feedback cycles that can significantly blur the roles of researchers, teachers, curriculum developers, instructional designers and assessment experts (Kelly, Lesh, Baek, & Bannan-Ritland, 2008). As educational researchers struggle to clarify this research method, they continue to raise significant questions such as how is design research different from the process of design? What are appropriate methods and processes that can be used in design research? How do we systematically create, test and disseminate design or teaching interventions that will have maximum impact on practice capitalizing on design research? How do we generate both theoretical and practical knowledge related to complex educational settings?

Kelly (2006) and others (Fishman, 2007; Zaritsky & Rogers, 2003; Collins, 1999; Design-based Research Collective, 2003) advocate that these emerging methods call for the articulation of new processes and criteria including factors such as the usefulness and usability of knowledge, its shareability, and marketability, how well it disseminates and the extent to which it positively impacts practice. Researchers in this area cite a need for organizational structure and protocol for the diffusion of research into practice and states that educational research situations are extremely complex systems that can benefit from integrated system research strategies (Bannan, 2012; Kelly, 2004; Kelly, Lesh, & Baek, 2008). There is a need for comprehensive models to guide design research addressing the process of designing, developing and assessing the impact of an educational innovation. In this chapter, I present an integration of existing design and research processes offering a guiding framework that goes beyond the individual domains of social science, behavioral science and communication theory and attempts to integrate the systematic processes of the related fields of instructional design, software engineering, product design, hence the name Integrative Learning Design Framework (ILDF). Building on the integration of processes from multiple fields such as instructional design, object oriented software development, product development and diffusion of innovations and educational research, the ILDF present a "meta-methodological" view that attempts to integrate the best of design, research and diffusion of educational innovations. This framework consisting of four phases (see Figure 1) challenges

researchers to provide improved articulation of design research processes by phase and to consider the entire scope of research from initial conceptualization to diffusion and adoption.

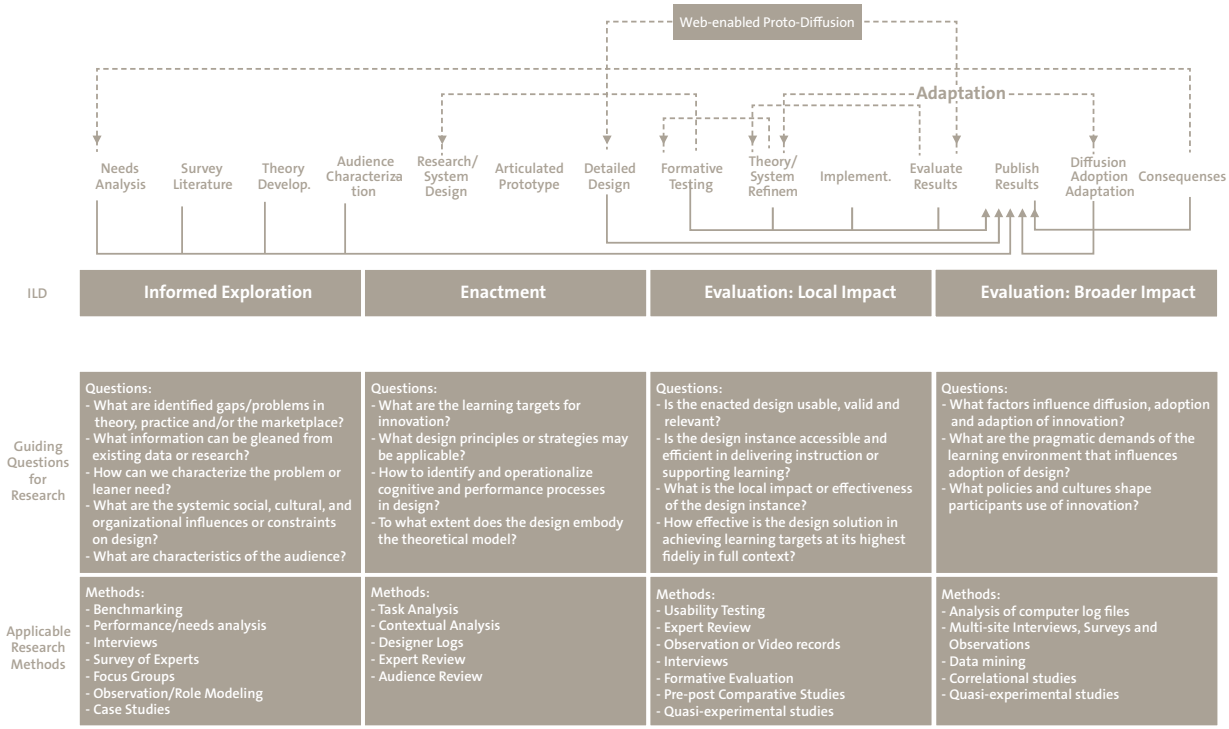


Figure 1: Questions and methods for design research by ILDF phase

The four phases of Informed Exploration, Enactment, Local Evaluation and Broad Evaluation presented in the ILDF encompass a process model for conducting design research based on several years of attempts to incorporate progressively more rigorous, research-based cycles within a technology-based instructional design effort. This type of effort is different than traditional instructional design as the iterative cycles are essentially micro-cycles of research (more comprehensive qualitative and quantitative research efforts than formative evaluation cycles) conducted to learn more than how to improve the technology system, although the studies may also result in that outcome. The reference to learning within the ILDF is to place emphasis on the learning that can result in the context and activity of design. For example, as researchers or instructional designers we may generate information about the teaching and learning process, participants, context, and culture that is often not attended to, discarded and captured in a rigorous manner for others to learn from

and capitalize on. Whether our design activity involves classroom-based interventions, technology or some combination of both, the interconnected design research cycles can generate knowledge about design principles but also provide rich information on aspects of learning, cognition, expert and novice perspectives, as well as stakeholder positions to direct design and design decision-making. The core issue at hand is that the rich, complex, design process may offer multiple opportunities to generate research-based knowledge however, much of it is lost and not documented in the creative design process. Our challenge, as design researchers, is to try to systematically gather, analyze, report and codify this information in a rigorous manner that strives toward some type of logical, argumentative grammar worthy of stringent research processes (Kelly, Lesh, & Baek, 2008; McKenney & Reeves, 2012).

Connected cycles of research cycles and design processes result in improved decision-making based on data-driven results for design, development and research purposes. Though clearly interventionist and primarily formative in nature, the ILDF process stands apart from traditional instructional design and research efforts. Throughout the multiple phases and cycles of integrated research and design processes valuable knowledge in the context of use is generated. We need to mine what is learned about important factors related to learning, context, culture, and technology within the design process (not separate from it in a controlled setting as evidenced in traditional research). If design researchers can articulate an integrative research and design process, it may have the potential to significantly improve our understanding of teaching, learning and training in-situ. The multiple macro and micro-cycles of data collection, analysis and most importantly, results-driven design decision-making is what sets design research apart from traditional formative evaluation in instructional design which is often conducted in a very limited manner or a single cycle of data-gathering and analysis.

Tessmer (1993) refers to formative evaluation as a "judgment of the strengths and weakenesses of instruction in its developing stages, for purposes of revising the instruction to improve its effectiveness and appeal" (p. 11). Although multiple methods may be used including expert review, one-to-one evaluation, small group and field testing, formative evaluation cycles in traditional instructional systems design may not always employ research methods that are specific to particular phases of an integrative and connected design research cycle. Formative evaluation, despite its most rigorous and comprehensive application does not progressively generate knowledge about cognition, context and culture of use but provides a limited focus on a particular technology system of instruction and judges its effectiveness, appeal and efficiency. In contrast, design research cycles are based on a thorough, systematic process integrated multiple design and research processes to progressively improve understanding about learners, learning, context, or culture as

well as iteratively improve an intervention. Therefore, formative evaluation methods are subsumed as one selected method in what could be described as a "meta-methodological" or involving multiple research methods across the design research process. What is critical in design research is the theoretical yield of the effort to be viewed as important as the improvement of the intervention (Plomp, 2013; Nieveen & Folmer, 2013).

Design research cycles are dynamic and integrate multiple exploratory, constructive and/or empirical research methods as well as multiple design/development techniques (see Figure 1). Exploratory research methods structure and identify new problems such as feasibility testing, benchmarking and qualitative research approaches. Constructive research develops solutions to problems and may include testing of a construct or theory against a predefined criteria and may, for example, include formative evaluation testing of an instructional technology system. In contrast, empirical research tests the feasibility of a solution using empirical or direct or indirect observation or evidence in the tradition of the scientific method. Design research may employ all three forms of research methods as well as incorporate formative evaluation methods at different phases in the process. However, traditional, formative evaluation perspectives while offering valuable iterative processes, do not in isolation, address the complexity inherent in educational practice. Most educational research projects advocate only one cycle of qualitative or quantitative empirical testing at a fixed point in time for a given instructional intervention for the sole purpose of generating knowledge. In contrast, design research attempts to progressively and dynamically generate (exploratory research), improve (constructive research) and learn about (empirical research) a particular phenomenon from interconnected research and design cycles.

In response to this challenge, the ILDF model attempts to provide a comprehensive yet dynamic and flexible guiding framework that positions multiple, micro and macro design research cycles as primarily socially-constructed, contextualized process of producing the most educationally effective product that has the best chance to be used in the classroom while also generating knowledge about teaching and learning within the activity of design. The model or framework attempts to move past isolated, individual efforts of educational research by clearly articulating a logically-ordered structural frame that considers the full spectrum of research methodology in advancing toward systemic impact in education and may be applied in a variety of contexts. In the early days of design research, Collins (1990; 1993) advocated for a similar overt, systematic methodology for conducting design experiments and states: "When designing a learning environment, whether computer based or not, there are a multitude of design decisions that must be made. Many of these design decisions are made unconsciously without any articulated view of the issues being addressed or the tradeoffs involved. It would be better if these design decisions were consciously considered, rather than unconsciously made" (1993, p.1).

The ILDF process presents one step toward a systematic framework for the articulation and documentation of common phases and complementary stages based on multiple design and research processes promoting more conscious design research (Bannan, Peters, & Martinez, 2010; Bannan, 2012; Collins, 1990; 1999). Although there are thousands of decisions made in a design research context, the major conjectures, learning targets, task analysis, design principles and evaluation or research decision-making resulting from exploratory, constructive and/or empirical research cycles may be uncovered by examining a rich case study as presented here entitled the LiteracyAccess Online (LAO) project. The ILDF is presented here as a starting point for researchers to consider as with the goal of eliciting questions, suggestions, limitations and criteria that may need to be considered as researchers struggle with the implications of this emerging form of educational research. In this chapter, I briefly describe the progression of the LAO design-based research study that encompassed four years of effort and illustrates the application of the ILDF. The LAO case study example is described according to broad phases including 1) the informed exploration phase; 2) the enactment phase; 3) the local impact phase; and 4) the broad impact phase as well as the multiple, potential applied and empirical research processes that align with each phase (see Figure 1). It is hoped that the LAO example will provide enough detail to potentially improve understanding of conducting cycles of design research related to a technology-based educational intervention.

LiteracyAccess Online – an integrative learning design study

The LiteracyAccess Online (LAO) project1 provides an example of an integrative learning design study based on the ILDF specifically illustrating the intersection and systematic expression of multiple design and research methods. LiteracyAccess Online is an effort to utilize Web-based technology to provide support for teachers, tutors, and parents (literacy facilitators) in addressing literacy goals for all children with a particular focus on those with disabilities. After four years of design research and development, LAO (http://literacyaccessonline.org/) now provides a technology-based learning environment that promotes the use of specific literacy strategies for the improvement of tutoring and reading performance as the child and literacy facilitator collaboratively engage in the process of reading online.

1) The LiteracyAccess Online (LAO) project was supported by the Office of Special Education Programs in the Department of Education Steppingstones of Technology Innovation for Students with Disabilities Grant CFDA84.327A

The informed exploration phase

The exploratory research objectives of the LAO integrative learning design study were two-fold; 1) to investigate the nature and effectiveness of a consistent technology-based, collaborative literacy environment as well as; 2) to generate knowledge about how literacy facilitators and children understand and employ reading support strategies. These objectives were originally conceived as research/evaluation questions (see Figure 2) and evolved from an extended, progressive investigation into the provision of literacy support for facilitators and children. This 'meta-methodological' design research process consisting of multiple research methods (e.g. survey, focus groups, interviews, expert reviews, etc) resulted in clearly articulated learning targets, task analyses of learning objectives, theoretical model embedded in a technology system design and congruent research/evaluation questions that drove more rigorous qualitative testing of the intervention whose results contributed to theory of literacy support for children with disabilities further elaborated in sections below (see Figure 2).

To begin exploratory research cycles, initial explorations into target audience and stakeholder perceptions, related products and literature and documentation of the complex nature of supporting literacy revealed many plausible paths for design research. The interdisciplinary research team involved in the LAO project were charged with determining the research direction and consisted of educational researchers, teachers, graduate students, content experts in literacy, special education and assistive technology as well as parents involved in an advocacy group for children with disabilities. The broad design research focus evolved from the team's perceived lack of support for children who were struggling with the literacy process, based on direct observations of this problem in both classroom and home environments which then manifested itself into several individual but connected research studies. For example, we conducted multiple interviews with parents with children with disabilities who were struggling with the reading process. We also invited several parents to participate on our design research team. The teams' analysis of the interview data as well as the design research team discussions revealed our initial approach. We had initially decided to design a tutorial-based intervention only for the child's benefit, however, a comment in a team meeting dramatically changed our design direction. In line with Collins' notion of conscious considerations of design as demonstrating core underlying design decision-making, one parent member/stakeholder on the team stated that she primarily read in conjunction with (not to) her son and wanted to do so online but with additional support of higher level reading strategies. Based on that input and follow-up micro-cycles of interview data collection and analysis to confirmation the viability of this design approach with parents, we conducted a series of interviews and surveys to determine the feasibility of this design direction. The results of our investigation and discussions evolved into an online collaborative performance support

system to support both literacy facilitators and their children in the literacy process as the determined design direction. This was a design decision based on data analyzed from multiple interviews and survey procedures. Aligned with Confrey and Lachance's (2000) notion of drawing key inferences from dissatisfaction with current educational practices and direct experiences with children, initial theoretical conjectures were developed based on the analyzed data that advocated for reading, writing and assistive technology support for children with or without disabilities to increase their engagement and performance in literacy.

While these initial theoretical conjectures provided a central premise and broad direction for design research, more information was needed to refine these conjectures resulting in a comprehensive needs analysis and literature review that provided a firm and complementary theoretical foundation for the intended design. Extensive exploration into appropriate literacy strategies, tutorial programs and processes, surveys of experts, teachers and parents as well as qualitative observation of children and facilitators engaged in a literacy experience all informed this phase of the research. This provided not only well-defined design directions but also added to the research literature regarding children with disabilities and their parents understanding of assistive technology and literacy learning (see Jeffs, Behrman, & Bannan-Ritland, 2006). Many potential design research directions were considered based on the initial conjectures, however, data drawn from conducted interviews, direct experience with potential research participants and literature review converged and pointed the team in a particular direction.

A prominent theme that emerged across initial interviews, surveys and observations with experts, parents, teachers and children revealed that literacy facilitators had a crucial role in providing support for children struggling to gain literacy skills and the question remained how to best support this role. These findings and related literature provided insight for informed theory and improved conjectures based on the aforementioned qualitative interviews and literature reviews. Results from data collection and literature review methods in the informed exploration phase indicated that 1) children can, but often do not use effective metacognitive reading strategies; 2) explicitly teaching these strategies can greatly enhance children's comprehension of text; 3) teachers (as well as other literacy facilitators) need to be trained in how to provide cognitive structure for their students so that children can learn to guide their own generative processes in reading; and 4) one-to-one tutoring is one of the most effective forms of instruction for improving reading achievement but increased success often depends upon the skill of the tutor or facilitator and the establishment of consistent roles and expectations (Wittrock, 1998; NRP, 2000; Wasik, 1998).

This exploration into the literature and perspectives of those involved in these issues greatly refined our initial theoretical conjectures and resulted in a dramatic change of our intended design direction for this research from a didactic, tutorial, child-focused intervention to a collaborative, story-based reading experience providing embedded metacognitive strategy support for both the literacy facilitator and the child's use. The rationale for this research direction was documented in a comprehensive needs analysis that detailed the data collection, conclusions and related literature review.

The next stage of our design research involved the analysis and description of the range of learners and facilitators that would potentially use the LAO system. Direct experience with 4^{th}-8^{th} grade children with or without disabilities, teachers, tutors, and parents provided data that characterized our audience. These descriptions were depicted as role models (Constantine & Lockwood, 1999) or personas (Cooper, 1999) that comprised abstract composite profiles of audience characteristics gleaned from actual interviews and observations and provided a focal point for design. Role models or personas are similar to Graue and Walsh's (1998) qualitative vignettes that strive to capture the substance of a setting, person or event to communicate a central theme of qualitative data, based on multiple direct observations and are employed here as also a focal point for design.

Exploring the nature of the identified educational problem, related products and literature as well as creating and refining theoretical conjectures and descriptions of the audience provided an informed perspective for grounded design of a learning environment based on articulated theory. These activities resulted in specific research artifacts including a needs analysis that contained an extensive literature review, an articulated and congruent design and research direction and detailed audience analysis based on qualitative and quantitative data. These documents were housed on a project website that provided a communication mechanism between team members as well as an archive of shareable design research processes, products and evidentiary data.

The enactment phase

The embodiment of the results of our informed exploration and theories about providing literacy support for children and literacy facilitators in a usable learning environment were collaboratively constructed across several stages and constructive research cycles that develop solutions to problems culminating in a web-based prototype. The initial design of the LAO learning environment resulted directly from the design implications articulated in the previous phase of exploratory research, analyses and review. These implications were translated into an articulated prototype initially developed by building an abstract, paper-based model of the system for researcher and teacher input according to procedures adapted from usage-centered design processes previously mentioned as role models

(Constantine & Lockwood, 1999). Role models are a technique to characterize primary and secondary target audiences for the purposes of design. For example, we created role models and personas for children with learning disabilities (such as attention deficit disorder) based on our direct experience with a child who was struggling in the reading process and his mother who did not have any knowledge of advanced reading support strategies. These techniques are based on real-world experiences with representatives of the target audiences your intervention or system is being designed for but evolve into a archetypal composite of the attributes of many individuals. Therefore, role models and persona's become a qualitative profile to continually target design efforts to maintain the audience(s) or user(s) perspectives.

Abstract or low-fidelity modeling/prototyping of the instantiated or enacted design provided opportunities for input and co-construction of LAO with several audience members prior to the more time-intensive computer-based production of the learning environment. We utilized Constantine and Lockwood's (see foruse.com) procedures of usage-centered design that encompassed low-fidelity representation and organization of all the features of the database-driven website. For LAO, we deliberately ultimately designed a web database system that would permit performance support for the parent-child dyad in providing meta-cognitive prompts for both participants based on research-based reading strategies throughout a collaborative and generative process of engaging with text.

In the context of a constructive research approach that attempts to validate a particular construct (e.g. theory, model, software or framework) against identified criteria or benchmarks, the team conducted several iterative cycles of data-gathering and analysis of expert reviews and target audience reviews. These progressive, micro-cycles of data collection and analysis resulted in data-driven cyclical revisions of the articulated prototype which were reflected in detailed design documentation including the production of flowcharts, technical specifications and storyboards. The design research process of employing micro-cycles of constructive research data gathering and analysis elicited feedback at each cycle and design revisions agreed upon by the team which resulted in the initial creation and then progressive improvement of a web-based prototype validated by data collected in a constructive research approach. As a team, we constructed specific criteria related to usability of system and observations and video analysis of actual use of the system by literacy facilitators and children. The specific methods of data collection employed at this stage included designer logs posted on the project website, expert panel reviews of the design and documented reviews of the design by content experts, audience members and the research team.

The Local Impact Phase

Once a physical web-based prototype was in place, the incorporation of formative evaluation and qualitative methods in an empirical manner could commence and began to characterize the rich, highly iterative nature of the local impact phase as it progressively informed, revised and refined our theoretical constructs as well as the web-based instructional design approach and redesign efforts. The complex interactions between facilitators and children that can occur in multiple settings formed the series of micro-cycles in LAO examining these specific constructs that grounded related research questions: 1) parent-child dyads in an informal setting with extensive involvement by researchers; 2) parent-child dyads in a structured workshop experience supported by researchers and; 3) pre-service-teacher dyads in a field trial progressing toward more closely modeling authentic conditions experienced with the prototype. When a fully functioning prototype was not yet available, studies attempted to closely mimic the tasks that would be embedded in LAO. The data gathering across these three studies incorporated observations, interviews, child and parent journal entries, videotaped use of system and pre- and post-online surveys (see Jeffs, 2006). This multi-tiered, multi-method evaluation scheme generated useful knowledge and subsequent results from each stage of inquiry were then cycled into changes of our theoretical conjectures, research design as well as system design. This process revealed insights into the core design principles (McKenney & Reeves, 2012; van den Akker, et al., 2006) that may support the collaborative learning and implementation of metacognitive processes by literacy facilitators and children in a technology-based environment. Our core design principles that evolved and were refined included the following that when engaged in a collaborative literacy process that provides high level reading strategy metacognitive support in a Web-based context: 1) parent literacy facilitators could develop greater awareness and skill in implementing reading activities and identify supports for their child in a structured setting; 2) children showed improvement in literacy skills using technology-based support when participating in a guided workshop environment; and 3) pre-service teachers felt that the strategies and activities embedded in the LAO environment facilitated children's comprehension, motivation and interest when working with them in this environment (see Jeffs, et. al. 2006). More rigorous evaluations are planned for the future to systematically increasing number of participants and varying contexts for the use of LAO in school, home and tutoring environments. These studies involve detailed tracking of computer-based activities of the dyads in school and home settings, assessment of facilitator and child use of metacognitive strategies prior to using LAO and pre- and post comprehension measures after several weeks of using the system.

In addition, a series of expert reviews, usability testing, one-to-to-one, small group, and field testing were implemented in progressively more authentic settings. The more intensive studies focused primarily on qualitative studies that characterized the target audience interaction with the enacted theoretical model in relation to the learning targets. Specifically, the team was interested in how facilitators and learners perceived and interacted with Web-based support in the collaborative literacy process (which included both reading and writing tasks). A pilot study was initially conducted that simulated some tasks within LAO and provided feedback on the emerging site with five dyads of mostly parent facilitators and one sibling facilitator. Methods included collecting data through semi-structured interviews and observations of parent-child interaction with the prototype and complementary assistive technologies (e.g. text-to-speech, etc.) that promoted in reading and writing activities.

The preliminary study revealed that the children were motivated to complete reading and writing activities on the Web and that facilitators developed awareness for implementing reading activities in a collaborative process but desired additional support for children's disabilities. While the Web-based activities and supports for the reading process were useful for providing more authentic and self-initiated reading and writing activities, the research also revealed that interaction between parent and child dyads during these activities often created tensions that were not present when children were working with non-family members. Revisions to the theoretical model and enacted design of LAO based on this cycle of evaluation included among others, behavioral prompts directed toward the parent-child dyad to potentially release tension (such as prompts to take a break, positive reinforcement techniques, etc.) when engaged in collaborative reading and writing tasks and additional reading strategy supports and activities.

To further investigate the enacted theoretical model, a follow-up small group qualitative study was conducted with eight parent/child dyads that represented a variety of skill levels and disabilities (Jeffs, 2000). The specific goals of this cycle of research was to identify the characteristics of parent/child dyads working together specifically in literacy skill development, depict the interactions of the dyad and investigate the impact of various forms of technology (Internet, EPSS and any assistive technology) on attitudes of the participants. Participants included parents and children with various disabilities in grades 4^{th} through 6^{th} who were reading at least two grades below grade level and had a tendency to avoid reading and writing tasks prior to participation in the study. The study revealed that parents recognized the importance of immediate feedback and assistive technology features in the provided tools. Other results revealed that with the support of their parents, children can select appropriate technologies and with integrated use of the Internet and assistive technologies, children's writing samples improved in both quantity and quality.

Suggested revisions for the LAO prototype based on these results included built-in assistive technology features (instead of merely references to outside resources) such as text-to-speech capabilities and reading selections reflecting varying abilities and areas of interest – features that were subsequently incorporated into the LAO design.

In each of these cycles of problem-state, data collection, analysis and subsequent design move or formative evaluation process, the theoretical model enacted within the LAO prototype expanded to incorporate new and revised elements based on targeted data collection and research results. At this point, traditional research and design processes somewhat diverge in that the analyzed results are not an end in and of themselves, but are used for data-driven decision making or problem solving to build upon or revise theoretical assumptions and improve design. Often, based on testing results, we would need to throw out previous prototype features and totally redesign, revise or add new features. The team's informed design judgment and collaborative social negotiation was key to this decision-making.

The local impact phase is a time-intensive phase with multiple cycles that strives to yield a usable and internally valid intervention. Testing the intervention in progressively more realistic settings provides valuable information to inform theoretical assumptions related to the design but also to begin to isolate variables that might be further empirically tested. In the LAO research conducted to date, the integration of reading strategy scaffolds and assistive technology supports in the collaborative literacy process between facilitators and children with a range of disabilities was identified as one factor, of many, that seem to hold promise for improving literacy skills. Conducting additional research to further investigate the collaborative process promoted by the technological environment as well as isolating the effects of the multiple reading supports and assistive technologies afforded by the prototype remains an important objective in this research.

Although the funding cycle for LAO has ceased, in order to progress from local effects to more externally generalizable effects, additional cycles of testing are needed to isolate and test particular variables using multiple sites, diverse participants and settings progressively limiting the researcher-participant interaction. Based on available funding, field tests or trials are planned for LAO to collect significant amounts of quantitative and qualitative data from several sites and over 50 participant dyads using selected measurements, online surveys and interviews including parents and children in home school environments, pre-service teachers and in-service teachers that could represent other literacy facilitators in several geographical locations interacting with children with a range of disabilities. This data would provide additional evidence for the effectiveness of enacted theoretical assumptions for the collaborative reading and literacy process as well as provide evidence for the effectiveness of the prototype at its highest fidelity in full context of the intended use.

The broad evaluation phase

The last phase of this design-based research effort involves disseminating LAO into the broad educational system. Although the LAO research has not yet fully progressed through this stage, initial explorations in this area have yielded some unique insights into the dissemination process. However, the reader should note that the dissemination process can encompass an entire research effort in itself. For example, Fishman (2006) has applied a design research framework related to the sustainability of technology-based curriculum interventions within an entire school district or system.

LAO, as a web-based learning environment, affords the opportunity to publish current working prototypes online for open use and input that has resulted in an early and unique diffusion and adoption process begun prior to the completion of a fully functioning system. While still in development, we have tracked over 100 potential adopters that have discovered and explored the LAO site. The profiling and data-base capabilities of the site permit tracking and analysis of this information that has provided detailed information on potential adopters of the system providing significant insight and impact on sources for our later diffusion efforts. We plan to incorporate more sophisticated computer-based data collection and analysis techniques such as datamining (Tsantis & Castellani, 2001) that may yield even more insights into early adopters' behaviors, profiles and use of this new tool. We have just begun to publish our results of the design based research conducted related to LAO in traditional academic journals and non-traditional Web publishing that provide avenues for additional forms of review and evaluation. The results of our initial studies have prompted new research directions such as exploring the interaction of an online community for parents of children with disabilities incorporated in the LAO environment. Given the iterative nature of this type of research, it is highly likely that determining the consequences of the LAO design research effort will yield new theoretical and applied questions that will prompt the entire process once again.

Theoretical yield of LiteracyAccess Online design research study

Given the design research process based on the Integrative Design Learning Framework described above, what did we learn? The characteristics of an intervention or as van den Akker, et. al. (2006) describe the "design principles" are an important yield of design research. In the LAO project, these design principles included providing metacognitive reading strategy support while a parent, teacher or tutor is engaged in the collaborative reading process with the child delivered through a comprehensive web-based performance support system.

Design research is often employed to begin to generate theory (McKenney & Reeves, 2012; Design-based Research Collective, 2003). With LAO, there were no literature sources, theoretical principles or research studies directly applicable to a Web-supported collaborative reading process so the team integrated insights from tutoring, reading strategies and real-time performance support. Zaritsky et al. (2003) speak to moving beyond traditional development and analysis to support and 'solidify' designs based on the appropriate empirical investigation of context with the appropriate methods at the appropriate phase with given resources. The LAO design research team went beyond traditional development with intensive cycles of interviews, surveys, observational studies as well as deep investigation of the one-on-one tutoring and reading strategies literature to build a new theoretical model of real-time metacognitive reading strategy and assistive technology support for both the literacy facilitator and the child with disabilities. Much of these insights were an integration of data analyses, direct experience with target audience members and a grounded literature in reading processes, tutoring and collaborative performance support.

The design research process was conducted systematically to:

1) uncover the initial conjectures about how learning might occur in this type of setting;
2) stated learning targets, task analyses (in this case based on Activity Theory),
3) the designed intervention which embodies the core design principles (metacognitive reading strategy support in a collaborative performance support context)
4) local impact or evaluation questions that drove the more intensive research cycles (see Figure 2).

This progression demonstrates an alignment or congruency from initial conjectures through local impact or evaluation questions that evolved during the design research study. The specific theoretical insights that were tested and revealed based on this process are included in Table 1. The multiple phases of the IDLF process uncovered many informal and formal theoretical insights based on macro and micro data collection and analysis cycles conducted within the process of design that can be typically overlooked in the traditional instructional design process. For example, extending beyond a traditional learner analyses, we conducted multiple cycles of surveys, interviews, and observations of target audience member interaction that revealed theoretical insights that go beyond just the design of the intervention. Our studies revealed that parents have little formal knowledge and use of good reading strategies when engaged with their child in the reading process. This insight parlayed into the design principles of LAO but also stand apart from it as a finding that may contribute to the literature in the reading field. By formalizing and extending the methods of traditional instructional design to promote rich cycles of data collection that then can inform our knowledge of particular audiences, learning contexts and processes –

separate but connected to the design of a particular intervention, we can begin to progress toward generating knowledge and useful theoretical insights that are typically overlooked in design. This becomes an information-loss process of learning about learners, contexts, and processes within the act of design that design research can recapture, which refers to the notion that in the context of both isolated design and research efforts, we do not take advantage of formalizing much of our learning in an exploratory, confirmatory or empirical manner. (Bannan-Ritland & Baek, 2008).

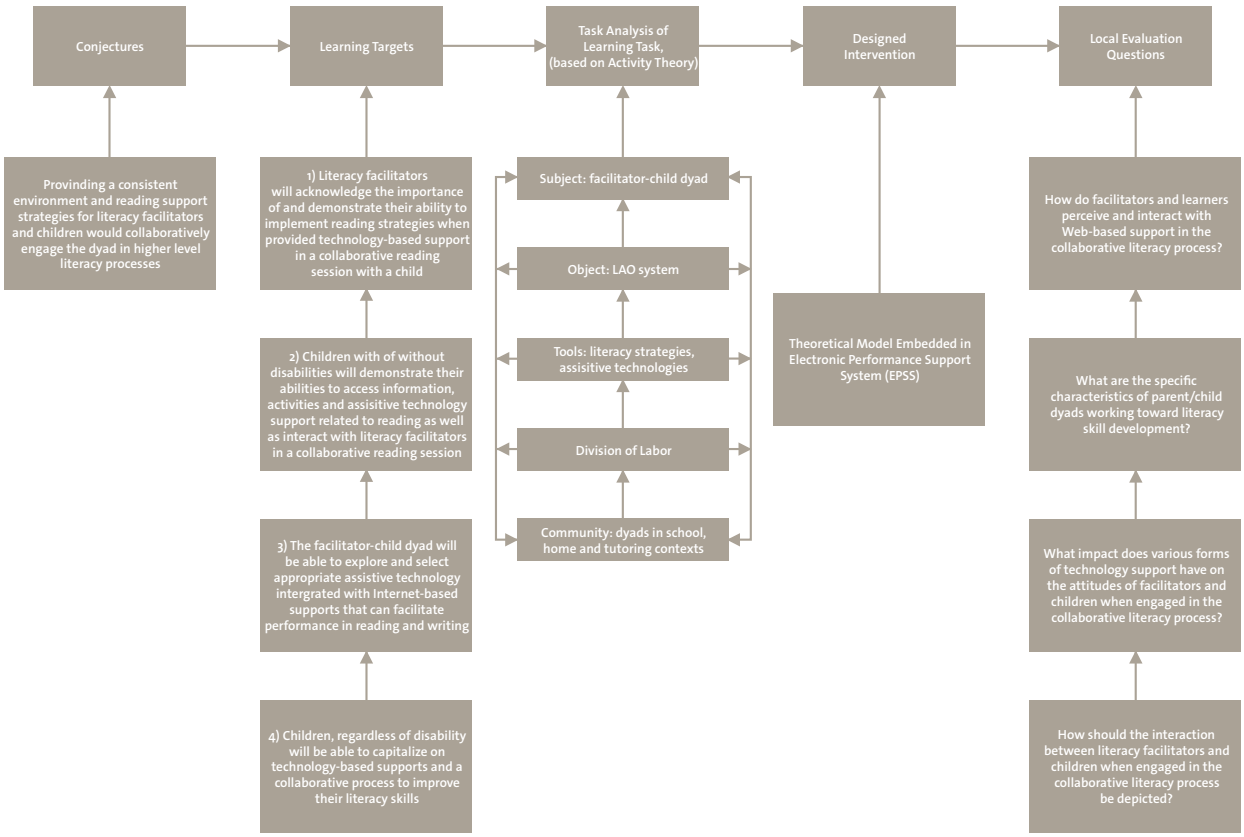


Figure 2: Initial progressive formulation from conjectures to local evaluation questions in LAO

Limitations will certainly also exist for the ILDF framework, as the knowledge generated is only as good as the rigor of the research methods employed. Integrating basic qualitative and quantitative research cycles to inform design at particular points and generate both design principles but also knowledge about learners, learning and learning contexts is the ultimate goal. Limitations may exist in time, quality of information uncovered

in data cycles that may impact design, small N to provide mostly qualitative insights initially and the failure inherent in the generation of theory in the discovery research process. However, it is through application in different design research contexts that more formalized processes will begin to be unveiled. The IDLF and LAO example are one case of a few currently for design researchers to uncover the logic and warrants of this new form of research (Kelly, Lesh, & Baek, 2008). There are many challenges that remain but capitalizing on the design process to generate research-based data-driven insights is a worthy goal, indeed.

Conclusion

This paper has presented a brief example and introduction to the ILDF framework that comprises a meta-methodological view of the design research process in an attempt to elucidate common phases and stages in this specific research methodology. The framework is presented to begin to establish common terminology and processes that can promote conscious design research. Most importantly, the ILDF framework is an attempt to provide a roadmap for future design researchers to investigate, articulate, document and inform educational practice.

* I am greatly indebted to Dr. Anthony E. Kelly whose insights and feedback on this chapter were invaluable in extending my thinking in this area. My appreciation also goes to Dr. Tjeerd Plomp, Dr. Nienke Nieveen and Dr. Jan van den Akker, esteemed colleagues and reviewers of this manuscript for their suggestions for revision.

References

Anderson, T., & Shattuck, J. (2012). Design-based research: A decade of progress in education research? *Educational Researcher, 41,*16-25.

Bannan-Ritland, B. & Baek, J. (2008). Investigating the act of design in design research: The road taken. In A.E. Kelly, R.A. Lesh, & J. Baek (Eds.), *Handbook of design research methods in education: Innovations in science, technology, mathematics and engineering* (pp. 299-319). Mahway, NJ: Taylor & Francis.

Bannan-Ritland, B. (2003). The role of design in research: The integrative learning design framework. *Educational Researcher, 32*(1), 21-24.

Bannan-Ritland, B. (2002) LiteracyAccess Online: The development of an online support environment for literacy facilitators working with children with disabilities. *Tech Trends, 45*(2), 17-22.

Bannan, B., Peters, E., & Martinez, P. (2010). Mobile, inquiry-based learning and geological observations: An exploratory study. *International Journal of Mobile and Blended Learning, 2*(3), 13-29.

Bannan, B. (2012). Design research and twice exceptional children: Toward an integration of motivation, emotion and cognition factors for a technology-based intervention. In D.Y. Dai (Ed.), *Design research on learning and thinking in education settings: Enhancing intellectual growth and functioning* (pp. 101-128). Mahwah, N.J.; Lawrence Erlbaum Associates.

Bannan-Ritland, B., Egerton, E, Page, J., & Behrmann, M. (2000) Literacy explorer: A support tool for novice reading facilitators. *Performance Improvement Journal, 39*(6) 47-54.

Cooper, A. (1999). *The inmates are running the asylum: Why high-tech products drive us crazy and how to restore the sanity*. Indianapolis, IN: SAMS.

Collins, A. (1990). *Toward a design science of education. Technical Report Issue No. 1 Center for Technology in Education*. New York, NY: Resources in Education (ERIC) ED 326179.

Collins, A. (1993). *Design issues for learning environments. Technical Report No. 27 Center for Technology in Education.* New York, NY: Resources in Education (ERIC) ED ED357733.

Collins, A. (1999). The changing infrastructure of education research. In E.C. Lagemann & L.S. Shulman (Eds.), *Issues in education research: Problems and possibilities* (pp.289-298). San Francisco: Jossey-Bass.

Confrey, J., & Lachance, A. (2000). Transformative teaching experiments through conjecture-driven research design. In A.E. Kelly & R.A Lesh (Eds.), *Handbook of research design in mathematics and science education* (pp.231-265). Mahway, NJ: Lawrence Erlbaum Associates.

Constantine, L.L., & Lockwood, L.A. (1999). *Software for use: A practical guide to the models and methods of usage-centered design*. Boston: Addison-Wesley.

The Design-Based Research Collective (2003). Design-based research: An emerging paradigm for educational inquiry. *Educational Researcher, 32*(1), 21-24.

Dick, W., & Carey, L. (1990). *The systematic design of instruction*. Glenview, IL: Scott, Foresman/Little.

Graue, M.E., & Walsh, D. (1998). *Studying children in context*. Thousand Oaks, CA: Sage

Isaac, S., & Michael, W.B. (1990). *Handbook in research and evaluation: A collection of principles, methods and strategies useful in the planning, design and evaluation of studies in education and the behavioral sciences*. San Diego, CA: EdITs publishers.

Jeffs, T., Behrmann, M., & Bannan-Ritland, B. (2006). Asssistive technology and literacy learning: Reflections of parents and children. *Journal of Special Education Technology 21*(1), 37-44.

Jeffs, T., Behrmann, M., & Bannan-Ritland, B. (2003). Asssistive technology and literacy difficulties: What parents and children have to say. *Journal of Special Education Technology, 19*.

Kelly, A.E., & Lesh, R.A. (Eds.). (2000). *Handbook of research design in mathematics and science education*. Mahway, NJ: Lawrence Erlbaum Associates.

LiteracyAccess Online Phase Four Project Web Site (2000). Retrieved August 30, 2002, from George Mason University, Immersion Program Web Site: http://chd.gse.gmu.edu/immersion/lao/deliver/needs.htm

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

Middleton, J., Gorard, S., Taylor, C., & Bannan-Ritland, B. (2008). The "compleat" design experiment: From soup to nuts. In A.E. Kelly, R.A. Lesh, & J. Baek (Eds.), *Handbook of design research methods in education: Innovations in science, technology, mathematics and engineering*. Mahway, NJ: Taylor & Francis.

Nieveen, N., & Folmer, E. (2013). Formative evaluation in educational design research. In T. Plomp & N. Nieveen (Eds.), *Educational design research – part A: An introduction* (pp. 152-169). Enschede, the Netherlands: SLO.

Plomp, T. (2013). Educational design research: An introduction. In T. Plomp & N. Nieveen (Eds.), *Educational design research – part A: An introduction* (pp. 10-51). Enschede, the Netherlands: SLO.

Rogers, E.M. (1995). *Diffusion of innovations*. New York: The Free Press.

Sabelli, N. (personal communication, May 15, 2001).

Tsantis, L., & Castellani, J. (2001). Enhancing learning environments through solution based data mining tools: Forecasting for self-perpetuating systemic reform. *Journal of Special Education Technology, 16*(4), 39-52.

Ulrich, K.T., & Eppinger, S.D. (2000). *Product design and development*. New York: McGraw-Hill.

Wasik, B. A. (1998). Using volunteers as reading tutors: Guidelines for successful practices. *Reading Teacher, 51*(7), 562-70.

Wittrock, M.C. (1998). Cognition and subject matter learning. In N.M. Lambert & B.L. McCombs (Eds.), *How students learn* (pp.143-152). Washington, DC: American Psychological Association.

5. When is Design Research Appropriate?

Anthony E. Kelly

Introduction

Design research has been described in detail in many publications, most recently by the Dutch (van den Akker, Gravemeijer, McKenney & Nieveen, 2006, with e.g., Kelly, 2006), and the Kelly, Lesh and Baek (2008) collection of papers (e.g., Kelly, Lesh, Baek, & Bannan-Ritland, 2008; Middleton, Gorard, Taylor, & Bannon-Ritland, 2008). Plomp (chapter 1 of this book) also provides an overview. For that reason, I will not reiterate the description here. Rather, I will assume that the reader is familiar with these sources and the special issues of journals (e.g., Barab & Squire, 2004; Kelly, 2003, 2004) that have appeared.

Instead, I wish to place design research within the frame of a larger context for research on interventions. In her seminal piece, Bannan-Ritland (2003) described a portfolio of research activities using the following categories:

- Informed Exploration
- Enactment
- Evaluation: Local Impact
  - Quasi-experimental designs
  - Randomized trials
  - Hierarchical Linear Modeling
- Evaluation: Broader Impact
  - Implementation in new contexts (Design and Research)
  - Implementation at Scale
  - Scaling up Design and Research
  - Web-enabled proto diffusion
  - Diffusion of Innovations (Rogers)
  - Adoption, adaptation, acceptance, rejection

Of course, this larger framework calls for many different research methods. In his paper (Plomp, chapter 1), briefly captures the functions of research methods:

- *survey:* to describe, to compare, to evaluate
- *case studies:* to describe, to compare, to explain
- *experiments:* to explain, to compare
- *action research:* to design/develop a solution to a practical problem
- *ethnography:* to describe, to explain
- *correlational research:* to describe, to compare
- *evaluation research:* to determine the effectiveness of a program

He then provides examples related to the Chinese context:

1. *to describe:* e.g., what is the achievement of Chinese grade 8 pupils in mathematics?; what barriers to students' experience in the learning of mathematical modelling?
2. *to compare:* e.g., what are the differences and similarities between the Chinese and the Netherlands curriculum for primary education?; what is the achievement in mathematics of Chinese grade 8 pupils as compared to that in certain other countries?
3. *to evaluate:* e.g., how well does a program function in terms of competences of graduates?; what are the strengths and weaknesses of a certain approach?; etc.
4. *to explain or to predict:* e.g., what are the causes of poor performance in mathematics (i.e. in search of a 'theory' predicting a phenomenon when certain conditions or characteristics are met)?
5. *to design and develop:* e.g., what are the characteristics of an effective teaching and learning strategy aimed at acquiring certain learning outcomes?

Both Bannan-Ritland and Plomp provide a broader context for research. Within this larger framework, we may ask, therefore: When is design research appropriate? We may approach an answer by asking, first, when is design research inappropriate?

When is design research inappropriate?

A review of the many published examples of design research (e.g., Kelly, Lesh, & Baek, 2008) demonstrate the heavy investment of time and resources necessary to make progress in the face of sometimes daunting circumstances. Design research requires investment of substantial resources at many levels: school district administrators, teachers, students, and the design research team (which may include education researchers, software developers, curriculum specialists, and so forth).

Thus, design research is inappropriate if the educational problem is fairly simple.

If the problem has a known or standard solution, and there is general agreement on when to apply the solution, and the solution has been regularly successfully applied in various settings, design research is probably a poor use of resources.

Even for more chronic learning problems such as learning to read, if there are adequate training programs, and clear measures of success or progress (e.g., use of phonics to teach decoding skills), design research is probably not indicated. If, however, new research suggests a powerful innovation, design research may be a reasonable choice (see below, and McCandliss, Kalchman, & Bryant, 2003)

Generally, design research is probably not recommended for closed problems (e.g., improving mathematics calculation fluency), where the:

- Initial state(s) are known (e.g., two numbers are to be multiplied; a chess board is ready to play)
- Goal state(s) are known (e.g., a product of two numbers is to produced; checkmate or stalemate in chess)
- Operators to move from initial states to goal states are known and can be applied. (e.g., the procedures of multiplication; the rules of chess).

When is design research appropriate?

Design research is recommended when the problem facing learning or teaching is substantial and daunting how-to-do guidelines available for addressing the problem are unavailable. Further, it is recommended when a solution to the problem would lead to significant advances in learning or at least a significant reduction in malfunction in the educational system.

There should be little agreement on how to proceed to solve the problem, and literature reviews together with an examination of other solutions applied elsewhere (i.e., benchmarking) should have proven unsatisfactory.
Design research is further suggested if prior training or interventions have consistently proven unsuccessful. Design research is often indicated for critical educational goals, even when there is not a clear definition of success, or designing adequate indicators of success is part of the overall problem.

In other words, design research is most appropriate for *open*, or more appropriately, wicked problems. The concept of a wicked problem was used by Rittel and Webber (1977) to indicate those problems that share the features of open problems, but that also engage elements that make their solution frustrating or potentially unattainable.

Following from the description of closed problems, above, in open problems, some or more of the following apply:

- Initial state(s) are unknown or are unclear.
- Goal state(s) are unknown or are unclear.
- Operators to move from initial states to goal states are unknown or how to apply the operators is unclear.

For wicked problems (e.g., Camillus, 2008; Horn & Weber, 2007; Richey, 2013), the character of open problems pertain. Plus, there are typically inadequate resources, unclear "stopping rules" (conditions that indicate a solution is at hand or the project should be abandoned), unique and complex contexts, and inter-connected systemic factors that impinge on progress. Most frustrating, these other factors may themselves be symptoms of problems of associated wicked problems. For example, attempting to teach numeracy in a society with high poverty and HIV rates.

Therefore, one of the broad goals of design research is to dynamically clarify the initial and goal states and the operators, and to illuminate the nature of the problem – i.e., to "tame" a wicked problem by better specifying its character and making it open to intervention. In educational settings, design research is recommended when one or more of the following conditions operate to make the problem more wicked and open than simple and closed, for example:

- When the content knowledge to be learned is new or being discovered *even by the experts*.
- When how to teach the content is unclear: pedagogical content knowledge is poor.
- When the instructional materials are poor or not available.
- When the teachers' knowledge and skills are unsatisfactory.
- When the educational researchers' knowledge of the content and instructional strategies or instructional materials are poor.
- When complex societal, policy or political factors may negatively affect progress.

A number of examples may be found in Kelly, Lesh and Baek (2008). Some other examples from mathematics, science, and reading are briefly presented in the next section.

Examples from mathematics, science and reading

This section presents briefly a number of examples of when applying design research is the appropriate research approach.

1. Introducing existing science or mathematics at earlier grade levels

For example, some education authorities have advocated the teaching of algebra in earlier grades (as early as the 8th grade in the US), see Foundations for Success: Report of the National Mathematics Advisory Panel (http://www.ed.gov/about/bdscomm/list/mathpanel/index.html). A few policymakers have even advocated starting algebra instruction in the early elementary.

How should one proceed to introduce ideas of algebraic reasoning in the early elementary grades? Is this recommendation even advisable? This issue clearly meets the criteria set out, above. Some of the complexities associated with answering this question can be gauged by reading some of the recent work on this topic by Carraher and colleagues (e.g., Carraher & Schliemann, 2007; Carraher, Schliemann, & Schwartz, 2007; Peled & Carraher, 2007; Schliemann, Carraher, & Brizuela, 2007).

2. Learning new or emerging science content (e.g., genetics)

Research in microbiology is in revolution with stunning findings appearing on front pages of newspapers, almost daily, worldwide. How can traditional science education be updated to prepare high school teachers and students to meet this challenge and opportunity? Moreover, how can high schools prepare students to be successful in emerging integrated biology programs such as the one at Princeton University (http://www.princeton.edu/integratedscience/)?
Rutgers University has explored this challenge through its microbiology program (http://avery.rutgers.edu/WSSP/Begin/index.html). A review of its varied solutions to this challenge exemplifies this rich context for design research.

3. Uncovering the potential contributions of neuroscience for mathematics learning

The author has joined other writers (e.g., Varma, McCandliss, & Schwartz, 2008) in outlining the case for cultivating the intersection of neuroscience and mathematics learning. (e.g., Kelly, 2002, 2008).

Why is there a growing interest in neuro-mathematics education? A number of factors have coincided to support a surge in interest in brain-based mathematics education research (see OECD, 2007 for a comprehensive review of brain-related research in education):

- Confidence due to recent gains in understanding the brain bases for processes of decoding in reading.
- Emergent findings in the neural bases for mathematical thought.
- Decades of behavioral and cognitive science findings on learning mathematics and related higher-order processes from which to draw.
- A desire to disambiguate and constrain research hypotheses at the behavioral, cognitive and social levels of analysis.
- A desire to sharpen and ground diagnosis and remediation of mathematical learning difficulties with improved assessments.
- A desire to construct new mixed-methods research methodologies for the social sciences.
- A desire to scientifically debunk learning and teaching "neuromythologies".

- A sense of urgency to bring scientific discourse, evidence and reasoning to the slate of ethical issues that are emerging that pertain both to learning and teaching.
- A goal to improve methods of teaching of mathematics.
- A goal to improve educational materials, including those that use computer hardware and software.
- More comprehensive and testable models of learning emerging from cognitive science (e.g., Bruer, 1997).
- A desire to understand and promote significant mathematical creativity.
- The challenge to neuroscientists to push the boundaries of imaging technologies, and to co-formulate with domain experts clinical learning tasks.

The point to be drawn here is that the coincidence of these factors, alone, does not dictate teaching or learning strategies or even provide principles, materials, curricula, interventions, or assessment approaches to support either learning or teaching. How, then, should researchers proceed to bring the laboratory findings of cognitive neuroscience into the classroom in viable ways? Again, the problem meets the above requirements for using design research methods.

4. Cyberinfrastructure

Cyberinfrastructure encompasses the use of distributed internet resources such as computing systems, data, information resources, networking, digitally enabled-sensors, instruments, virtual organizations, and observatories (NSF, 2007). It allows to link groups of scientists to attack multi-level complex problems. These problems will have associated challenges for learning, teaching, and assessment.

Important questions center in how education should capitalize on cyberinfrastructure resources. What does it mean to study science content within a cyberinfrastructure framework, and what curricular, instructional design, assessment, teacher professional development, and policy questions are raised, and how must they be answered to fully exploit the high-technology investment in science at this level. As important, what are the methodological challenges in studying learning within a cyberinfrastructure project? For example, how are claims of causality handled in a complex networked and nested learning environment, and what evidence would make such claims credible (e.g., Kelly & Yin, 2007)? This is a clear example, spanning many science disciplines, for which design research is an appropriate investment.

The Appendix to this chapter discusses in more detail the meaning and possibilities of cyberinfrastructure, or e-science in general, for education.

5. Reading and inquiry science

The reader's attention is drawn to two examples from Bannan-Ritland. Her analysis of how design research works within her integrative learning design framework (Bannan-Ritland, 2003; Bannan-Ritland & Baek, 2008; see also chapter 4 in this volume) provides examples in narrower, if no less important, applications.

For example, design research is appropriate when developing creative or innovative educational products, blueprints or designs that are directed at chronic educational problems. A number of papers (see LiteracyAccess Online, Bannan-Ritland & Baek, 2008; http://immersion.gmu.edu/lao/spring2003/projectResources.htm), and Bannan-Ritland's chapter 4 (this volume) describe the processes undertaken to address a chronic problem in most countries, how to teach reading to struggling readers.

Following her work on LiteracyAccess Online, Bannan-Ritland extended her work on reading design into the learning of inquiry science at the 4th grade. Based on this experience, she significantly added to the broadening use of design research principles by methodologically incorporating teachers as designers in the overall design research paradigm. This exciting new direction, called teacher design research (which dovetails with work by Zawojewski, Chamberlin, Hjalmarson and Lewis, 2008 is described in Bannan-Ritland (2008). The area of application in the report is earth sciences in the early elementary school.

The growing need for design assessment research

A recent review of contributions to design research show an increasing awareness of the need for tackling the problem of how to assess learning in emerging areas of learning, particularly when there is an emphasis on innovation in instructional practices (Kelly, Baek, Lesh, & Bannan-Ritland, 2008, p.7). They note:

> "In design research as currently practiced, assessment is not directed at some summative sense of learning, though a summative measure of student learning would be central to later attempts at confirmatory studies, i.e. to show local impact (Bannan-Ritland, 2003). . . . Design research also differs from formative assessment with regard to the student's knowledge end state and how feedback loops are enacted. Formative assessment is the gathering of data relative to some predetermined fixed point, providing feedback that informs the students and teacher of their current knowledge state in relation to some end state (see Black & William, 1998). In design research, assessment may be used formatively in order to dynamically determine progress toward mastery of disciplinary knowledge (e.g., Cobb & Gravemeijer, [Kelly, Lesh, & Baek, 2008]) or to guide the design of a prototype

> and to inform its iterative re-design as necessary or both. In fact, sensitivity to assessment practices themselves may inform changes to the act of assessment itself (e.g., Lobato, [Kelly, Lesh, & Baek, 2008]; Lesh et al., [Kelly, Lesh, & Baek, 2008]). Ultimately, design researchers are challenging the assumptions about learning, teaching, and knowing that underlie available assessment techniques, not only in terms of the psychometric assumptions (like item response theory), but also the function of assessment itself within and across the stages of design research (see Sloane & Kelly, [Kelly, Lesh, & Baek, 2008])."

What is the *evidence* to support claims of effectiveness during iterations, and later, as the innovation is subject to more rigorous tests?

In other words, when a suitable context for design research is identified, to the extent that the application is novel (e.g., teaching algebraic concepts in the early elementary grades, reading comprehension) or the knowledge unfolding (e.g., genomics, cyberinfrastructure), there will be a requirement and a responsibility for researchers not only to iteratively investigate the impact of learning prototypes, but also to address directly the question of how this impact will be measured. The point here is not that assessment is necessary, rather that the targets for assessment may arise dynamically in the course of design research and measures may not be available apriori. As a result, many of the questions about the validity and reliability of measures have to be actively reconsidered. In practice, too often, prototypes are redesigned without specifying the evidence base (via assessment design) for the redesign. In many cases, design researchers appear to rely on personal judgment or subjective factors. Adding to the unfolding need for new methods in design assessment will be a major challenge and opportunity for scholars in the next decade (e.g., Kelly, 2005a, 2005b).

Design research in general practice

The goal of this chapter was to characterize design research at a broad level, and to provide some examples of where the significant resources associated with design research might be spent. I will finish with a general outline of how design research cycles unfold within a larger framework of research (Bannan-Ritland, 2003; Plomp, chapter 1 this volume). Using cognitive science, cognitive psychology and other social science methods such as surveys, case studies, clinical interviews, ethnography:

- Identify or characterize the initial states. Clarify the initial knowledge and goal knowledge states (of students, teachers, researchers, experts) using the interventions.
- Identify or characterize the goal states. Design formative assessments to monitor progress toward the goal state.

- Identify or characterize the operators. Dynamically using the cognitive and other analyses, iteratively design and specify the operators (interventions, supports, environments) to support learning. See, in particular, the work of Bannan-Ritland (2008) and Zawojewski et al., (2008).
- Inform re-design cycles or iterations using data gathered from unfolding, and parallel work in design assessment.
- Work toward developing a mature prototype that can be subject to a more definitive test (e.g., randomized clinical trial), see Bannan-Ritland's (2003) local impact phase.

One final note: Prototyping and theory building

By perturbing the system using the interventions in this iterative research process, design research transcends each of the local methods used. In other words, design research involves not only the use of different methods (e.g., surveys, case studies, clinical interviews), but combines the fruits of each method, over time, to specify theory and models related to learning, teaching and assessing the target knowledge (see Cobb & Gravemeijer, 2008). Thus, design research *goes beyond simple development of an intervention* and *goes beyond standard cognitive analyses* and allows theory and modeling that accounts for the content, the cognition, and the enactment by real people in real and rich contexts with real limits on resources (see Zaritsky, Kelly, Flowers, Rogers and O'Neill, 2003).

The question of the "theoretical yield" of design research is not a simple one. Note that this chapter was framed in terms of complex, open and wicked problems. For such problems, there exists no simple theoretical model (at least none is perceived at the time). For that reason, if "theory" is something that is assumed to be informed by hypothesis testing of a somewhat definitive question, then design research (in early stages) will likely not pose or easily answer simple hypotheses, and thus not have simple theoretical yield. Schwartz, Chang and Martin (2008; in Kelly, Lesh, & Baek) view the design research cycles as preparatory for theoretical yield from later randomized clinical trials or other laboratory tests. If the observation is borne out that much of educational intervention occurs in complex systems, then the theoretical yield will not be associated with one theory, but many (perhaps interdependent) subtheories. If so, then the yield may be diffuse and obfuscated by the influence of many factors that are not controlled in design research settings. Some researchers have attempted to frame design research within an overarching theory (say, "variation" theory, Holmqvist, Gustavsson, & Wernberg, 2008). The pay-off of this approach will inform us greatly about the role of theory in design research.

Some writers use the word "theory" more generally to encompass "design principles," and it may be the case that such principles can indeed be identified (see Kali, 2008). Such recommendations for design practice are useful heuristics. If these heuristics show evidence of durable applicability across many projects and contexts, it is likely that some necessary (as opposed to contingent) principles are being evoked (see Kelly, 2004), which would open these heuristics to theoretical analysis.

References

Bannan-Ritland, B. (2003). The role of design in research: The integrative learning design framework. *Educational Researcher 32(1), 21-24*.

Bannan-Ritland, B. (2008). Teacher design research: An emerging paradigm for teachers' professional development. In A.E. Kelly, R. Lesh, & J. Baek (Eds.), *Handbook of design research methods in education: Innovations in science, technology, engineering, and mathematics learning and teaching,* (pp. 246-262). New York: Routledge.

Bannan-Ritland, B. , & Baek, J. Y. (2008). Investigating the act of design in design research. In A.E. Kelly, R. Lesh, & J. Baek (Eds.), *Handbook of design research methods in education: Innovations in science, technology, engineering, and mathematics learning and teaching,* (pp. 299-319). New York: Routledge.

Barab, S. & Squire, K. (2004). Design-based research: Putting a stake in the ground. *Journal of the Learning Sciences, 13*(1), 1-14.

Camillus, J.C. (2008). Strategy as a wicked problem. *Harvard Business Review*, 1-11.

Carraher, D.W., & Schliemann, A.D. (2007). Early algebra and algebraic reasoning. In F. Lester (Ed.), *Second handbook of research on mathematics teaching and learning: A project of the National Council of Teachers of Mathematics.* Vol II. (pp. 669-705). Charlotte, NC: Information Age Publishing.

Carraher, D.W., Schliemann, A.D., & Schwartz, J. (2007). Early algebra is not the same as algebra early. In J. Kaput. D. Carraher, & M. Blanton (Eds.), *Algebra in the early grades* (pp. 235-272). Mahwah, NJ, Erlbaum.

Chin, G., & Lansing, C.S. (2004). Capturing and supporting contexts for scientific data sharing via the biological sciences collaboratory. *Proceedings of the 2004 ACM conference on computer supported cooperative work*, (pp. 409-418). New York: ACM Press.

Cobb, P. & Gravemeijer, K. (2008). Experimenting to support and understand learning processes. In A.E. Kelly, R. Lesh, & J. Baek (Eds.), *Handbook of design research methods in education: Innovations in science, technology, engineering, and mathematics learning and teaching* (pp. 68-95). New York: Routledge.

Cogburn, D.L. (2003). HCI in the so-called developing world: What's in it for everyone, *Interactions*, *10*(2), 80-87.

Holmqvist, M., Gustavsson, L., & Wernberg, A. (2008). Variation theory: An organizing principle to guide design research in education. In A.E. Kelly, R. Lesh, & J. Baek (Eds.), *Handbook of design research methods in education: Innovations in science, technology, engineering, and mathematics learning and teaching.* (pp. 111-130). New York: Routledge.

Horn, R.E. (2001). *Knowledge mapping for complex social messes.* A presentation to the "Foundations in the Knowledge Economy" conference at the David and Lucile Packard Foundation.

Kali, Y. (2008). The design principles database as means for promoting design-based research. In A.E. Kelly, R. Lesh, & J. Baek (Eds.), *Handbook of design research methods in education: Innovations in science, technology, engineering, and mathematics learning and teaching.* (pp. 423-438). New York: Routledge.

Kelly, A.E. (2003). The role of design in research. *Educational Researcher, 32*(1), 3-4. Theme issue guest editor.

Kelly, A.E. (2004). Design research in education: Yes, but is it methodological? *Journal of the Learning Sciences, 13*(1), 115-128.

Kelly, A.E. (2005a). *How do modeling perspectives inform program assessment?* Presentation at the 27th annual meeting of the North American Chapter of the International Group for the Psychology of Mathematics Education. Roanoke, VA.

Kelly, A.E. (2005b). *The intersection of design and assessment for mathematics learning.* Presentation, National Science Foundation Workshop, Park City, UT.

Kelly, A.E. (2006). Quality criteria for design research. In J.J.H. van den Akker, K. Gravemeijer, S. McKenney, & N. Nieveen (Eds.), *Educational design research.* London: Routledge.

Kelly, A.E. (2008). Brain research and education: Potential implications for pedagogy. In *Education, sciences cognitives et neurosciences.* Paris: Presses Universitaries de France.

Kelly, A.E., Baek, J., Lesh, R., & Bannan-Ritland, B. (2008). Enabling innovations in education and systematizing their impact. In A. E. Kelly, R. Lesh, and J. Baek (Eds.), *Handbook of design research methods in education: Innovations in science, technology, engineering, and mathematics learning and teaching.* (pp. 3-18). New York: Routledge.

Kelly, A.E. & Yin, R. (2007). Strengthening structured abstracts for educational research: The need for claim-based structured abstracts. *Educational Researcher,* 36(3), 133-138.

McCandliss, B.D., Kalchman, M., & Bryant, P. (2003). Design experiments and laboratory approaches to learning: Steps toward collaborative exchange. *Educational Researcher, 32*(1), 14-16.

Middleton, J., Gorard, S., Taylor, C., & Bannan-Ritland, B. (2008). The "compleat" design experiment: From soup to nuts. In A.E. Kelly, R. Lesh, & J. Baek (Eds.), *Handbook of design research methods in education: Innovations in science, technology, engineering, and mathematics learning and teaching.* (pp. 21-46). New York: Routledge.

National Science Foundation (2007). *Cyberinfrastructure vision for 21st Century discovery.* Office of Cyberinfrastructure.

OECD (2007). *Understanding the brain: The birth of a learning science.* Paris: OECD.

Olson, G.M., Teasley, S., Bietz, M. J., & Cogburn, D.L. (2002). Collaboratories to support distributed science: The example of international HIV/AIDS research. P*roceedings of the 2002 annual research conference of the South African institute of computer scientists and information technologists on enablement through technology,* 44–51.

Peled, I., & Carraher, D.W. (2007). Signed numbers and algebraic thinking. In J. Kaput, D. Carraher, & M. Blanton (Eds.), *Algebra in the early grades*, (pp. 303-327). Mahwah, NJ: Erlbaum.

Richey, T (2013). Wicked Problems: Modeling Social Messes with Morphological Analysis. *Acta Morphologica Generalis, 2*(1), 1-8.

Schliemann, A.D., Carraher, D.W., & Brizuela, B. (2007). *Bringing out the algebraic character of arithmetic: From children's ideas to classroom practice.* Mahwah, NJ: Lawrence Erlbaum Associates.

Schwartz, D.L., Chang, J., & Martin, L. (2008). Instrumentation and innovation in design experiments: Taking the turn towards efficiency. In A.E. Kelly, R. Lesh, & J. Baek (Eds.), *Handbook of design research methods in education: Innovations in science, technology, engineering, and mathematics learning and teaching.* (pp. 47-67). New York: Routledge.

Thomas, J.J., & Cook, K.A. (Eds.). (2005). *Illuminating the path: The research and development agenda for visual analytics.* Los Alamitos, CA: IEEE Computer Society.

Van den Akker, J.J.H., Gravemeijer, K., McKenney, S., & Nieveen, N. (Eds.) (2008). *Educational design research.* London: Routledge.

Varma, S., McCandliss, B.D., Schwartz, D.L. (2008). Scientific and pragmatic challenges for bridging education and neuroscience. *Educational Researcher, 37*(3), 140-152.

Wulf, W. (1989). The national collaboratory. In *Towards a national collaboratory.* Unpublished report of a National Science Foundation invitational workshop. New York: Rockefeller University.

Zaritksy, R., Kelly, A.E., Flowers, W., Rogers, E., & O'Neill, P. (2003). Clinical design sciences: A view from sister design efforts. *Educational Researcher, 32*(1), 32-34.

Zawojewski, J., Chamberlin, M., Hjalmarson, M.A., & Lewis, C. (2008). Developing design studies in mathematics education professional development: Studying teachers' interpretive systems. In A.E. Kelly, R. Lesh, & J. Baek (Eds.), *Handbook of design research methods in education: Innovations in science, technology, engineering, and mathematics learning and teaching.* (pp. 216-245). New York: Routledge.

Appendix

For the US National Science Foundation (NSF, 2007), the opportunities in the complementary areas that make up cyberinfrastructure: computing systems, data, information resources, networking, digitally enabled-sensors, instruments, virtual organizations, and observatories, along with an interoperable suite of software services and tools provide challenges along three lines: (a) data, data analysis, and visualization; (b) virtual organizations for distributed communities; and (c) learning and workforce development.

A major parallel activity in cyberinfrastructure is underway in Europe, which is labeled "e-science". E-science describes similar activities to the US cyberinfrastructure. Not unlike early visions of US cyberinfrastructure, the UK launching document, Developing the UK's e-infrastructure for science and innovation (http://www.nesc.ac.uk/documents/OSI/report.pdf), did not explicitly list education as one of the key areas of concern in setting up a cyber infrastructure. It focused, rather, on networks, middleware, digital libraries, and computational resources. As in the US, this imbalance is being recognized. In Europe, it being addressed by the creation of ICEAGE: "The international collaboration to extend and advance grid education". ICEAGE, while international, is primarily a European effort, with branches in Edinburgh, Scotland, University of Catania, Sicily, SPACI (Southern Partnership for Advanced Computational Infrastructures), an Italian university-based effort (http://www.spaci.it/), CERN, near Geneva (http://public.web.cern.ch/Public/Welcome.html), the Royal Institute of Technology in Sweden (http://www.kth.se/?l=en_UK), and The Computer and Automation Research Institute, Hungarian Academy of Sciences (http://www.sztaki.hu/institute).

Cyberinfrastructure describes the use of distributed internet resources to link groups of scientists to attack multi-level complex problems. These problems will have associated challenges for learning, teaching, and assessment. For example, a design research problem would be how to describe and credit a student's learning in a cyberinfrastructure research collaboratory in geosciences:
Scientifically, a crucial concern in detecting earthquakes is to measure minute changes in elevation. Traditional radar, which uses radio waves as the means of detecting distances from the source, are of limited value in precise measurements due to the length of the radio waves. The use of LiDAR (Light Detection and Ranging) technology allows the use of wavelengths in the ultraviolet, visible, or near infrared range (from about 10 micrometers to the UV (ca. 250 nm). These shorter wavelengths allow detection of smoke and other diffuse particulates, which has led to the use of LiDAR in meteorology.
For earthquake prediction, LiDAR can be used to locate faults, and to measure uplift. Faults describe the line of fracture and demarcation between plates (McKnight & Hess, 2000). Uplift is typically due to tectonic plate activity (Kearney & Vine, 1990), technically "orogenic

uplift" or due to the removal (due to erosion) of heavy material, technically "isostatic uplift." The significant advantage of LiDAR over radar is that LiDAR can generate digital elevation models (DEMs) of the shape the earth's surface at resolutions not previously possible. Complexifying the problem, an earthquake is sometimes associated with volcanic activity. For example, the "Pacific Rim of Fire" is associated with colliding tectonic plates. In such cases, LiDAR may be used not only to make precise measurements of elevation, but also to characterize the density and even the chemical makeup of the gases and ash emitted by a volcano.

Learning about geomorphology using LiDAR is complex, and some publicly available web sites have attempted to provide instruction (e.g., http://lidar.cr.usgs.gov/). The most comprehensive activity has been conducted by the GEON network (http://www.geongrid.org/). This network is part of a cyberinfrastructure research collaborator. Tutorials on the use of LiDAR within and outside of geoscience (e.g., coastal erosion, flooding, river courses, forest mapping and mining) may be found here;
http://home.iitk.ac.in/~blohani/LiDAR_Tutorial/Airborne_AltimetricLidar_Tutorial.htm.

We can now see just a fraction of the associated scientific concepts that are pertinent in understanding the use of LiDAR in understanding geoscience: e.g., radar technology vs LiDAR technology, the science of plate tectonics, digital elevation models, reading and understanding computer visualizations, modeling complex inter-related scientific processes, reasoning about implications for human activity, including urban growth, and so forth. Which of these (or other related concepts) are most pertinent for scientists in a cyberinfrastructure research collaboratory will be an empirical question. How to identify the central constructs pertinent to a high-school science education will provide a significant measurement challenge, including how to design authentic assessments to measure understanding of these concepts. Identifying and mapping out the content and cognitive demands of such measurement could be a major focus of the design research work. Of particular interest will be how to establish content, construct, predictive, concurrent and other forms of validity for these measures.

Factors converging to support the development of cyberinfrastructure.

1. Existing computing data grids in the US and overseas
   a. The XSEDE project combines the power of NCSA, SDSC, Argonne National Laboratory, CACR, PSC, ORNL, TACC, and various university partners integrated by the Grid Infrastructure Group at the University of Chicago. European e-science links facilities on the Continent with those in the UK. Similar activities occur in Japan. Industry partners include IBM, Intel, Hewlett-Packard and Oracle.
2. The availability of massive data storage capacity and speed
   a. The XSEDE currently offers over 100 teraflops of computing power; and over 3 petabytes of rotating storage
3. The development of middleware and software to gather and analyze stored data
   a. The XSEDE supports data analysis and visualization production interconnected at 10-30 gigabits/second.
4. The emergence of large teams of scientists dedicated to solving shared science problems (acting through science "collaboratories" and "gateways")
   a. A collaboratory (Wulf, 1989) is "more than an elaborate collection of information and communications technologies; it is a new networked organizational form that also includes social processes; collaboration techniques; formal and informal communication; and agreement on norms, principles, values, and rules" (Cogburn, 2003, p. 86). Collaboratories exist in many areas of science, including biology, chemistry, geoscience and astronomy (e.g., Chin & Lansing, 2004; Olson, Teasley, Bietz, & Cogburn, 2002).
   b. Science gateways are web-based portals or interfaces for the structures and data of the cyberinfrastructure in many science areas (for a listing of gateways, see http://rt.uits.iu.edu/visualization/gateways/index.php).
5. Developments in scientific visualization.
   Scientific visualization draws on human spatial and visual processing in order to model and analyze computationally intense the graphic display of complex data (for a comprehensive review, see Thomas & Cook, 2005). Existing methods and models for scientific visualization are significantly challenged by cyberinfrastructure (e.g., Chin et al., 2006).
6. Funding.
   The establishment and funding of national and international efforts to coordinate and develop the infrastructure to better serve science and, more recently, education (e.g., the Office of Cyberinfrastructure – NSF; CERN, Dutch (VL-e) and UK initiatives). The promise of cyberinfrastructure for education is that the vast investment by US agencies (upwards of $250M over the next 5 years, alone) will provide test-beds for exploration.


40
40
40
40
48
48
52
52
100
52
52
48
48
40

6. Formative Evaluation in Educational Design Research

Nienke Nieveen & Elvira Folmer

Introduction

In the general introduction of this book, two possible purposes of educational design research were identified leading to a distinction between development studies and validation studies (Plomp, 2013). This chapter starts from the perspective of *development studies*. In line with the definition introduced by Plomp (2013), we define educational design research as *the systematic analysis, design and evaluation of educational interventions with the dual aim of generating research-based solutions for complex problems in educational practice, and advancing our knowledge about the characteristics of these interventions and the processes of designing and developing them*. This type of design research has a twofold yield. The first yield comprises high-quality, research-based interventions designed to solve complex problems in educational practice. This type of output brings forward the practical relevance of design research. It is for that reason that design research is also labeled as being use-inspired, applied oriented and/or socially responsible research (van den Akker, 1999; Reeves, 2000). The second main yield is a set of well-articulated design principles (Linn, Davis, & Bell, 2004; van den Akker, 1999) that provide insight into the:

- purpose/function of the intervention
- key characteristics of the intervention (substantive emphasis)
- guidelines for designing the intervention (procedural emphasis)
- its implementation conditions and
- theoretical and empirical arguments (proof) for the characteristics and procedural guidelines.

These comprehensive design principles serve several purposes for a variety of target groups:

- for researchers (these principles show the contribution of design research to the existing knowledge base with information on how the intervention works in practice, the effects of using the intervention and explanation of the working mechanisms)
- for educational designers (these principles carry rich information on how to design similar interventions for similar settings)
- for future users (the principles provide information needed for selecting and applying interventions in the specific target situation and provide insights in the required implementation conditions)
- finally, for policy makers (these principles assist in making research-based decisions to address for solving complex educational problems).

In order to reach both ends (high-quality interventions and design principles), design researchers carefully combine and interweave design and research activities. Usually, these design research studies start with a *preliminary research phase*, comprising various systematically performed analysis activities, such as needs and context analysis and a review of relevant literature, a conceptual framework for the study, set of tentative design guidelines and an accompanying design proposal. Next, the *prototyping or development phase* is carried out, including several cycles/iterations of analysis, design and formative evaluation, eventually leading to a complete intervention and a set of final design principles. Many design research studies end with a *summative evaluation phase*, leading to more confident assertions about the results of the study. The outcomes of the summative evaluation can also be used as input during a preliminary research phase of another, new design research study.

In this chapter we will first briefly elaborate the first (preliminary research) and last (summative evaluation) phases. Then we will more extensively branch out the role that formative evaluation plays within the development or prototyping phase.

Preliminary research phase

The preliminary research phase is needed to gain insight into the educational problem at stake (the gap between the current and desired situation). The core question is: which educational problem does the intervention need to address? This phase will contribute to the quality of the future intervention by obtaining insight in the existing situation, the needs of those involved, the conditions for innovation and potential approaches to address the problem. The aim of this phase is two-fold:

- to gain insight into the existing problem situation and the possibilities for improvement and innovation; and
- to specify the desired tentative features of the intervention (tentative design principles) and how these can be developed.

Important activities that are typically performed during a preliminary research phase include an analysis of the user practice (needs and context analysis) and an exploration of the scientific knowledge base (literature review and expert appraisal). A needs analysis looks into the perceptions of stakeholders on the current situation—what works well, what should be changed—and the features of a more desirable situation. A context analysis is aimed at exploring the problem environment and mapping out the scope for innovation. Questions to be asked during a context analysis include: What does the user context look like? What is the innovation scope, considering needs and abilities of those involved, e.g. their willingness to change, and conditions in the school, e.g. room for collaboration? What means, including time, finances, and staff, are available for development? Methods that

are frequently used in needs and context analyses include interviews, focus groups, lesson observations, document analysis and case studies.
In order to make relevant and valid design decisions, it is important to gain insight into the state of the art knowledge base. This can be done by means of literature review, expert appraisal, and the analysis and evaluation of existing projects and products that address similar problems. Questions asked during the knowledge-base analysis focus on the following: What recent insights from educational research and subject matter discipline may be used in the design? And what available (related, promising) interventions could serve as a source of inspiration and what lessons may be learned from the implementation and the impact of these products?

Summative evaluation phase

At the very end of a design research study, the summative evaluation phase is aimed at determining the actual effectiveness of the complete intervention (which has resulted from the development or prototyping phase). The focus is on the extent to which implementation of the intervention leads to the desired outcomes. These desired outcomes are related to the intended outcomes of the study.

It is important not to carry out a summative evaluation until the intervention is developed to such an extent that it has sufficient *potential* effectiveness. In order to have this potential effectiveness, the intervention should at least be relevant for the educational problem or need at hand, and it should be logically designed and practical in use. This means that design researchers, before entering the summative evaluation phase, need to be able to provide convincing evidence for the quality of the intervention so far, on the basis of the formative evaluation activities undertaken during the development or prototyping phase. The decision whether or not to plan a summative evaluation depends also on the kind of impact of the intervention. If a new nation-wide curriculum is being developed, than it seems to be justified to invest the required means for a thorough summative evaluation. If the impact of an intervention is much smaller, it may be decided not to do so. Under these circumstances decisions for future implementation of the intervention will be based on the results of final formative evaluation activities (such as trials in pilot settings).

The reason for not performing a full-fledged summative evaluation has to do with the fact that these types of evaluations are costly, time consuming and need to meet criteria that are hard to meet in educational settings. Being more precise, the most powerful research design to reveal cause-effect relationships (e.g. did the intervention cause the increase in learner results?) is a (quasi-)experiment. Table 1 elaborates some criteria that need to be taken into consideration when planning and performing a (quasi-)experiment and its accompanying

issues when applying them in educational settings (Gravemeijer & Kirschner, 2008; Rossi, Lipsey, & Freeman, 2004; Gay, Mills, & Airasian, 2012; Wayne, Yoon, Zhu, Cronen, & Garet, 2008).

Table 1: Criteria and issues related to (quasi-)experimental designs in educational settings

Criteria for (quasi-) experimental designs	Issues related to effect studies in education
Similarity between intended and implemented intervention	When implementing an intervention in educational practice, the various parties involved, such as teachers and pupils, will usually carry out this intervention according to their own needs and wishes. This enactment may affect the learning results, which means that insight into the implemented intervention is important for the interpretation and explanation of the effect results.
Comparability of groups	In educational settings it is often not possible to randomly assign respondents to groups. If that is the situation then it is essential to ensure that groups are made comparable by means of matching or statistical control. The groups must be comparable in characteristics that influence the effectiveness of the intervention. Furthermore, it is important to look for an adequate representation of the population of schools, allowing generalization of the results.
Overlap between the intervention and the test	In effect studies, already available tests are often used. When these tests do not measure the new outcomes, this will lead to validation problems and possible effects may not be revealed. Moreover, the moment of the assessment should be well chosen: it may take a long time before learning effects can be measured.

Educational design researchers have found alternative ways for planning and performing summative evaluation, for example, by combining a large-scale survey (for instance on the implementation of new examination programs and corresponding results of pupils) with several in-depth case studies (focusing on the teaching practice linked to these new programs). Although this type of summative evaluation cannot detect cause-effect relationships it can give fruitful information on the effectiveness of the intervention in a cost-effective way.

Development or prototyping phase

The development or prototyping phase commences after the preliminary research phase has ended and a first set of tentative design guidelines and an accompanying design proposal for the intervention are in place. As design research studies typically aim at innovative and complex interventions, with few experiences from which to draw, applying a prototyping approach is recommended for this phase. The term prototype refers to a

tentative version of the whole (or part of an) intervention before full commitment is made to implement it. During the development or prototyping phase, several prototypes are being developed, evaluated and revised, which makes this phase highly iterative. Each iteration or cycle helps to develop and improve both end results of design research efforts: the educational intervention under development; and its accompanying design principles.

Prototypes may be used in at least two ways. A prototype may be continually refined (based on formative evaluation results and reflections of developers on the prototype) and evolve towards a final deliverable. This refining approach can be termed *evolutionary* prototyping. For instance, the development of an innovative learning and teaching situation usually involves this approach. Besides, many design research studies also make use of *throw-away* prototypes, such as paper-based prototypes (Nieveen, 1999; 2013). In the case of computer-based educational interventions, a paper-based prototype (comprising a set of papers representing all screens which may appear during the use of the software) can be used in the process of user-interface design. Future users may 'walk through' the screens to get an idea of the intentions of the software application. After being evaluated, a throw-away prototype will be discarded and its evaluation results taken into account in the next prototype.

To make the prototyping process more systematic and manageable, the idea of 'think big, but start small' is helpful in two ways. First of all, one can start by developing a small part of the entire proposed intervention (for instance, one complete lesson of an innovative learning and teaching situation, or one module of a course). By evaluating this first lesson or module with teachers and learners, one can learn from inaccuracies and apply successes when designing the subsequent parts, eventually leading to a final version of the entire intervention. Secondly, it can also be functional to break down the final intervention into several components and to develop these components separately. Educational interventions can be broken into at least two key parts (cf. Nieveen, 1999; Nieveen & van den Akker, 1999):

- the conceptual framework of the intervention, referring to all notions underlying the intervention. For instance, if the intervention is based on problem-based learning, usually one will focus first on defining problem-based learning and its consequences for the objectives, the learner and teacher activities and assessment strategies that need to be included in the curriculum; and
- the presentation-mode of the intervention, referring to the format that assures the intervention is usable for its target group. For instance, improve the usability all interventions need to have a sound and consistent layout.

Although components of the intervention may be in different stages of development in each prototype, in the final and complete version, all components need to be assembled and made consistent. For example, in a math project aimed at solving the problems low-achieving students have with measuring quantities, lesson materials were developed to help these

students to acquire the required mathematical problem-solving skills. The design research team put much effort into the development of innovative learning and teaching activities geared to the problems of these students. During the formative evaluation of this first prototype, the design research team was especially interested in the quality of the new learning and teaching activities and less so in the layout of the materials. However, towards the end of the study, the layout of the materials got specific attention in order to improve the overall practicality of the materials for the learners and teachers. The prototyping process finishes when all uncertainties are covered and the final intervention can be delivered and implemented.

Formative evaluation within the development/ prototyping phase

In order to gain insight into the quality of tentative interventions and their design principles and to get revision decisions for developing the next—improved—prototypes, empirical data are needed. For that reason, formative evaluation is a crucial feature of the development or prototyping phase. Results of the formative evaluation give ground for both outputs of a design research study: improving the prototype of the intervention towards a high-quality and completed intervention; and sharpening the underlying tentative design principles towards a final set of design principles.

As far as the term evaluation is concerned, the Joint Committee on Standards for Educational Evaluation (1994) uses the following definition: "Evaluation is the systematic assessment of the worth or merit of some object." Merit refers to the object's inherent, intrinsic value, while its worth is defined as its contextually determined, place-bound value (Lincoln & Guba, 1979). The function of formative evaluation is 'to improve'. It focuses on uncovering the shortcomings of an object during its development process with the purpose of generating suggestions for improving it. The function of summative evaluation is 'to proof'. A summative evaluation is carried out to gain evidence for the effectiveness of the intervention and to find arguments that support the decision to continue or terminate the project. Based on a comparison and synthesis of the definitions of various scholars in the field of formative evaluation (cf. Brinkerhoff, Brethouwer, Hluchyj, & Nowakowski, 1983; Flagg, 1990; Scriven, 1967; Tessmer, 1993), we define formative evaluation in the context of educational design research as *a systematically performed activity (including research design, data collection, data analysis, reporting) aiming at quality improvement of a prototypical intervention and its accompanying design principles.*

As a design research project usually needs several cycles or iterations before an optimal solution for the complex problem can be reached, each design research cycle or iteration has its specific formative evaluation with its specific research questions and thus needs its

own appropriate research design (see also chapter 1 where Plomp stresses that each cycle is a micro-cycle of research). The remainder of this section will elaborate on the appropriate research design for each cycle.

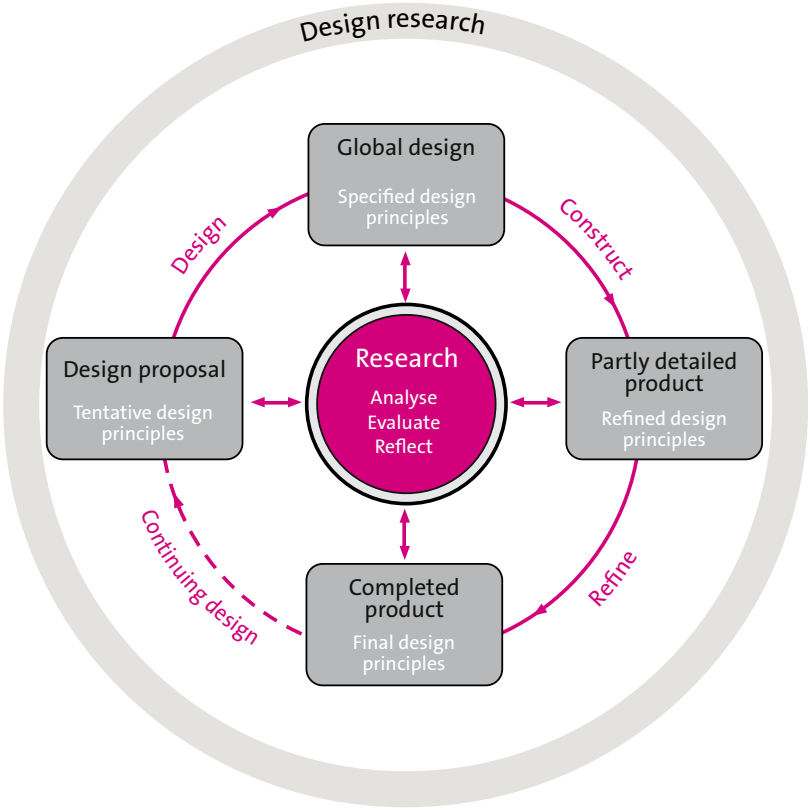


Figure 1: Development or prototyping phase in an educational design research study

Successive approximation through stages of development

In Figure 1 we illustrate the successive approximation of both outputs of a design research study during the development or prototyping phase. This model distinguishes four stages:

- Design proposal: A prototype in this first stage contains a general description of the future intervention in which attention is paid to its substantive parts. Usually, this 'sketch' is written based on preliminary research results, including needs and context analysis and a review of relevant literature.
- Global design: A prototype at this stage provides tentative details of some or all components of the future intervention. This could also be termed a 'horizontal prototype'. It gives an idea of how the intervention will eventually appear, although

it cannot yet be used in practice. For example, in the case of the development of an innovative curriculum at this stage the prototype could take the form of a table of contents with a brief description of the sub-components or modules.

- Partly detailed intervention/product: At this stage, a part of the future intervention has been elaborated to such an extent that it can be used in practice by the target group. This could be called a 'vertical prototype'. For instance, next to having an overview of all modules of the prospective learning materials, only one module addresses all details. Based on the results of the formative evaluation, this first detailed module will be adapted and the remaining modules will be constructed.
- Completed intervention/product: The intervention is complete and can be used in the intended user-setting.

Figure 1 provides a model with all four stages of the product and their accompanying design principles. The development or prototyping phase starts with a design proposal based on a tentative set of design principles (as a result of the preliminary research phase). After a number of iterations per stage, the project evolves via a global and partly detailed intervention towards a completed one. Together with the intervention, the design principles evolve from a tentative set via a more specific and refined set towards a final set of design principles. Moreover, each stage may consist of several cycles or iterations of analysis, design and formative evaluation before the prototype will grow into a next development stage.

Quality criteria for interventions

Educational design research strives to design a high-quality solution for a complex problem in educational practice. When it comes to the concept of quality, we distinguish four quality criteria that are applicable to a wide array of educational interventions (cf. Nieveen, 1999; 2013):

- Relevancy: There is a need for the intervention, and its design is based on state-of the art (scientific) knowledge – also called content validity
- Consistency: The intervention is 'logically' designed – also called construct validity
- Practicality:
  - Expected: The intervention is expected to be usable in the setting for which it has been designed
  - Actual: The intervention is usable in the setting for which it has been designed
- Effectiveness
  - Expected: Using the intervention is expected to result in desired outcomes
  - Actual: Using the product results in desired outcomes.

At the end of a design research study, the intervention should suffice for all of these criteria. They are linked to one another in a hierarchical way, as can be illustrated by a 'rhetoric' question like: "if an intervention is *not* practical, why would it make sense to investigate its effectiveness?" (see also Plomp, 2013). This logical hierarchy implies that in different development stages different quality criteria are emphasized, see also Table 2. For instance,

during the first stages (design proposal and global design), most attention will usually be given to the relevancy and consistency of the prototypes, whereas the design research team will be concerned about practicality when parts of the prototypr are elaborated in detail. Effectiveness will become increasingly important in the latest stages.

Moreover, it is important to point out the distinction between expected and actual practicality and expected and actual effectiveness. Only when the target users have had practical experience using the intervention will one be able to get data on the actual practicality of the prototype. Similarly, only when target users have had the opportunity to use the intervention in the target setting, will the evaluator get data on the actual effectiveness. In all other instances, the researcher will only get data on the expected practicality and/or expected effectiveness.

Formulating research questions for each development stage

Each evaluation starts with a main research question. The building blocks for the main research question(s) of a formative evaluation are provided by the kind of value judgment that the research team is looking for (focus on what quality criteria?) and the current development stage of the intervention (design proposal, global design, partly-detailed or completed product). The syntax of these questions is: *'What is the [quality criterion a, b, c and/or d] of the prototype that is in [development stage w, x, y, z].* Examples of such questions are:

- What is the relevancy [quality criterion] of the content of a quick reference manual for using Chinese characters that is in a global design stage [development stage]?
- What is the internal consistency [quality criterion] of the attainment targets for science in upper secondary education in which three out of seven domains are elaborated in detail [development stage]?
- What is the practicality [quality criterion] of the innovative Math module that is in a completed stage [development stage]?
- What is the effectiveness [quality criterion] of the complete/fully detailed professional development module [development stage] on the most recent changes in civics education?

Building on earlier research (Nieveen, 1997, 1999), Table 2 provides an overview of the relationship between the quality criteria (on the vertical axis) and the development stages of an intervention (horizontal axis).

Selecting appropriate methods

For each stage, design researchers need to select those formative evaluation methods that fit the research question(s). In Table 2 suitable formative evaluation methods are indicated in the cells (one may also refer to the evaluation matchboard in appendix 1). Here we distinguish the following methods and corresponding data collection activities:

- Screening: Members of the design research team check the design. Data could be collected by using a checklist containing the required characteristics of the intervention.

- Focus group (also referred to as expert appraisal): A group of respondents reacts to a prototype of a product. Data could be collected by organizing interviews.
- Walkthrough: The design research team and representatives of the target group together go over the prototype (like in a theater play). Possible data collection activities are: using a checklist, interviewing and observing the respondents when they are running through the prototype.
- Micro-evaluation: A small group of target users (e.g. learners and/or teachers) uses parts of the product outside its normal user setting. Possible data collection activities are interviewing, observing, administering a questionnaire, and assessing the performance of learners through a test, a learner report and/or a portfolio.
- Try-out: The target group uses the product in practice. If the evaluation focuses on the practicality of the intervention, the following evaluation activities are common: observation, interviewing, requesting logbooks, and administering questionnaires. If the evaluation focuses on the effectiveness of the intervention, evaluators may decide to assess the performance of learners through a test, a learner report and/or a portfolio.

For more information on each of the methods, please also refer to Tessmer (1993) or Brinkerhoff, et al. (1983).

Table 2: Table for selecting formative evaluation methods

Quality criteria	**Design stage →**	**Design proposal**	**Global design**	**Partly detailed intervention/ product**	**Completed intervention/ product**
Relevancy		- Screening - Focus group	- Screening - Focus group	- Screening - Focus group	- Screening - Focus group
Consistency		- Screening - Focus group	- Screening - Focus group	- Screening - Focus group	- Screening - Focus group
Practicality	**expected**	- Screening - Focus group	- Screening - Focus group	- Focus group - Walkthrough	- Focus group - Walkthrough
	actual			- Micro-evaluation	- Micro-evaluation - Try-out
Effectiveness	**expected**	- Screening - Focus group	- Screening - Focus group	- Focus group	- Focus group
	actual			- Micro-evaluation	- Micro-evaluation - Try-out

Note: grey indicates that with the shift in focus (from relevancy and consistency to practicality and effectiveness) from one stage to the other, other, more suitable, evaluation methods will also come into play.

Sampling - selecting respondents

To be able to answer the research questions with the chosen evaluation methods, the required type and number of respondents need to be discussed. The type and sample size depend on the research questions. With respect to the type of respondents, one needs to select those respondents that can help answer the research questions. For instance, in case design researchers want to gain insight into the relevancy of the design from a subject matter perspective, they may select a number of experts in that specific domain to do a focus group. If insights are needed in the actual practicality of a learning package for learners by performing a micro-evaluation, students need to be sampled who will have to work with the intervention. Moreover, the main purpose of the evaluation also influences the sample *size*. In case of a formative evaluation during the early stages of the project, the main purpose is to locate shortcomings in the intervention and to generate suggestions for improvement (see also definition of formative evaluation), for which the number of respondents is less critical: a remark of only one respondent could be highly valuable because of its salience. Small samples of respondents are usually sufficient if they are carefully selected. Samples are usually deliberately chosen (also referred to as purposive sampling where subjects are selected because of some characteristic) in such a way that the comments and reactions will be as information-rich as possible. This means that for instance, for organizing a micro-evaluation in order to gain insights into the practicality of a prototype of some learner materials, next to high-achieving students as well as a group of low-achieving and a group of average students need to be selected. Triangulation is important here in order to enhance the reliability and internal validity of the findings (cf. Miles & Huberman, 1994). One could triangulate by combining different data collection activities (including different types of respondents, different instruments, different times, and different places). The effectiveness of triangulation rests on the premise that the weaknesses in each single data collection activity will be compensated for by the counterbalancing strength of another.

When inviting respondents for a formative evaluation it is necessary to clarify their role. They could fulfill the role of learner, critic and/or reviser (Weston, McAlpine, & Bordonaro, 1995). Respondents with a learner role are not necessarily expert in the subject matter which is covered by the materials. One could think of students who learn a new subject, as well as teachers who have not taught in a certain manner before. In many cases experts in a certain area can also take the learner role in formative evaluation of an intervention. In many cases experts represent this category. For instance, educational technology experts do not always have expertise in the subject matter domain of the educational intervention. They will take the role of a learner first, before giving comments on matters related to educational technology (in which they are experts). Critics are respondents who are asked to comment on the materials from the perspective of their expertise. This group consists, for instance, of subject matter experts and teachers who are invited to make statements about

the difficulty or readability of learner materials. Revisers will not only give comments on the materials (like critics do), but they will also provide suggestions for improvements. For instance, a subject matter expert may indicate what 'state-of-the-art knowledge' is missing in the learner materials and where this knowledge can be found. It is important to note that individuals may play several roles simultaneously during the formative evaluation. The next sub-section will elaborate on the role of researchers during a formative evaluation.

Researchers' role during formative evaluation

Since a design research project comes into play when there is a need to solve a complex educational problem for which no ready-made solutions are available, often a multi-disciplined team is brought together to work on it. Such teams usually comprise experts in domains that were identified when breaking down the intervention (e.g. from a conceptual point of view: subject matter experts, pedagogical experts, instructional designers; from a presentation-mode point of view: e.g. graphic and user-interface designers) as well as members of the target group. Practitioner involvement is one of the key characteristics of educational design research. Involving future users in a design research team has several advantages: more accurate information about the complexity of the problem at hand, more intensive discussions about the requirements of the intervention, increase of user commitment and ownership of the final deliverable, increase of insights into the requirements of the context in which the intervention will be used, and stimulation of the professional development of all participants.

Next to designing and constructing the prototypes of the intervention, one of the key responsibilities of the design research team is to work on the formative evaluation of the prototypes. For reasons of scientific rigor, it is often recommended to look for external evaluators. However, certainly in the early stages of a design research study, it seems legitimate or even advisable that design researchers *themselves* carry out the formative evaluation of the prototype. Engaging in formative evaluation activities tends to lead design researchers toward important learning experiences. They will experience for themselves the problems that may occur and hear first hand the suggestions for improvement that respondents come up with during their use of a prototype (for example, by observing or interviewing teachers or students). This usually has stronger and more direct impact on their thinking and design activities than cases where external evaluators report the results to the developers.

Of course, design researchers need to be aware of several pitfalls when they are involved in the formative evaluation of the intervention they are also designing (cf. McKenney, Nieveen, & van den Akker, 2006; Plomp, 2013). They may easily become too 'attached' to their prototype which could lead to a less objective view toward problems and comments from the respondents. In this respect, Scriven (1991) warns against what he calls 'a (co-)authorship

bias'. Moreover, respondents could be biased during the evaluation. For instance, if they know how much effort the design research team has put into the design of the prototype, they may hesitate to be fully critical of it. This is another reason why it is important in an educational design research project to perform formative evaluations at an early stage in the design process and to apply various ways of triangulation. In the final stages of the development or prototyping stage, the design research team usually organizes try-outs with a larger group of representatives of the target group and with external evaluators responsible for data collection and reporting.

Closing remark

This chapter focuses on the empirical data the design research team collects during formative evaluation activities of each prototype in the development or prototyping phase of educational design research studies. Each stage (or prototype) gives the design research team firmer ground and arguments for the final intervention the team is working on in order to solve a complex educational problem. The formative evaluation results will not only provide suggestions for improving the prototypical intervention, but will also assist in sharpening the accompanying design principles. The development or prototyping phase will eventually lead to the final phase of the scientific cycle in which claims of causality can be studied in summative evaluation settings (cf. Nieveen, McKenney, & van den Akker, 2006).

We are aware that there is much more to say about formative evaluation in general, and how it can be integrated into design research projects in particular. For instance, in this chapter we did not pay attention to evaluation instrument development, data collection, data analysis and reporting. Several helpful books and articles are available to assist in systematically conducting formative evaluation in educational settings (cf. Brinkerhoff, et al., 1983; Flagg, 1990; Tessmer, 1993, and more recently, McKenney & Reeves, 2012). Although some of these sources were not written with the specific needs and wishes of educational design researchers in mind, they can provide ample inspiration.

Appendix 1

To help educational design researchers plan formative evaluations, Nieveen, Folmer and Vliegen (2012) developed the 'Evaluation Matchboard' (see Figure 2 and 3)[1].

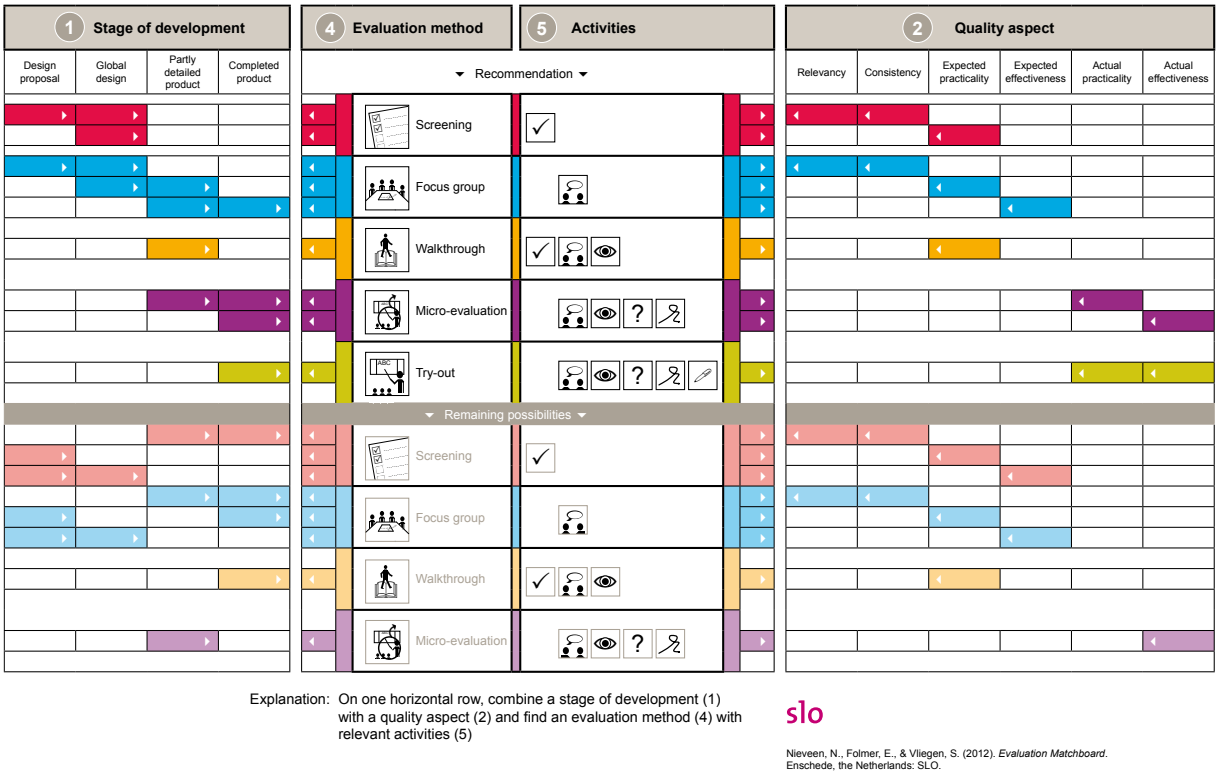


Figure 2: Frontside of the 'Evaluation Matchboard' (Nieveen, Folmer and Vliegen, 2012)

The frontside of the matchboard represents the information given in Table 2. When using the matchboard one needs to combine the stage of development (prototype) [1] with a quality aspect [2] in order to find an appropriate evaluation method [4] with relevant data collection activities [5]. For example, if design researchers will evaluate a *partly detailed product* and they wish to focus the formative evaluation on the *expected practicality* of this prototype, then the matchboard shows a 'match in colors' with respect to a focus group (blue) and a walkthrough (orange), both being suitable methods for this combination of development stage and quality criteria.

1) The 'Evaluation Matchboard' is printed on A3-format (for a pdf of the matchboard please refer to http://leerplanevaluatie.slo.nl/english/.

When deciding to organize a focus group, the matchboard recommends that data can be collected by interviewing the respondents. In case of organizing a walkthrough, using a checklist and carrying out observations are suitable data collection activities, according to the matchboard. At the backside of the matchboard (see Figure 3) information is given concerning the definitions of the stages of development, quality aspects, evaluation methods and activities.

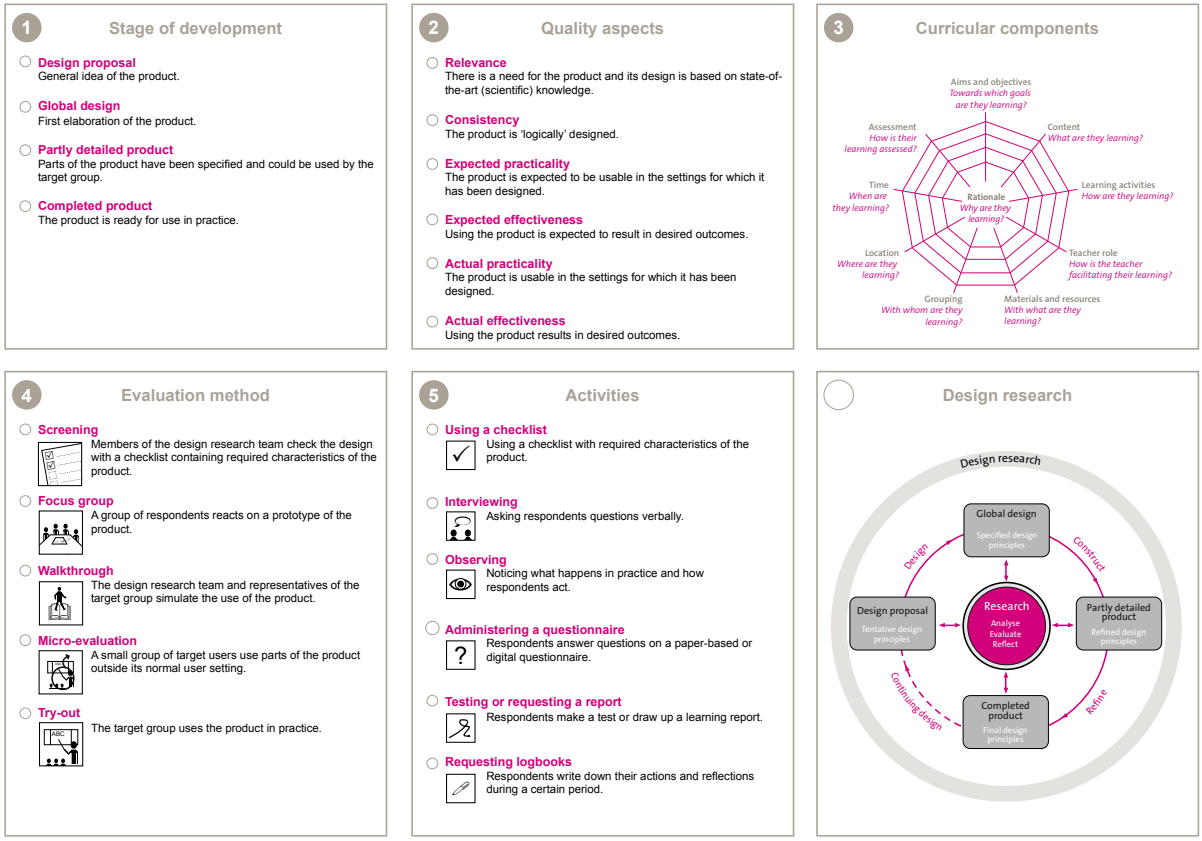


Figure 3: The backside of the 'Evaluation Matchboard' (Nieveen, Folmer and Vliegen, 2012)

References

Brinkerhoff, R.O., Brethouwer, D.M., Hluchyj, T., & Nowakowski, J.R. (1983). *Program evaluation: A practitioner's guide for trainers and educators.* Boston: Kluwer-Nijhoff.

Denscomb, M. (2007). *The good research guide for small-scale social research projects.* Maidenhead (UK): Open University Press.

Flagg, B.N. (1990). *Formative evaluation for educational technologies*. Hillsdale, NJ: Lawrence Erlbaum Associates.

Gay, L.R., Mills, G.E., & Airasian, P.W. (2012). *Educational research: Competencies for analysis and applications*. London: Pearson.

Gravemeijer, K.P.E., & Kirschner, P.A. (2008). Naar meer evidence-based onderwijs? [Towards more evidence-based education?] *Pedagogische Studiën, 84*, 563-571.

Joint Committee on Standards for Educational Evaluation (1994). *The program evaluation standards.* Thousand Oaks, CA: Sage Publications.

Lincoln, Y., & Guba, E.G. (1979). The distinction between merit and worth in evaluation. *Paper presented at the Annual Meeting of the Evaluation Network* (5th, Cincinnati, OH, September 24-26, 1979) retrieved October 20, 2008, http://www.eric.ed.gov/ERICDocs/data/ericdocs2sql/content_storage_01/0000019b/80/32/b5/c8.pdf

Linn, M., Davis, E.A., & Bell, P. (2004). *Internet environments for science education.* Hillsdale: Lawrence Erlbaum.

McKenney, S., Nieveen, N., & van den Akker, J. (2006). Design research from a curriculum perspective. In J. van den Akker, K. Gravemeijer, S. McKenney, & N. Nieveen (Eds.), *Educational design research* (pp. 67-90). London: Routledge.

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

Miles, M.B., & Huberman, A.M. (1994). *Qualitative data analysis: An expanded source book.* London: Sage.

Nieveen, N.M. (1997). *Computer support for curriculum developers: A study on the potential of computer support in the domain of formative curriculum development.* Doctoral dissertation. Enschede, the Netherlands: University of Twente.

Nieveen, N. (1999). *Prototyping to reach product quality.* In J. van den Akker, R. Branch, K. Gustafson, N. Nieveen, & T. Plomp (Eds.), *Design approaches and tools in education and training* (pp. 125-36). Dordrecht, the Netherlands: Kluwer Academic Publishers.

Nieveen, N. (2013). A four-year design research study improving curriculum developers' formative evaluation through an Electronic Performance Support System (EPSS). In T. Plomp & N. Nieveen (Eds.), *Educational design research - Part B: Illustrative cases* (pp. 1101-1123). Enschede, the Netherlands: SLO.

Nieveen, N.M., & van den Akker, J.J.H. (1999). Exploring the potential of a computer tool for instructional developers. *Educational technology research & development, 47*(3), 77-98.

Nieveen, N., McKenney, S., & van den Akker, J. (2006). Educational design research: The value of variety. In J. van den Akker, K. Gravemeijer, S. McKenney, & N. Nieveen (Eds.), *Educational design research* (pp. 151-158). London: Routledge.

Nieveen, N., Folmer, E., & Vliegen, S. (2011). *Evaluation matchboard*. Enschede: SLO. Available at http://leerplanevaluatie.slo.nl/english

Plomp, T. (2013). Educational design research: An introduction. In T. Plomp & N. Nieveen (Eds.), *Educational design research - Part A: An Introduction* (pp. 10-51). Enschede, the Netherlands: SLO.

Reeves, T. (2000). *Enhancing the worth of instructional technology research through "design experiments" and other developmental strategies*. Retrieved Oct. 20, 2006 from http://it.coe.uga.edu/~treeves/AERA2000Reeves.pdf

Rossi, P. H., Lipsey, M.W., & Freeman, H. E. (2004). *Evaluation: A systematic approach.* Thousand Oaks, CA: Sage.

Scriven, M. (1967). The methodology of evaluation. In R.W. Tyler, R.M. Gagné, & M. Scriven (Eds.), *Perspectives of curriculum evaluation*. AERA Monograph series on curriculum evaluation. nr.1. Chicago, MI: Rand McNally.

Scriven, M. (1991). Beyond formative and summative evaluation. In M.W. McLaughlin, & D.C. Phillips (Eds.), *Evaluation and education: At quarter century* (pp. 19-64). Chicago: University of Chicago.

Tessmer, M. (1993). *Planning and conducting formative evaluations: Improving the quality of education and training.* London: Kogan Page.

Van den Akker, J. (1999). Principles and methods of development research. In J. van den Akker, R. Branch, K. Gustafson, N. Nieveen, & T. Plomp (Eds.), *Design approaches and tools in education and training* (pp. 1-15). Dordrecht, the Netherlands: Kluwer Academic Publishers.

Wayne, A.J., Yoon, K.S., Zhu, P., Cronen, S., & Garet, M.S. (2008). Experimenting with teacher professional development: Motives and methods. *Educational Researcher, 37*(8), 469-479.

Weston, C., McAlpine, L., & Bordonaro, T. (1995). A model for understanding formative evaluation in instructional design. *Educational training research & technology, 43*(3), 29-48.

7. References and Sources on Educational Design Research

Tjeerd Plomp and Nienke Nieveen

Introduction

This bibliography has been compiled to support researchers and graduate students in getting access to key publications on design research. We do not claim that the selection of sources included in this chapter is complete – it is coloured by our background and bias as well as our knowledge and familiarity with publications. Important criteria for us to include titles in this bibliography are (i) proven usefulness of sources for our own work, and (ii) representing important perspectives and groups that are (or have been) actively working in this domain. We have also asked a few colleagues for suggestions.

In the first section we present an overview of relevant sources available. In the following two sections we point the reader to selected journal articles and book chapters on the concept and methodology of design research and on design research in domains such as curriculum, instructional technology, and the learning of reading and writing, mathematics and science. In the final section, we list references (and URLs) of a number of doctoral theses utilizing design research as a research approach.
As stated, our selection is coloured by our bias and experience, but all these publications refer to a wide range of writings on design research and we trust that they therefore serve as a useful introduction to the reader.

Overview of sources

This section presents titles and references to various special issues of journals and books that have been published about design(-based) research. Besides a number of websites will be listed.

But first reference is made to the sources presented in Part B of this book comprising of 51 cases of successful design research.

Selected key sources on design research listed in Part B of this book

All chapters in Part B: Illustrative Cases (Plomp & Nieveen, 2013) of this book present a specific case of design research in a way that a reader can understand how the research has been designed and conducted. As we could imagine that a reader with a specific interest in

the theme and topic of that particular chapter would like to know about the details of the research reported, we asked all authors to make a clear distinction between the references at the end of their chapter and one or a few key sources specifically relevant for the research reported in the chapter. There were no restrictions given for these key sources – so that these can be e.g. the URL of a research report or doctoral thesis, references to articles in journals, chapters in a book, and the like.
As a result, each of the chapters in Part B at **http://international.slo.nl/edr** presents at the end a few *Key Sources for the Research Reported* and the general *References* for the research reported.

Special issues of journals

In this section we give an overview of a number of special issues that have been published early this century on design(-based) research.

• Educational Researcher 32(1), January/February 2003

Special issue prepared by A.E. Kelly, including:

- Kelly, A.E. (2003). Theme Issue: The role of design in educational research. *Educational Researcher, 32*, 3-4.
- The Design-Based Research Collective (2003). An emerging paradigm for educational inquiry. *Educational Researcher, 32*, 5-8.
- Cobb, P., Confrey, J., diSessa, A., Lehrer, R., & Schauble, L. (2003). Design experiments in educational research. *Educational Researcher, 32*, 9-13.
- McCandliss, B.D., Kalchman, M., & Bryant, P. (2003). Design experiments and laboratory approaches to learning: Steps toward collaborative exchange. *Educational Researcher, 32*, 14-16.
- Lobato, J. (2003). How design experiments can inform a rethinking of transfer and vice versa. *Educational Researcher, 32,* 17-20.
- Bannan-Ritland, B. (2003). The role of design in research: The integrative learning design framework. *Educational Researcher, 32,* 21-24.
- Shavelson, R.J., Phillips, D.C., Towne, L., & Feuer, M.J. (2003). On the science of education design studies. *Educational Researcher, 32,* 25-28.
- Sloane, F.C, & Gorard, S. (2003). Exploring modeling aspects of design experiments. *Educational Researcher, 32,* 29-31.
- Zaritsky, R., Kelly, A.E., Flowers, W., Rogers, E., & O'Neill, P. (2003). Clinical design sciences: A view from sister design efforts. *Educational Researcher, 32*, 32-34.

• Journal of the Learning Sciences, 13(1), 2004

Special issue, including:

- Barab, S., & Squire, K. (2004). Design-based research: Putting a stake in the ground. *Journal of the Learning Sciences, 13*(1), 1-14.

- Collins, A., Joseph, D., & Bielaczyc, K. (2004). Design research: Theoretical and methodological issues. *Journal of the Learning Sciences, 13(*1), 15-42.
- Fishman, B., Marx, R.W., Blumenfeld, P., Krajcik, J., & Soloway, E. (2004). Creating a framework for research on systemic technology innovations. *Journal of the Learning Sciences, 13*(1), 43-76.
- diSessa, A.A., & Cobb, P. (2004). Ontological innovation and the role of theory in design experiments. *Journal of the Learning Sciences, 13*(1), 77-103
- Dede, C. (2004). If design-based research is the answer, what is the question? A commentary on Collins, Joseph, and Bielaczyc; diSessa and Cobb; and Fishman, Marx, Blumenthal, Krajcik, and Soloway in the JLS special issue on design-based research. *Journal of the Learning Sciences, 13*(1), 105-114.
- Kelly, A. (2004). Design research in education: Yes, but is it methodological? *Journal of the Learning Sciences, 13(*1), 115-128.

• Educational Psychologist, 39(4), 2004

Special issue 'Design-based research methods for studying learning in context', edited by W. Sandoval & P. Bell, including:

- Sandoval, W.A., & Bell, P. L. (2004). Design-based research methods for studying learning in context: Introduction. *Educational Psychologist, 39*(4), 199-201.
- Hoadley, C. (2004). Methodological alignment in design-based research. *Educational Psychologist, 39*(4), 203-212.
- Sandoval, W.A. (2004). Developing learning theory by refining conjectures embodied in educational designs. *Educational Psychologist, 39*(4), 213-223.
- Tabak, I. (2004). Reconstructing context: Negotiating the tension between exogenous and endogenous educational design. *Educational Psychologist, 39*(4), 225-233.
- Joseph, D. (2004). The practice of design-based research: Uncovering the interplay between design, research, and the real-world context. *Educational Psychologist, 39*(4), 235-242.
- Bell, P.L. (2004). On the theoretical breadth of design-based research in education. *Educational Psychologist, 39*(4), 243-253.
- Also includes commentary by Angela O'Donnell.

• Educational Technology, 45(1), 2005

Special issue prepared by C. Dede, including:

- Dede, C. (2005). Why design-based research is both important and difficult. *Educational Technology, 45*(1), 5-8.
- Squire, K.D. (2005). Resuscitating research in educational technology: Using game-based learning research as a lens for looking at design-based research. *Educational Technology, 45*(1), 8-14.
- Barab, S.A., Arici, A., & Jackson, C. (2005). Eat your vegetables and do your homework: A design-based investigation of enjoyment and meaning in learning. *Educational Technology, 45*(1), 15–21.

- Nelson, B., Ketelhut, D.J., Clarke, J., Bowman, C., & Dede, C. (2005). Design-based research strategies for developing a scientific inquiry curriculum in a multiuser virtual environment. *Educational Technology, 45*(1), 21–28.
- Kafai, Y.B. (2005). The classroom as "living laboratory": Design-based research for understanding, comparing, and evaluating learning science through design. *Educational Technology, 45*(1), 28–34.
- Hay, K.E., Kim, B., & Roy, T.C. (2005). Design-based research: More than formative assessment? An account of the Virtual Solar System Project. *Educational Technology*, *45*(1), 34-41.
- Hoadley, C. (2005). Design-based research methods and theory building: A case study of research with SpeakEasy. *Educational Technology, 45*(1), 42-47.
- Reeves, T.C. (2005). Design-based research in educational technology: Progress made, challenges remain. *Educational Technology, 45*(1), 48-52

Books

• Van den Akker, J., Gravemeijer, K, McKenney, S.,& Nieveen, N. (Eds). (2006). *Educational design research*. London: Routledge. ISBN10: 0-415-39635-2 (pbk) (163 pages)
Available at http://www.taylorandfrancis.co.uk/shopping_cart/products/product_detail.asp?sku=&ppid=118302&isbn=9780415396356
This book comprises the papers presented at a seminar organized by the Netherlands Organization for Scientific Research, in particular by the Program Council for Educational Research. The seminar, conducted in December 2003, has been a meeting place of design researchers from the USA and The Netherlands. The book reflects the various angles from which researchers in the domains of curriculum, instructional technology and (mathematics and science) education address the need to develop research based solutions (interventions) to problems for which no guidelines to solutions are available. The book illustrates that authors with various backgrounds have clearly a common ground when reflecting on design research as a research approach. The book has four parts:
Part 1. What and why
1. Introducing educational design research - Jan van den Akker, Koeno Gravemeijer, Susan McKenney, Nienke Nieveen
2. Toward productive design studies - Decker Walker

Part 2. Examples from the field
3. Design research from the learning design perspective - Koeno Gravemeijer, Paul Cobb
4. Design research from the technology perspective - Thomas Reeves
5. Design research from a curriculum perspective - Susan McKenney, Nienke Nieveen, Jan van den Akker

Part 3. Quality

6. Assessing the quality of design research proposals: Some philosophical perspectives - D.C. Phillips
7. Balancing innovation and risk: Assessing design research proposals - Daniel C. Edelson
8. Quality Cciteria for design research: Evidence and commitments - Anthony E. Kelly

Part 4. Moving ahead

9. From design research to large-scale impact: Engineering research in education – Hugh Burkhardt
10. Educational design research: The value of variety - Nienke Nieveen, Susan McKenney, Jan van den Akker

- Kelly, A.E., Lesh, R.A. & Baek, J.Y. (Eds). (2008). *Handbook of design research methods in education innovations in science, technology, engineering, and mathematics learning and teaching*. New York: Lawrence Erlbaum Associates. ISBN: 978-0-8058-6059-7 (pbk) (560 pages)

  Available at http://www.routledgeeducation.com/books/Handbook-of-Design-Research-Methods-in-Education-isbn9780805860597

  The announcement of the book states that the handbook presents the latest thinking and current examples of design research in education. Design-based research involves introducing innovations into real-world practices (as opposed to constrained laboratory contexts) and examining the impact of those designs on the learning process. Designed prototype applications (e.g., instructional methods, software or materials) and the research findings are then cycled back into the next iteration of the design innovation in order to build evidence of the particular theories being researched, and to positively impact practice and the diffusion of the innovation.

  The Handbook of Design Research Methods in Education is meant to fill a need in how to conduct design research by those doing so right now. The chapters represent a broad array of interpretations and examples of how today's design researchers conceptualize this emergent methodology across areas as diverse as educational leadership, diffusion of innovations, complexity theory, and curriculum research.

  The handbook has eight sections:

  - Design research and its argumentative grammar
  - Modeling student learning during design research
  - Modeling teacher learning using design research
  - Modeling stakeholders commitments using design research
  - Reflecting on design research at the project level
  - Reflecting on design research at the program level
  - Extending design research methodologically
  - Tracking the diffusion of design research.

- McKenney, S., & Reeves, T.C. (2012). *Conducting educational design research.* London / New York: Routledge. ISBN: 978-0-415-61804-5 (pbk) (244 pages). Also as hardcover and e-book.
  This book integrates multiple perspectives of educational design research throughout this three-part book.
  Part I Foundations clarifies the educational design research origins, approach and outcomes. It also presents a generic model. Chapters:
  1. About educational design research
  2. Contributions to theory and practice: concepts and examples
  3. Toward a generic model for educational design research.

  Part II Core Processes discusses the constituent elements of the model in detail. Chapters:
  4. Analysis and exploration
  5. Design and construction
  6. Evaluation and reflection
  7. Implementation and spread.

  Part III Moving forward offers recommendations for proposing, reporting and advancing educational design research. Chapters:
  8. Writing proposals for educational design research
  9. Reporting educational design research
  10. Looking back and looking ahead.

- Reinking, D., & Bradley, B.A. (2008). *On formative and design experiments: Approaches to language and literacy research.* New York & London: Teachers College, Columbia University. ISBN: 978-0-8077-4841-1 (pbk) (134 pages)
  This booklet provides a nice introduction into formative and design experiments, a term synonymous for what we call design research and others design-based research. It provides a thorough, but practical and useful overview of design research addressing the following questions:
  - What are formative and design experiments? (Ch1)
  - What are the methods of formative and design experiments? (Ch2)
  - What are some good examples of formative and design experiments? (Ch3)
  - Is there a formative or design experiment in your future? (Ch4)

- Richey, R., & Klein, J.D. (2007). *Design and development research: Methods, strategies, and issues*. London: Routledge. ISBN 080585732X, 9780805857320 (180 pages)
  This volume discusses methods and strategies appropriate for conducting design and development research. Rich with examples and explanations, the book describes actual strategies that researchers have used to conduct two major types of design and development research: 1) product and tool research and 2) model research. Common challenges confronted by researchers in the field when planning and conducting a study

are explored and procedural explanations are supported by a wide variety of examples taken from current literature.

- Fishman, B.J., Penuel, W. R., Allen, A.-R., & Cheng, B.H. (Eds.). (In press). Design-based implementation research: Theories, methods, and exemplars. National Society for the Study of Education Yearbook. New York, NY: Teachers College Record.
  From the book announcement:
  "Design-Based Implementation Research applies design-based perspectives and methods to address and study problems of implementation...DBIR challenges education researchers to break down barriers between sub-disciplines of educational research that isolate those who design and study innovations within classrooms from those who study the diffusion of innovations."
  This book will have five sections with in total 15 chapters:
  Section 1 – Introduction to DBIR
  1. Design-based implementation research: An emerging model for transforming the relationship of research and practice
  2. Theories and research methodologies for design-based implementation research: Examples from four cases

  Section 2 – Taking a cross-setting perspective in DBIR
  3. Taking a societal sector perspective on youth learning and development
  4. Adaptation by design: A context-sensitive, dialogic approach to interventions

  Section 3 – Designing across levels for DBIR
  5. Negotiating problems of practice in research-practice design partnerships
  6. Beyond the policy memo: Designing to strengthen the practice of district central office leadership for instructional improvement at scale
  7. Supporting teachers in schools to improve their instructional practice
  8. Designing for productive adaptations of curriculum interventions

  Section 4 – Forms of evidence in DBIR
  9. Design research with educational systems: Investigating and supporting improvements in the quality of mathematics teaching and learning at scale
  10. Towards an evidence framework for design-based implementation research
  11. Situated research design and methodological choices in formative program evaluation

  Section 5 – Infrastructures in support of DBIR
  12. The SERP approach to problem-solving research, development, and implementation
  13 Beyond classrooms: Scaling and sustaining instructional innovations
  14 More than a network: Building professional communities for educational improvement
  15 Empowering DBIR: The need for infrastructure

Websites

Of the many websites for design research or design-based research, we mention a few, as through these ones access can be obtained to many other websites and sources.

• **http://international.slo.nl/edr** (checked December 19, 2013)
This website has links to a number of publications related to this book published by SLO – Netherlands Institute for Curriculum Development, Enschede (The Netherlands), namely:

- *Educational Design Research – Part A: Introduction.*
- *Educational Design Research – Part B: Illustrative cases*; this is a collection of 51 cases of successful design research. Each case chapter is a separate pdf file allowing the reader making his own purposive selection
- Table of Content of Part B with a 'portrait' per case
- Case selection tool to assist in selecting appropriate cases of design research.

• **http://edutechwiki.unige.ch/en/Design-based_research** (checked on 25 August 2013; on that date last modified 25 April 2013)
This website defines design-based research (DBR), discusses characteristics and present example approaches. For methodological aspects, it refers to a 'Methodology tutorial - design-oriented research designs' (created by Daniel K. Schneider, 2008). An useful section is 'Writing up a DBR' with a number of suggestions for structuring how to report DBR. This website presents also a number links to other sites about the topic. It has an extensive list of references with authors mainly from North-America (one of the exceptions is Van den Akker, Graveijer, McKenney & Nieveen, 2006).

• **http://en.wikipedia.org/wiki/Design-based_research** (checked on 25 August 2013; on that date last modified on 25 August 2013)
This website is a nice introduction, but has an exclusive North-American orientation – no references seem to be included about design (based) research in other parts of the world.

In the first edition (2009) of this book, we mentioned the following websites that are still frequently referred to:

• **http://projects.coe.uga.edu/dbr/index.htm** (last update November 2006)
titled 'Design-based Research EPSS'– created by Instructional Technology Ph.D. students at The University of Georgia under supervision of Tom Reeves (comprehensive till last update of November 2006).

• **http://cider.athabascau.ca/CIDERSIGs/DesignBasedSIG/**
CIDER is the Centre for Distance Education Research at Athabasca University, Canada's Open University (Edmonton, Alberta, Canada). CIDER has special interest groups (SIGs)

and this is the website of the Design-Based Research SIG. It has, amongst others, a link to a bibliography of DBR drawn up by Terry Anderson
Anderson calls it a snapshot of most current (early 2005) literature related to discussion, exploration and examples of design-based research. The references are presented with URLs (if available) along with abstracts and occasionally quotations or annotations by Anderson. Has overlap with the University of Georgia website.

• **http://www.designbasedresearch.org/index.html** (last update not clear, but no references later than 2004)
This is the website of the Design-Based Research Collective, a small group of researchers who engage in design-based research, often in technology enhanced learning environments. It contains references of a number of publications, as well as a number of links to relevant related websites.

Selected journal articles and book chapters

Apart from the sources mentioned above, many articles and book chapters have been published dealing with conceptual and/or methodological aspects of design research, or reporting about design research projects. Many of these references (plus abstracts) can be found on the websites mentioned in this section, but we have selected a number which are summarized in the final section of this chapter.

Selected journal articles and book chapters on the concept and methodology of design research

There are so many publications on educational design research that it is impossible to draw up a comprehensive bibliography.
However we want to point the reader to a number of articles and chapter that have helped us to get involved in design research and to understand the main issues in our field. Given this rationale for selecting these titles, the reader will find that some of the titles are also referred to websites mentioned above.

- Aken, J. van (2013). Design Science: Valid knowledge for socio-technical system design. Accepted for publication in the *Proceedings of the European Design Science Symposium 2012* to be published by Springer Verlag AG.
  Abstract: This article shows how one can develop design science, i.e. valid knowledge produced by rigorous research to support designing, for the social world. The nature of the aforementioned methodological problem will be discussed, followed by a presentation of a research strategy by which one can overcome this problem. This strategy, objective and systematic social experimental learning, will be discussed and will be illustrated by some examples from the field of organization and management.

Finally some suggestions are given on the development of design science for the social components of information systems.

- Akker, J. van den (1999). Principles and methods of development research[1]. In J. van den Akker J.J.H. van den Akker, R. Branch, K. Gustafson, N.M. Nieveen, & Tj. Plomp. (Eds.), *Design approaches and tools in education and training* (pp. 1-14). Dordrecht: Kluwer Academic Publishers.
  Abstract: This chapter discusses the role of research in relation to educational design and development activities. The first part of the chapter focuses on the rationale and basic principles of development research by outlining motives for conducting formative research, analyzing definitions and aims of various types of development research, and discussing several of its key characteristics. The second part of the chapter deals with methods of development research, exploring some of its typical problems and dilemmas, and discussing several challenges for further action and reflection.

- Akker, J. van den, & Plomp, Tj. (1993). *Development research in curriculum: propositions and experiences.* Paper presented at the annual conference of the American Educational Research Association, April 1993, Atlanta (GA, USA).
  Abstract: Reason to include this paper is that it is the first paper from the group at the University of Twente on what they called at that time development research. Based on the assertion that both curriculum development and curriculum research have much relevance to gain from a close liaison, the authors suggest that boundaries between the two should fade, which can be done in a new research strategy called development research. The paper presents the purpose, a conceptual framework and some characteristics of development research in curriculum.
  See: https://instrumenten.slo.nl/leerplanevaluatie/Documents/vandenAkker_Plomp_1993.pdf

- Anderson, T., & Shattuck, J. (2010). Design-based research: A decade of progress in education research? *Educational Researcher, 41*(1), 16-25.
  Note: this article should be read in combination with McKenney & Reeves (2013) – see below.
  Abstract: Design-based research (DBR) evolved near the beginning of the 21st century and was heralded as a practical research methodology that could effectively bridge the chasm between research and practice in formal education. In this article, the authors review the characteristics of DBR and analyze the five most cited DBR articles from each year of this past decade. They illustrate the context, publications, and most popular

1) The concept development research, used in some titles, is synonymous to design research.

interventions utilized. They conclude that interest in DBR is increasing and that results provide limited evidence for guarded optimism that the methodology is meeting its promised benefits.

- Bannan-Ritland, B. (2003). The role of design in research: The integrative learning design framework. *Educational Researcher, 32(*1), 21-24.
  Abstract (from UGA website): In this article, a general model is proposed for design research in education that grows out of the author's research and work in related design fields. The model emphasizes the stage sensitivity of (a) research questions, (b) data and methods, and (c) the need for researchers to design artifacts, processes, and analyses at earlier stages in their research that can then be profitably used (perhaps by different researchers) in later stages.

- Barab, S.A., & Squire, K.D. (2004). Design-based research: Putting a stake in the ground. *Journal of the Learning Sciences, 13*(1), 1-14.
  Abstract (from UGA website): The article highlights and problematizes some challenges that are faced in carrying out design-based research. It states that the emerging field of learning sciences is one that is interdisciplinary, drawing on multiple theoretical perspectives and research paradigms so as to build understandings of the nature and conditions of learning, cognition and development. A fundamental assumption of many learning scientists is that cognition is not a thing located within the individual thinker but is a process that is distributed across the knower, the environment in which knowing occurs and the activity in which the learner participates. In other words, learning, cognition, knowing and context are irreducibly co-constituted and cannot be treated as isolated entities or processes.

- Brown, A.L. (1992). Design experiments: Theoretical and methodological challenges in creating complex interventions in classroom settings. *Journal of the Learning Sciences, 2*, 141-178.
  Abstract (from UGA website): This is the seminal article on design research. Discusses theoretical and methodological challenges in creating complex interventions in classroom settings. Movement from the classical psychological position of concentrating on a theoretical study of the learning processes of individual students to a concentration on conceptual change in teachers and students; classroom restructuring; design experiments; experiences on learning theory.

- Design-Based Research Collective. (2003). Design-based research: An emerging paradigm for educational inquiry. *Educational Researcher, 32*(1), 5-8.
  Abstract (from UGA website): The authors argue that design-based research, which blends empirical educational research with the theory-driven design of learning

environments, is an important methodology for understanding how, when, and why educational innovations work in practice. Design based researchers' innovations embody specific theoretical claims about teaching and learning, and help us understand the relationships among educational theory, designed artifact, and practice. Design is central in efforts to foster learning, create usable knowledge, and advance theories of learning and teaching in complex settings. Design based research also may contribute to the growth of human capacity for subsequent educational reform.

- Eilks, I., & Ralle, B. (2002). Participatory Action Research in chemical education. In B. Ralle, & I. Eilks (Eds.), *Research in Chemical Education - What does this mean?* (pp. 87-98). Aachen: Shaker.
  (http://www.idn.uni-bremen.de/chemiedidaktik/material/Symp 2002 Eilks Ralle PAR.PDF)
  Abstract: Participatory Action Research is recommended as a method for conducting research within chemical education. This can provide research on curriculum development and improved teaching strategies with a well-founded methodological framework. The objective is to establish an accepted methodological foundation for education research, to fit curriculum development better to the needs of practice, to ensure that research is of value for practical use, and thus to close the gaps between curriculum development, empirical research and teaching practice. The main aspects of the research strategy are discussed here. In addition, initial experiences that were made using this method are presented. They refer to a project that was designed to develop new, more efficient approaches to teaching the particulate nature of matter in lower secondary school chemistry.

- Kelly, A.E. (2006). Quality criteria for design research. In: J. van den Akker, K. Gravemeijer, S. McKenney, & N. Nieveen (Eds.), *Educational design research* (pp. 107-119). London: Routledge.
  Abstract: this chapter discusses for each of three different uses for design research in education a number of characteristics and exemplary examples. It introduces the notion of the commissive space of design research, meaning that (amongst other characteristics) design research does not strive for context-free claims but sees contexts as central to its conceptual domain, that design research is experimental but not an experiment, and that design researchers choose to work in the "context of discovery", rather than in the "context of verification" utilizing randomized trials.

- McKenney, S., & Reeves, T.C. (2013). Systematic review of design-based research progress: Is a little knowledge a dangerous thing? *Educational Researcher, 42*(2), 97-100.
  Note: this article should be read in combination with Anderson & Shattuck (2010) – see above.

Abstract: Sufficient attention and resources have been allocated to design-based research (DBR) to warrant review concerning if and how its potential has been realized. Because the DBR literature clearly indicates that this type of research strives toward both the development of an intervention to address a problem in practice and empirical investigation yielding theoretical understanding that can inform the work of others, thoughtful assessment of DBR progress must devote substantial attention to each of these aspects. This requires an in-depth analysis of full-text reports of DBR, framed by a refined conceptualization of the intended outputs of DBR, and ideally, complemented by empirical investigation involving design-based research participants directly.

- Penuel, W.R., Fishman, B.J., Cheng, B.H., & Sabelli, N. (2011). Organizing research and development at the intersection of learning, implementation, and design. *Educational Researcher, 40*(7), 331-337.
  Abstract: This article describes elements of an approach to research and development called *design-based implementation research.* The approach represents an expansion of design research, which typically focuses on classrooms, to include development and testing of innovations that foster alignment and coordination of supports for improving teaching and learning. As in policy research, implementation is a key focus of theoretical development and analysis. What distinguishes this approach from both traditional design research and policy research is the presence of four key elements: (a) a focus on persistent problems of practice from multiple stakeholders' perspectives; (b) a commitment to iterative, collaborative design; (c) a concern with developing theory related to both classroom learning and implementation through systematic inquiry; and (d) concern with developing capacity for sustaining change in systems.

- Reeves, T. (2000). *Enhancing the worth of instructional technology research through "design experiments" and other developmental strategies.* Paper presented at the Annual Meeting of the American Educational Research Association, April 2000, New Orleans (LA, USA).
  Retrieved Oct. 20, 2006 from http://it.coe.uga.edu/~treeves/AERA2000Reeves.pdf
  Abstract: The author argues that in general research in the area of instructional technology is poor, not providing practitioners with sufficient guidance. He discusses various types of instructional technology research goals and methods and suggests that 'use-inspired basic research' is needed in the domain of instructional technology referring to approaches like development research and design experiments. He presents a framework and characteristics for development research in the area of instructional technology.

Selected journal articles and book chapters on design research in domains

Over the last few years, increasingly examples of design research have been published. This section contains just a few exemplary references to articles and chapters in books of design research in various domains, of which a few are taken from the UGA website

Domain of mathematics education

Some references for design research cases in the domain of mathematics education conducted in the USA that were included in the 2009-edition are:

Bowers, J.S., Cobb, P., & McClain, K. (1999). The evolution of mathematical practices: A case study. *Cognition and Instruction, 17*, 25-64.

Cobb, P. (1999). Individual and collective mathematical learning: The case of statistical data analysis. *Mathematical Thinking and Learning, 1*, 5-44.

Cobb, P., McClain, K., & Gravemeijer, K. (2003). Learning about statistical covariation. *Cognition and Instruction, 21*, 1-78.

Confrey, J., & Smith, E. (1995). Splitting, covariation, and their role in the development of exponential functions. *Journal for Research in Mathematics Education, 26*, 66-86.

Lehrer, R., & Schauble, L. (2004). Modeling natural variation through distribution. *American Educational Research Journal, 41*, 635-679.

Lobato, J. (2003). How design experiments can inform a rethinking of transfer and vice versa. *Educational Researcher, 32*(1), 17-20.

Simon, M.A. (1995). Reconstructing mathematics pedagogy from a constructivist perspective. *Journal for Research in Mathematics Education, 26*, 114-145.

Stephan, M., Bowers, J., & Cobb, P. (Eds.). (2003). Supporting students' development of measuring conceptions: Analyzing students' learning in social context. *Journal for Research in Mathematics Education Monograph No. 12.* Reston, VA: National Council of Teachers of Mathematics.

Further:

- Cobb, P., & Gravemeijer, K. (2008). Experimenting to support and understand learning processes. In A.E. Kelly, R.A. Lesh, & J.Y. Baek (Eds.), *Handbook of design research methods in education. Innovations in science, technology, engineering and mathematics learning and teaching* (pp. 68-95). New York: Lawrence Erlbaum Associates.
  Abstract: In this chapter, we describe an approach to design research that we have refined while conducting a series of design research projects in mathematics education over a 10-year period. Our intent in doing so is to highlight a number of issues that we believe are essential to consider when conducting a design experiment regardless of the specific approach followed. For the purpose of this chapter, we define design research as a family of methodological approaches in which instructional design and research are interdependent. On the one hand, the design of learning environments serves as

the context for research, and, on the other hand, ongoing and retrospective analyses are conducted in order to inform the improvement of the design. This type of research involves attempting to support the development of particular forms of learning and studying the learning that occurs in these designed settings. In each of these cases, design research enables us to investigate simultaneously both the process of learning and the means by which it is supported and organized. We focus specifically on design experiments in classrooms in which a research team assumes responsibility for a group of students' learning both because they are the most common type of design research and because most of our work has involved experiments in classrooms. We discuss the three phases of conducting a design experiment: preparing for the experiment, experimenting to support learning, and conducting retrospective analyses of the data generated during the course of the experiment.

- Doorman, M., Drijvers, P., Gravemeijer, K., Boon, P., & Reed, H. (2012). Tool use and the development of the function concept: from repeated calculations to functional thinking. *International Journal of Science and Mathematics Education, 10*(6), 1243-1267.
  Abstract: The concept of function is a central but difficult topic in secondary school mathematics curricula, which encompasses a transition from an operational to a structural view. The question in this paper is how to design and evaluate a technology-rich learning arrangement that may foster this transition. With domain-specific pedagogical knowledge on the learning of function as a starting point, and the notions of emergent modeling and instrumentation as design heuristics, such a learning arrangement was designed for grade 8 students and field tested. The results suggest that these design heuristics provide fruitful guidelines for the design of both a hypothetical learning trajectory and concrete tasks, and can be generalized to other design processes.

- Gravemeijer, K., & van Eerde, D. (2009) Design research as a means for building a knowledge base for teachers and teaching in mathematics education. *Elementary School Journal, 109*(5), p.p. 510-524.
  Abstract: Building on the claim that teachers need to know how innovative instructional approaches work to be able to adapt them to their own classrooms, design research is presented as a research method that aims to offer exactly that kind of information. We elaborate design research that aims to develop a local instruction theory—a theory about the process by which students learn a given topic in mathematics and theories about the means of support for that learning process. We will illustrate this with the example of design research on a local instruction theory on addition and subtraction to 100. We further discuss design research that combines the two goals of teacher learning and student learning in one project as a special case of research on teacher learning. In closing, we briefly look into the relation between design research and teacher research.

- An illustrative example of design research in the *context of a developing country* is: Vos, P., Devesse, T.G., & Pinto, A.A.R. (2007). Designing mathematics lessons in Mozambique: Starting from authentic resources. *African Journal of Research in SMT Education, 11*(2), pp. 51-66
  Abstract: This article describes research on the design of student-centred instruction in Mozambique. The starting point was the use of real-life resources, such as traditional art craft objects and authentic newspaper clippings. The research was based on an instructional design model, which attempts to align theory with practice and which is geared towards improving practice. In two parallel studies, one on geometry and one on statistics, student-centred instruction was facilitated through the use of worksheets with open-ended questions tailored for group work. In a cyclic process, the prototype materials and the associated instructional method were formatively evaluated. The evaluations showed that the designs were useful even in classrooms packed with more than sixty students.

Domain of science education

- Burmeister, M., & Eilks, I. (2013). Using participatory action research to develop a course module on education for sustainable development in pre-service chemistry teacher education. *Centre of Educational Policy Studies Journal, 3*(1), 59-78.
  Abstract: This paper describes the development of a course module on sustainability issues and Education for Sustainable Development in German pre-service chemistry teacher education. The module was inspired by empirical research findings about the knowledge base of student teachers. It was created and cyclically refined using Participatory Action Research. Experience gained during its three-year application will be reflected upon here, including feedback collected from student evaluation sheets. In the end, the participants responded extremely positively to the course. The student teachers stated that the module was interesting, relevant and valuable for their later profession as high school chemistry teachers. They also emphasised that they now felt more competent in the area of sustainability and ESD.

- Hoadley, C.M., & Linn, M. C. (2000) Teaching science through online, peer discussions: SpeakEasy in the knowledge integration environment. *International Journal of Science Education, 22*(8), 839-857.
  Abstract: This article discusses whether students can learn science from carefully designed online peer discussions. Contrasts two formats of contributed comments–historical debate and narrative text–and assesses the impact of an asynchronous discussion on student understanding of the nature of light. It also reports that students gain integrated understanding of the nature of color from both discussion formats.

- Höttecke, D., Henke, A., & Ries, F. (2012). Implementing history and philosophy in science teaching: Strategies, methods, results and experiences from the European HIPST project. *Science & Education, 15*(1), 1233-1261.
  Abstract: The paper presents the results of a symbiotic developmental project called HIPST (History and Philosophy In Science Teaching) where researchers from physics education and physics teachers collaborated. The project benefited from the special skills, creative potentials and experiences of teachers and their capacities for evaluating the materials designed through the collaborative process. The project let to a series of historical case studies for teaching and learning physics with its history. Teaching methods comprise student-centered activities as creative writing, the use of replications of historical apparatus and new ways for explicitly and reflectively addressing the nature of science.

- Juuti, K., & Lavonen, J. (2006). Design-based research in science education: one step towards methodology. *Nordina*, *4*(06), 54–68.
  Abstract: Recently, there has been critiques towards science education research, as the potential of this research has not been actualised in science teaching and learning praxis. The paper describes an analysis of a design-based research approach (DBR) that has been suggested as a solution for the discontinuation between science education research and praxis. We propose that a pragmatic frame helps to clarify well the design-based research endeavour. We abstracted three aspects from the analysis that constitute design-based research: (a) a design process is essentially iterative starting from the recognition of the change of the environment of praxis, (b) it generates a widely usable artefact, (c) and it provides educational knowledge for more intelligible praxis. In the knowledge acquisition process, the pragmatic viewpoint emphasises the role of a teacher's reflected actions as well as the researches' involvement in the authentic teaching and learning settings.

- Kafai, Y.B., & Ching, C.C. (2001). Affordances of collaborative software design planning for elementary students' science talk. *Journal of the Learning Sciences, 10*(3), 323-363.
  Abstract: This article investigates whether science permeates the design environment and is thus contexted within the other activities of collaborative management and technology. Focuses on which contexts gave rise to science talk. Studies a classroom with (n=33) students divided into seven teams

- Knippels, M.C.P.J., Waarlo, A.J., & Boersma, K.Th. (2005). Design criteria for learning and teaching genetics. *Journal of Biological Education, 39*(3), 108-112.
  Abstract: While learning and teaching difficulties in genetics have been abundantly explored and described, there has been less focus on the development and field-testing

of strategies to address them. To inform the design of such a strategy a review study, focus group interviews with teachers, a case study of a traditional series of genetics lessons, student interviews, and content analysis of school genetics teaching were carried out. Specific difficulties reported in the literature were comparable to those perceived by Dutch teachers and found in the case study and the student interviews.The problems associated with the abstract and complex nature of genetics were studied in more detail. The separation of inheritance, reproduction and meiosis in the curriculum accounts for the abstract nature of genetics, while the different levels of biological organisation contribute to its complex nature. Finally, four design criteria are defined for a learning and teaching strategy to address these problems: linking the levels of organism, cell and molecule; explicitly connecting meiosis and inheritance; distinguishing the somatic and germ cell line in the context of the life cycle; and an active exploration of the relations between the levels of organisation by the students.
Key words: Biology education; Genetics; Learning and teaching difficulties; Design criteria

- Kock, Z.D.Q.P., Taconis, R., Bolhuis, S.M, & Gravemeijer, K.P.E. (2013). Some key issues in creating inquiry-based instructional practices that aim at the understanding of simple electric circuits. *Research in Science Education*, *43*(2), 579-597.
Abstract: Many students in secondary schools consider the sciences difficult and unattractive. This applies to physics in particular, a subject in which students attempt to learn and understand numerous theoretical concepts, often without much success. A case in point is the understanding of the concepts current, voltage and resistance in simple electric circuits. In response to these problems, reform initiatives in education strive for a change of the classroom culture, putting emphasis on more authentic contexts and student activities containing elements of inquiry. The challenge then becomes choosing and combining these elements in such a manner that they foster an understanding of theoretical concepts. In this article we reflect on data collected and analyzed from a series of 12 grade 9 physics lessons on simple electric circuits. Drawing from a theoretical framework based on individual (conceptual change based) and socio-cultural views on learning, instruction was designed addressing known conceptual problems and attempting to create a physics (research) culture in the classroom. As the success of the lessons was limited, the focus of the study became to understand which inherent characteristics of inquiry based instruction complicate the process of constructing conceptual understanding. From the analysis of the data collected during the enactment of the lessons three tensions emerged: the tension between open inquiry and student guidance, the tension between students developing their own ideas and getting to know accepted scientific theories, and the tension between fostering scientific interest as part of a scientific research culture and the task oriented school culture. An outlook will be given on the implications for science lessons.

- Lijnse, P.L. (1995). "Developmental research" as a way to an empirically based "didactical structure" of science. *Science Education, 29*(2), 189-199.
  Abstract: The author argues that developmental research (in this book called ´design research') is needed in which small-scale curriculum development is cyclically coupled to in-depth classroom research of teaching-learning processes. Such research should result in worked out examples of successful ways of teaching, according to new conceptual curriculum structures. Designing such `didactical´ structures constitutes a longer term research program, which asks for international exchange and cooperation.

- Marks, R., & Eilks, I. (2010). Research-based development of a lesson plan on shower gels and musk fragrances following a socio-critical and problem-oriented approach to chemistry teaching. *Chemistry Education Research and Practice, 11*(2), 129-141
  Abstract: A case is described of the development of a lesson plan for 10th grade (age range 15-16) chemistry classes on the chemistry of shower gels. The lesson plan was developed within the framework of a Participatory Action Research project. From the accompanying evaluation based on teachers' feedback, written student questionnaires and a study based on students' group discussions, the lesson plan was refined in different cycles of development, testing, evaluation and reflection. In the end, the lesson plan was found to be highly feasible, motivating, and an initiator of intense discussions among pupils. The overall approach seems to be promising for promoting higher-order cognitive skills, *i.e.* reflection and evaluation within the framework of science, technology and society. The article describes Participatory Action Research as a research model and a reflection thereof.

- Penuel, W.R., & Fishman, B.J. (2012). Large-scale science education intervention research we can use. *Journal of Research in Science Teaching, 49*(3), 281-304.
  Abstract: This article develops an argument that the type of intervention research most useful for improving science teaching and learning and leading to scalable interventions includes both research to develop and gather evidence of the efficacy of innovations and a different kind of research, *design-based implementation research* (DBIR). DBIR in education focuses on what is required to bring interventions and knowledge about learning to all students, wherever they might engage in science learning. This research focuses on implementation, both in the development and initial testing of interventions and in the scaling up process. In contrast to traditional intervention research that focuses principally on one level of educational systems, DBIR designs and tests interventions that cross levels and settings of learning, with the aim of investigating and improving the effective implementation of interventions. The article concludes by outlining four areas of DBIR that may improve the likelihood that new standards for science education will achieve their intended purpose of establishing an effective, equitable, and coherent system of opportunities for science learning in the United States.

Domain of reading - writing

- Abbott, S.P., Reed, E., Abbott, R.D., & Berninger, V.W. (1997). Year-long balanced reading/writing tutorial: A design experiment used for dynamic assessment. *Learning Disability Quarterly, 20*(3), 249-263.
  Abstract: Sixteen children with severe reading problems in first grade received a year-long individual tutorial intervention. Growth curve analyses found significant gains on measures of orthographic and phonological coding, word identification, word attack skills, reading comprehension, letter automaticity, and spelling and marginally significant gains in writing composition.

- DeCorte, E., Verschaffel, L., & van de Ven, A. (2001). Improving text comprehension strategies in upper primary school children: A design experiment. *The British Journal of Educational Psychology, 71*, 531-559.
  Abstract: With respect to the acquisition of competence in reading, new standards for primary education stress more than before the importance of learning and teaching cognitive and metacognitive strategies that facilitate text comprehension. Therefore, there is a need to design a research-based instructional approach to strategic reading comprehension. The design experiment aimed at developing, implementing and evaluating a research-based, but also practically applicable learning environment for enhancing skilled strategy use in upper primary school children when reading a text. This design experiment shows that it is possible to foster pupils' use and transfer of strategic reading comprehension skills in regular classrooms by immersing them in a powerful learning environment. But this intervention does not automatically result in improvement of performance on a standardized reading comprehension test.

- Neuman, S.B. (1999). Books make a difference: A study of access to literacy. *Reading Research Quarterly, 34*(3), 286-311.
  Abstract: This article examines the impact of an intervention targeting economically disadvantaged children that flooded over 330 child-care centers with high-quality children's books and provided 10 hours of training to child-care staff. It examines the project's impact and gives support for the physical proximity of books and the psychological support to child-care staff on children's early-literacy development.

Domain of instructional technology

- Bannan-Ritland, B. (2003). The role of design in research: The integrative learning design framework. *Educational Researcher, 32*(1), 21-24.
  Abstract: (from UGA website): In this article, a general model is proposed for design research in education that grows out of the author's research and work in related design fields. The model emphasizes the stage sensitivity of (a) research questions, (b) data and

methods, and (c) the need for researchers to design artifacts, processes, and analyses at earlier stages in their research that can then be profitably used (perhaps by different researchers) in later stages.

- Herrington, J., & Oliver, R. (1997). Multimedia, magic and the way students respond to a situated learning environment. *Australian Journal of Educational Technology, 13*(2), 127-143. Available at: http://www.ascilite.org.au/ajet/ajet13/herrington.html
  Abstract: This article presents a design of an interactive multimedia learning environment entitled Investigating assessment strategies in mathematics classrooms, which represents the operationalized characteristics of situated learning. The authors also suggest the critical guidelines for the design of the multimedia software to enable it to support a situated learning environment. They then report a study that investigates patterns of behavior of students immersed in this multimedia situated learning environment. The findings suggest that the use of the situated learning model is successful in providing guidelines for the development of an interactive multimedia program. They also reveal that in instances where learners are empowered and are enabled to assume higher degrees of responsibility for their activity and conduct in a learning setting, the researchers need to be cognizant of the various design factors which can impede or enhance learning. In multimedia environments, these include such elements as the motivational aspects of the environment, the interface design, and the navigation elements employed. In conclusion, the authors suggest that it is also important to practice research which explores the impact of the more tangible aspects of multimedia design such as those explored in this study.

- Herrington, J., & Oliver, R. (2000). An instructional design framework for authentic learning environments. *Educational Technology Research and Development, 48*(3), 23-48. Available at: http://edserver2.uow.edu.au/~janh/Assessment/Authentic%20Assessment_files/ETR%26D.pdf
  Abstract: The instructional technology community is in the midst of a philosophical shift from a behaviourist to a constructivist framework, a move that may begin to address the growing rift between formal school learning and real-life learning. One theory of learning that has the capacity to promote authentic learning is that of situated learning.
  The purpose of this three part study was firstly, to identify critical characteristics of a situated learning environment from the extensive literature base on the subject; secondly, to operationalise the critical characteristics of a situated learning environment by designing a multimedia program which incorporated the identified characteristics; and thirdly, to investigate students' perceptions of their experiences using an multimedia package based on a situated learning framework.

The learning environment comprised a multimedia program for preservice teachers on assessment in mathematics, together with recommended implementation conditions in the classroom. Eight students were observed and interviewed to explore their perceptions of the situated learning environment. Findings suggest that the use of the situated learning framework appeared to provide effective instructional design guidelines for the design of an environment for the acquisition of advanced knowledge.

- Reeves, T. (2006). Design research from a technology perspective. In: J. van den Akker, K. Gravemeijer, S. McKenney, & N. Nieveen (Eds.), *Educational design research* (pp. 52-67). London: Routledge.
  Abstract: The effectiveness of the field known as educational technology in fundamentally enhancing teaching and learning has increasingly been called into question, as has the efficacy of educational research in general. Doubts about educational technology research primarily stem from decades of an arguably flawed research agenda that has been both pseudoscientific and socially irresposible. It is proposed that progress in improving teaching and learning through technology may be accomplished using design research as an alternative model of inquiry. Design research protocols require intensive and long-term collaboration involving researchers and practitioners. It integrates the development of solutions to practical problems in learning environments with the identification of reusable design principles. Examples of design research endeavors in educational technology are described here. The chapter ends with a call for the educational technology research community to adopt design research methods more widely.

- Reeves, T. C., Herrington, J., & Oliver, R. (2004). A development research agenda for online collaborative learning. *Educational Technology Research and Development, 52*(4), 53-65.
  Abstract: Although important, traditional basic-to-applied research methods have provided an insufficient basis for advancing the design and implementation of innovative collaborative learning environments. It is proposed that more progress may be accomplished through development research or design research. Development research protocols require intensive and long-term collaboration among researchers and practitioners. In this article, we propose guidelines for implementing development research models more widely, and conclude with a prescription for an online collaborative learning research agenda for the next five to ten years.

- Reinking, D., & Watkins, J. (2000). A formative experiment investigating the use of multimedia book reviews to increase elementary students' independent reading. *Reading Research Quarterly, 35*(3), 384-419.
  Abstract: This study investigates how a computer-based instructional intervention (creating multimedia reviews of books) might increase fourth and fifth graders' independent reading. The study finds that the success of the intervention was related to the mediating effects of using technology, changes in the interactions among students and teachers, and students' engagement in relation to their reading ability. It also notes several other factors.

Domain of curriculum

- McKenney, S., & van den Akker, J. (2005). Computer-based support for curriculum designers: A case of developmental research. *Educational Technology Research & Development, 53*(2), 41-66.
  Abstract: In this article, we explore the potential of the computer to support curriculum materials development within the context of secondary level science and mathematics education in southern Africa. During the four-year course of the study, a computer program was developed named CASCADE-SEA, which stands for Computer Assisted Curriculum Analysis, Design and Evaluation for Science (and mathematics) Education in Africa. By carefully documenting the iterative process of analysis, prototype design, evaluation, and revision, we sought insight into the characteristics of a valid and practical computer-based tool that possesses the potential to affect the performance of its users. The results of this study include the CASCADE-SEA program itself, which assists users in producing better quality materials than they otherwise might, while they also learn from the development process. Further, this research has contributed to the articulation of design principles and related developmental research methods. This article highlights the research and development that took place, and only briefly addresses the tool itself.

- McKenney, S., Nieveen, N., & van den Akker, J. (2002). Computer support for curriculum developers: CASCADE. *Educational Technology Research and Development, 50*(4), 25-35.
  Abstract: This article examines research on a computer-based tool, CASCADE (Computer Assisted Curriculum Analysis, Design and Evaluation), that was developed at the University of Twente (Netherlands) to assist in curriculum development. The article discusses electronic performance support systems and the need for increased attention to implementation and impact studies.

- Nieveen, N.M. (1999). Prototyping to reach product quality. In J.J.H. van den Akker, R. Branch, K. Gustafson, N.M. Nieveen, & Tj. Plomp (Eds.), *Design approaches and tools in education and training* (pp. 125-136). Dordrecht: Kluwer.
  Abstract: This chapter provides a framework for product quality consisting of the following three criteria: validity, practicality and effectiveness, and provides insight into the applicability of the framework in various domains of educational product development. In order to reach product quality, the prototyping approach is seen and understood as a suitable approach. This chapter discusses three significant characteristics of a prototyping approach: extensive use of prototypes, high degree of iteration and the role of formative evaluation, and the paramount importance of user involvement. The chapter illustrates the way the prototyping approach has been instrumental in developing a computer support system for instructional developers. During the prototyping process, the framework assisted in deciding the focus of each prototype and enhanced the transparency of the entire process.

- Nieveen, N.M., & van den Akker, J.J.H. (1999). Exploring the potential of a computer tool for instructional developers. *Educational Technology Research & Development, 47*(3), 77-98.
  Abstract: Information and communication technology tools currently permeate almost every professional domain. Those geared toward the field of instructional development have emerged in recent years. This article explores the potential for linking the domains of computer support and instructional development. This article reports on the design and evaluation of CASCADE (Computer Assisted Curriculum Analysis, Design and Evaluation), a computer system that supports instructional developers during formative evaluation efforts. Five prototypes of the system were created and evaluated on the basis of their validity (reflection of state-of-the-art knowledge and internal consistency); practicality (ability to meet the needs, wishes and contextual constraints of the target group); and effectiveness (improved user task performance). The results of this study suggest that the use of CASCADE could: (a) improve the consistency of formative evaluation plans and activities; (b) motivate developers by elevating their confidence in using formative evaluation activities; (c) save time; and (d) help to provide justifications for decisions made.

Some PhD theses utilizing design research as a research design

Over the years, various PhD these have been written in which design research has been applied as the main research approach. In this section we just mention a few in the order of year of defense. Some more can be found in Part B of this book, as a number of chapters in that book are reports of PhD research. Of course, many more dissertations can be traced via the internet.

Domain of curriculum development

- Nieveen, N. (1997). *Computer support for curriculum developers*. Doctoral thesis. Enschede (The Netherlands): University of Twente. Retrieved from http://projects.edte.utwente.nl/cascade/original/

- McKenney, S. (2001). *Computer-based support for science education materials developers in Africa: exploring potentials*. Doctoral thesis. Enschede (The Netherlands): University of Twente. Retrieved from http://projects.edte.utwente.nl/cascade/seastudy/

- Kouwenhoven, W. (2003). *Designing for competence in Mozambique: towards a competence-based curriculum for the Faculty of Education of the Eduardo Mondlane University*. Doctoral thesis. Enschede (The Netherlands): University of Twente. Retrieved from http://doc.utwente.nl/41442/1/thesis_Kouwenhoven.pdf
  Note: this is an example of design research in which the researcher was not actively involved in all phases of the design process.

- Akomaning, E. (2012). *Improving student internship through collaborative curriculum design in Ghanaian polytechnics*. Doctoral thesis. Enschede (The Netherlands): University of Twente.

Domain of distance education/mobile learning/networked learning

- Nihuka, K. (2011). *Collaborative course design to support implementation of e-learning by instructors.* Doctoral thesis. Enschede (The Netherlands): University of Twente. (also domain of Professional development) Retrieved from http://doc.utwente.nl/78096/1/thesis_K_Nihuka.pdf

- Palalas, A. (2012). *Design guidelines for a Mobile-Enabled Language Learning system supporting the development of ESP listening skills.* Doctoral Dissertation. Athabasca University (Canada). Retrieved from https://dt.athabascau.ca/jspui/handle/10791/17

- Ostashewski, N. (2013) *Networked Teacher Professional Development: Assessing K-12 Teacher Professional Development within a social networking framework*. Doctoral dissertation. Athabasca University (Canada). Retrieved from https://dt.athabascau.ca/jspui/handle/10791/26

- Shattuck, J. (2013) *Training higher education adjunct faculty to teach online: A design-based research study*. Doctoral dissertation. Athabasca University (Canada). Retrieved from https://dt.athabascau.ca/jspui/handle/10791/27

Domain of instructional technology

- Porcaro, D. (2011). *Omani undergraduate student reactions to collaborative knowledge building: A design research study.* Doctoral dissertation. University of Georgia (Athens, GA, USA). Retrieved from https://getd.libs.uga.edu/pdfs/porcaro_david_s_201105_phd.pdf

Domain of teacher education and professional development of teachers

- Hadi, Sutarto (18/12/2002). *Effective Teacher Professional Development for Implementation of Realistic Mathematics Education in Indonesia.* Enschede: University of Twente. Retrieved from http://doc.utwente.nl/58708/1/thesis_Hadi.pdf

- Teclai Tecle, Andemariam (2006). *The potential of a professional development scenario for supporting biology teachers in Eritrea*. Doctoral thesis. Enschede (The Netherlands): University of Twente. Retrieved from http://purl.org/utwente/55985

- Alayyar, G. (2011). *Developing pre-service teacher competencies for ICT integration through design teams.* Doctoral thesis. Enschede (The Netherlands): University of Twente. Retrieved from http://doc.utwente.nl/77918/1/thesis_G_Allayar.pdf

- Bakah, M.A.B. (2011). *Teacher professional development through collaborative curriculum design in. Ghanas polytechnics.* Doctoral thesis. Enschede (The Netherlands): University of Twente. Retrieved from http://doc.utwente.nl/78306/1/thesis_MAB_Bakah.pdf

- Agyei, D.D. (2012). *Preparation of pre-service teachers in Ghana to integrate Information and Communication Technology in teaching mathematics*. Doctoral thesis. Enschede (The Netherlands): University of Twente. Retrieved from http://doc.utwente.nl/80660/1/thesis_D_Agyei.pdf (also domain of ICT in education).

Domain of environmental education

- Wynveen, B.J. (2013). *Attaining sustainable behavior among non-environmentally-motivated individuals: A formative experiment*. Doctoral thesis. Clemson University (South Carolina, USA). Retrieved from http://etd.lib.clemson.edu/documents/1371151621/

Domain of literacy education

- Tracy, K.N. (2009). *A formative experiment investigating the use of nonfiction texts in writing workshop to assist fourth-grade readers and writers*. Doctoral thesis. Clemson University (South Carolina, USA). Retrieved from http://etd.lib.clemson.edu/documents/1252424273/Tracy_clemson_0050D_10292.pdf

- Abdallah, M.M.S. (2011). *Web-based new literacies and EFL curriculum design in teacher education: A design study for expanding EFL student teachers' language-related literacy practices in an Egyptian pre-service teacher education programme*. Doctoral thesis. College of Social Sciences and International Studies, University of Exeter, UK. Retrieved from https://eric.exeter.ac.uk/repository/handle/10036/3202

- Colwell, J. (2012). *A formative experiment to promote disciplinary literacy in middle-school and pre-service teacher education through blogging*. Doctoral dissertation. Clemson University (South Carolina, USA). Retrieved from http://etd.lib.clemson.edu/documents/1349104087/

- Taveggia, D.E. (2013). *Adapting a vocabulary notebook strategy to the needs of community college English language learners*. Doctoral Dissertation. Lawrence, Kansas (USA). University of Kansas. Retrieved from UMI Dissertation Publishing, http://disexpress.umi.com/dxweb, Order No. 3559060
  Note: This is a formative experiment.

Domain of mathematics education

- Armanto, Dian (2002). *Teaching multiplication and division realistically in Indonesian primary schools: a prototype of local instructional theory*. Doctoral thesis. Enschede (The Netherlands): University of Twente. Retrieved from http://purl.org/utwente/58710

- Fauzan, Ahmad (2002). *Applying realistic mathematics education (RME) in teaching geometry in Indonesian primary schools*. Doctoral thesis. Enschede (The Netherlands) University of Twente. Retrieved from http://purl.org/utwente/58707

- Bakker, A. (2004). *Design research in statistics education : on symbolizing and computer tools*. Doctoral thesis. Utrecht (The Netherlands): University of Utrecht. Retrieved from http://igitur-archive.library.uu.nl/dissertations/2004-0513-153943/inhoud. htm

- Larson, C. (2010). *Conceptualizing matrix multiplication: A framework for student thinking, an historical analysis, and a modeling perspective*. Doctoral dissertation. Indiana University, Bloomington (IN, USA). Retrieved from http://search.proquest.com/docview/748218708

- Wawro, M. (2011). *Individual and collective analyses of the genesis of student reasoning regarding the invertible matrix theorem in linear algebra.* Doctoral dissertation. University of California, San Diego (CA, USA). Retrieved from http://search.proquest.com/docview/885366193

- Sweeney, G.F. (2012). *Negotiating meaning for the symbolic expressions for vectors and vector equations in a classroom community of practice.* Doctoral dissertation. San Diego State University (USA). Retrieved from http://sdsu-dspace.calstate.edu/handle/10211.10/3063

Domain of science education

- Ottevanger, W. (2001). *Materials development as a catalyst for science curriculum implementation in Namibia*. Doctoral thesis. Enschede: University of Twente. Retrieved from http://doc.utwente.nl/87327/1/thesis_W_Ottevanger.pdf

- Knippels, M.C.P.J. (2002). *Coping with the abstract and complex nature of genetics in biology education: The yo-yo learning and teaching strategy*. Doctoral thesis. Utrecht (The Netherlands: University of Utrecht. Retrieved from http://igitur-archive.library.uu.nl/dissertations/2002-0930-094820/inhoud

- Tilya, F.D. (2003). *Teacher support for the use of MBL in activity-based physics teaching in Tanzania*. Doctoral thesis. Enschede (The Netherlands): University of Twente. Retrieved from http://purl.org/utwente/41462

- Westbroek, H.B. (2005). *Characteristics of meaningful chemistry education - The case of water quality*. Doctoral thesis. Utrecht (The Netherlands): University of Utrecht. Retrieved from http://igitur-archive.library.uu.nl/dissertations/2005-0922-200121/index.htm

- Mafumiko, F.S.M. (2006). *Micro-scale experimentation as a catalyst for improving the chemistry curriculum in Tanzania*. Doctoral thesis. Enschede (The Netherlands) University of Twente. Retrieved from http://purl.org/utwente/55448

- Prins, G.T. (2010). *Teaching and learning of modelling in chemistry education. authentic practices as contexts for learning*. Utrecht (The Netherlands): University of Utrecht. Retrieved from http://igitur-archive.library.uu.nl/dissertations/2010-0525-200212/UUindex.html

- Feierabend, T. (2011). *Concepts for socio-critical and problem-oriented science education to foster evaluation competence in the domain of climate change*. Doctoral Dissertation. University of Bremen (Germany). Available via author: feierabend.t@gmx.de

- Burmeister, M. (2012). *Education for sustainable development in chemistry education – A project of innovating practice by strengthening the expertise of teachers*. Doctoral Dissertation. University of Bremen (Germany). Available via author: mareike@familie-burmeister.net

Domain of school effectiveness

- Archer, E. (2010). *Bridging the gap: optimising a feedback system for monitoring learner performance*. Pretoria: University of Pretoria. Doctoral thesis. Retrieved from http://upetd.up.ac.za/thesis/available/etd-02022011-124942/

Author Biographies

Tjeerd Plomp (1938) is professor emeritus of the University of Twente, Enschede, The Netherlands, where he has been professor of curriculum from 1981 – 2002. He was in charge of teaching educational design methodology in the (at that time) Faculty of Educational Science and Technology. He was chair of the IEA, the International Association for the Evaluation of Educational Achievement, from 1989 - 1999. In the IEA he served as chair for the 'Computers in Education' study (Comped), the Third International Mathematics and Science Study (TIMSS) and the IEA Second International Technology in Education Study (SITES). His research interests are educational design and design research, international comparative research, information technology in the curriculum and 21st century skills in education. He has been recently involved as advisor in various research projects and programs utilizing design research, both in The Netherlands and internationally. Tjeerd Plomp has written many articles and chapters in books and has co-edited special issues of international journals and books.

Email: t.plomp@utwente.nl

Jan van den Akker is Director General of SLO (Netherlands Institute for Curriculum Development). Besides, for many years, he was affiliated with the University of Twente as professor in the domain of Curriculum Design and Implementation.
His main areas of expertise are:

- curriculum policy making
- curriculum development in interaction with teacher learning and school development
- design and evaluation of curriculum materials
- methodology of design research in education.

Email: j.vandenakker@slo.nl

Koeno Gravemeijer (1946) is professor emeritus in Science and Technology Education at Eindhoven University of Technology in the Netherlands. Earlier he was professor in Mathematics Education at Utrecht University (Utrecht, the Netherlands) and Vanderbilt University (Nashville, USA). His research interests concern design research as a research method, domain-specific instruction theories (such as the theory for realistic mathematics education, RME), and students' use of symbols and modeling. In relation to the latter he developed the so-called 'emergent modeling' design heuristic. Next to his research activities, he also has been involved in instructional design. He has been the leading author of an RME textbook series for Dutch primary schools, and he was involved in the development of the textbook series for the American Middle School, "Mathematics in Contexts"—a collaborative project of the Dutch Freudenthal Institute and the University of Wisconsin-Madison.

Email: koeno@gravemeijer.nl

Paul Cobb is professor of Mathematics Education at Vanderbilt University. He teaches courses in elementary mathematics teaching, design research methodology, and the institutional setting of mathematics teaching and learning. His research focuses on improving the quality of mathematics teaching and thus student learning on a large scale, and on issues of equity in students' access to significant mathematical ideas. He is an elected member of the National Academy of Education and an Invited Fellow of the Center for Advanced Studies in the Behavioral Sciences. He received Hans Freudenthal Medal for cumulative research program over the prior ten years from the International Commission on Mathematics Instruction (ICMI) in 2005, and the Silvia Scribner Award from the American Education Research Association in 2008. An overview of his work can be found in Yackel, E., Gravemeijer, K., & Sfard, A. (2011) (Eds.) *A Journey into mathematics education research: Insights from the Work of Paul Cobb.* New York: Springer.

Email: paul.cobb@vanderbilt.edu

Email: bbannan@gmu.edu

Brenda Bannan (1964) is an associate professor in Instructional Design and Technology Program in the College of Education and Human Development at George Mason University, Fairfax, VA, in the U.S. Dr. Bannan's research interests involve the articulation of methods related to design research in the learning technologies field. Dr. Bannan has authored numerous articles and book chapters in the areas of design research, design process, mobile learning, and augmented reality in several areas including educational technology, special education, language learning, and geoscience. Her work links design process, research and development of innovative learning technology systems and solutions.

Email: akelly1@gmu.edu

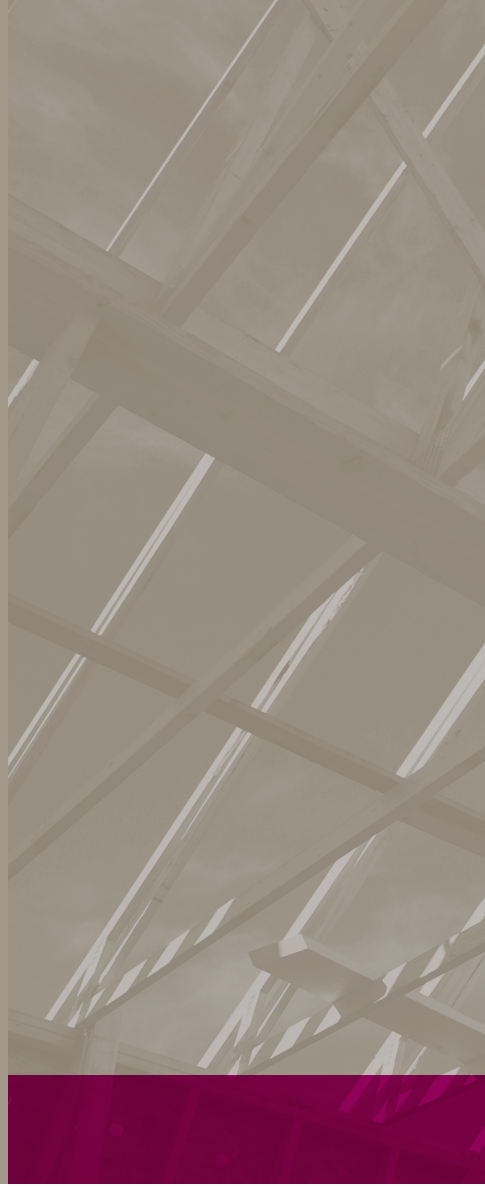
Anthony Kelly is a professor of Educational Psychology at George Mason University, Fairfax, Virginia, USA. His work centres on research methods in education. He has recent grants from the US National Science Foundation on big data applications in learning. He is interested design-based research methods, and in the design of assessments. He is lead editor of the Handbook of design research methods in education with Richard Lesh and John Baek (2008).

Nienke Nieveen (1969) is senior researcher at SLO [Netherlands Institute for Curriculum Development]. Her work centres on coordinating the Institute's evaluation activities and the thematic strand 'Curriculum and teachers'. She has been engaged in several projects related to school-based curriculum development and professional development of teachers in the field of curriculum development. Her dissertation, in 1997, was based on a four-year design research project in the field of curriculum design and evaluation. From 1997 to 2007 she was assistant professor at the University of Twente (Enschede, The Netherlands), specializing in curriculum design research and school-based curriculum development. The following co-edited books represent her orientations: Design approaches and tools in education and training (1999), Educational design research (2006), Introduction to educational design research (2009) and Schools as curriculum agencies: Asian and European perspectives on school-based curriculum development (2010).

Email: n.nieveen@slo.nl

Elvira Folmer (1976) is senior researcher at SLO [Netherlands Institute for Curriculum Development], since 2007. Her work centres on coordinating the Institute's evaluation activities and the thematic strand 'Curriculum monitoring and evaluation'. She has been engaged in the evaluation of several curriculum renewal projects in secondary education. The subject of her dissertation, 2004, was the effects of class size reduction in the lower grades of primary education in the Netherlands. After finishing her PhD she worked as a researcher at the Radboud University.

Email: e.folmer@slo.nl



SLO is the Netherlands institute for curriculum development. We are bridging the contexts of policy, research, and practice. Our expertise focuses on the development of curricular goals and content for various educational levels, from national policy to classroom practices. We closely collaborate with many different stakeholders from policy circles, schools (boards, principals, teachers), research, civic organizations, and the society at large. This allows us to design and validate relevant curriculum frameworks, to elaborate exemplary materials and to evaluate these in school practices.
Our products and services support both policy makers and schools and teachers in making substantive curricular decisions and in elaborating these into relevant, inspiring and effective education.

ISBN: 978-90-329-2334-1

SLO

PO box 2041
7500 CA Enschede
The Netherlands

T +31(0)53 484 08 40
E info@slo.nl

www.slo.nl

slo