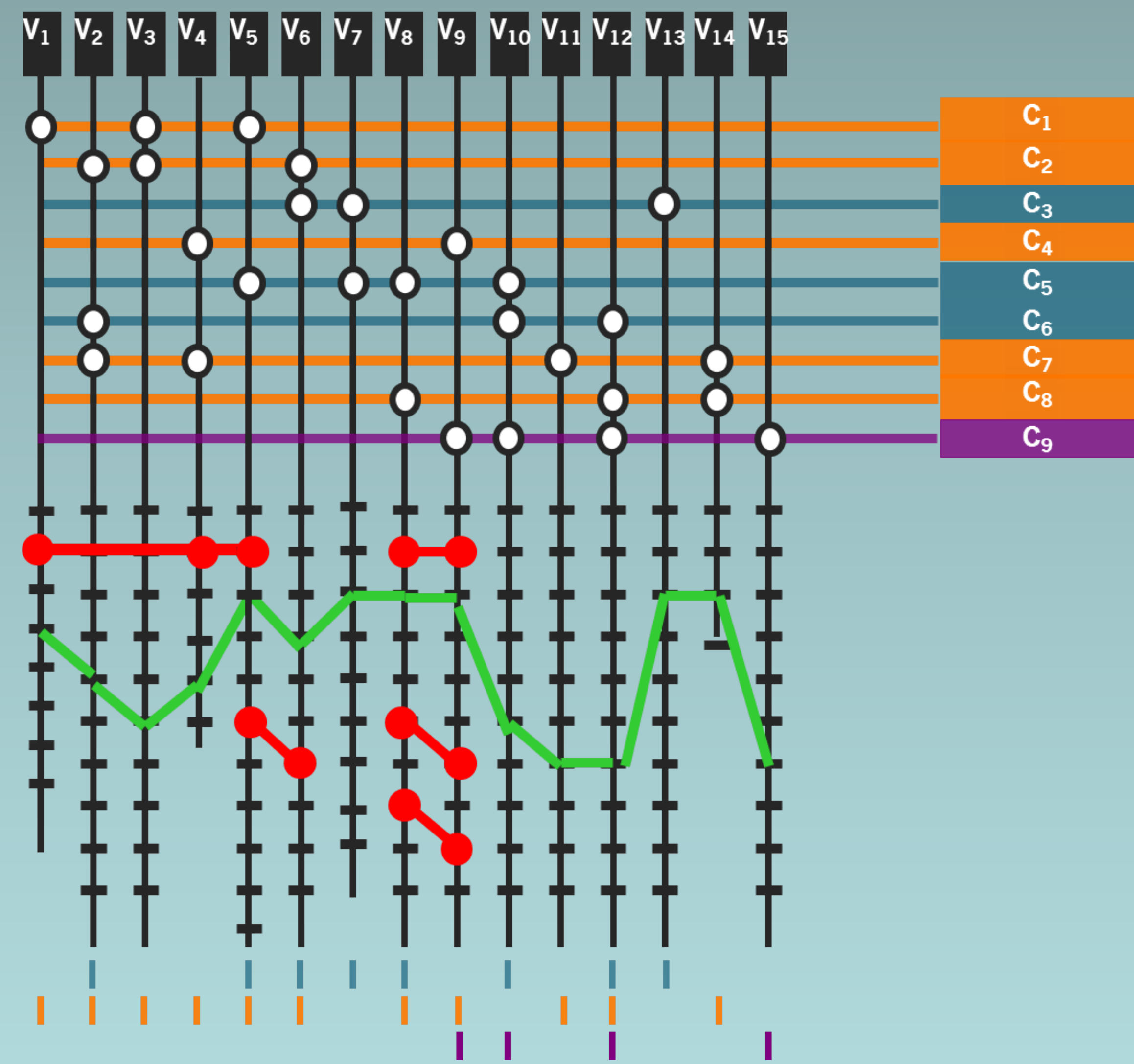
On design-oriented research and digital learning materials in higher education
Rob J M Hartog 2012

On design-oriented research and digital learning materials in higher education



Rob J.M. Hartog

On design-oriented research and digital learning materials in higher education

Rob J.M. Hartog

Thesis committee

Promotors
Prof. ir. A.J.M. Beulens
Professor of Information Technology
Wageningen University

Prof. dr. ir. J.Tramper
Professor of Bioprocess Engineering
Wageningen University

Other members
Prof. dr. T. Bisseling, Wageningen University
Prof. dr. M.J. Goedhart, University of Groningen
Prof. dr. ir. H. Gruppen, Wageningen University
Prof. dr. J.L. Top, VU University Amsterdam

This research was conducted under the auspices of
Wageningen School of Social Sciences

On design-oriented research and digital learning materials in higher education

Rob J.M. Hartog

Thesis
submitted in fulfilment of the requirements for the degree of doctor
at Wageningen University
by the authority of the Rector Magnificus
Prof. dr. M.J. Kropff,
in the presence of the
Thesis Committee appointed by the Academic Board
to be defended in public
on Tuesday 18 December 2012
at 1.30 p.m. in the Aula

On design-oriented research and digital learning materials in higher education,

276 pages.

PhD thesis, Wageningen University, Wageningen, NL (2012)

With references, with summaries in Dutch and English

ISBN 978-94-6173-508-9

Contents

Glossary

---

action

In digital learning materials typical actions are: the selection of an option, the selection of a tool, setting a parameter, making a connection et cetera.

activating learning material

Learning material that requires the student to take conscious actions in order to progress through the material.

affordance

The set of perceived and actual properties that determine just how a thing could possibly be used.

applet

A small application that runs within a browser.

articulation of design goals

In this thesis, the term articulation of design goals is used to cover elaboration, expansion or refinement of design goals and requirements into operational design requirements.

In much literature the term ‘to derive’ is used; this term is avoided in this thesis because in some disciplines ‘to derive’ is close to ‘logical inference’.

In much literature the term 'to analyze' is used; this term is avoided in this thesis. 'To analyze' does not adequately convey that articulation of design goals almost always implies addition of information and knowledge. Furthermore, while articulation to a certain extent might involve analysis, it also involves to a large extent synthesis.

artifact

Artifacts, are "constructs (e.g., concepts, terminologies, and languages), models, methods, and instantiations (i.e. concrete solutions implemented as prototypes or production systems). Concrete manifestations of such artifacts can be axioms, guidelines, frameworks, norms, patents, software (with open source code), business models, enterprise start-ups, and much more."[16]

assessment

(in the direct context of learning) A process aimed to assess if the student has achieved a set of operationally defined learning objectives or intended learning outcomes. Assessment implies a measurement or observation and an interpretation of this measurement or observation. In higher education, an exam is intended to be an assessment.

to direct attention to ...

To allocate perception and information processing capacity to … .

authenticity - degree of

The extent to which a task is similar to an authentic task.
The difference between an authentic task and a task with a certain degree of authenticity is often based on
(1) forms of support for the student such as hints and on
(2) simplifications such as the skipping of details.

authentic task

A task that occurs in real life. Usually we will refer to a task that will occur naturally in one of the professional contexts at which the curriculum program of the student is aimed. For university education this professional context is often, but by far not exclusively, defined by a research position.

capacity building

"…enhancement of capabilities of people and institutions to improve their competence and problem solving capacities in a sustainable manner." [17]

capacity of a resource (1)

The maximum amount of output per unit of time that can be delivered by this resource. This can be the number of actions per unit of time, the number of hours of work per week that a lecturer has available for educational duties, the amount of data processing per unit of time, the number of bits per second that can be transferred by a data connection (i.e. the bandwidth) et cetera.

capacity of a resource (2)

Sometimes the maximum amount of something else, for instance in case of the capacity of a room: the maximum number of seats in the room.

case (part of learning material)

A case is a combination of a situation, an assignment and a role for the student. A case is constructed as a network of interactions. For instance a case might be a situation in which the student is placed in a role as bioprocess engineering consultant and in which the student has to design a purification process for a given fermentation broth.

case study

In order to avoid confusion we will only use the term case study for a set up intended to study or evaluate the use of a body of learning material. The term case study will refer to an implementation and a process of using the learning material in a learning scenario. Thus, whenever the term 'case' is used we will not refer to a 'case study'.

closed question

A question where the required response should be based on options or value ranges offered to the respondent (usually the student). In general, a score for the closed question can be generated automatically by comparison of selected values with intended values.

compelling comparison

Comparison of two innovations under otherwise similar conditions [18].

competence
" - the integrated set of capabilities (or competencies);
- consisting of clusters of knowledge, skills, and attitudes;
- necessarily conditional for task performance and problem solving;
- and for being able to function effectively (according to certain expectations or standards); and
- in a certain profession, organization, job, role and situation."
[19]

competency
" - a situated element of competence, which can be;
- behaviour-oriented and/or ;
- task-oriented; and
- meaningful in a specific context and at a sufficient level of specification."
[19]

component
A constituting element. Sometimes the meaning is more specific: A module with one or only a few functions.

complex
The term *complex* is avoided in this thesis. Only when we refer to literature and when clear reference to this literature induces the occurrence of the term we will use the term *complex*.

concept
"A mental representation representing a class of objects, events or other entities by their characteristic features and/or mental images." [14]

concept map
A graphical representation of a set of names of concepts and the relationships between those concepts.

constructive alignment

Systematical alignment of "the teaching/learning activities, and the assessment tasks to the intended learning outcomes, according to the learning activities required in the outcomes."[9]

constraint

A constraint on a set of variables defines a relation (in the mathematical sense of the word) over the domains of these variables. The domains of these variables define a Cartesian product. The constraint defines a subset of this Cartesian product.

The constraint can be defined extensionally (i.e. every element is explicitly listed) or intensionally (i.e. by some expression).

content

The aggregation of representations of knowledge: texts, pictures, video clips, audio tracks et cetera.

content management system

A system that is dedicated to manage a large volume of content.

course content

All the texts, pictures, video clips and other objects that represent knowledge covered by the course.

context of a system

(see also outer environment)

The context of a system is everything that is
(1) not part of the system and

(2) is related to the system by an interface

Note that the context **of** a piece of learning material consists of the learning management system, the students, the teacher, et cetera.

context within a piece of learning material

Within a piece of learning material, a phenomenon or concept may be presented 'within a context'. In that case, the term 'context' refers to a context of the phenomenon or concept that is illustrated. For instance, we might illustrate the concept of 'confounding' in the context of a study on the relationship between fiber-intake and blood pressure.

criterion

A criterion is a component of an operational requirement that defines which values of the variables (dimensions) in this requirement satisfy the requirement.

A one-dimensional criterion is a set of one or more values in a one-dimensional requirement. Sometimes the term 'criterion' refers to the dimension on which the criterion has been defined. In particular when this dimension has only two values, one satisfactory and one unsatisfactory, it is common usage to use one of these values to refer to the dimension as well. Thus we may say that 'ethical acceptability' is a criterion for a goal.
In practice, this often implies that we want to revert to a broader definition of criterion in cases where we fail to define variables and values that adequately express our intentions.

design (verb)

The systematic, intelligent **generation** and **evaluation** of **specifications** of artifacts whose form and function achieve objectives and satisfy specified constraints. In this definition, **generation** should be understood to include **decision-making** as well as **formulating the specifications** for the artifact. The design is a model of the intended instantiation. When evaluation has to imply tests of its concrete manifestation in operation, the design process includes the realization, implementation and use of that concrete manifestation. Finally, the process of stating objectives is also often considered to be part of the design process. This definition of a design process is an extension of a definition given by Dym & Little [20].

design (noun)

= design proper

A design is a model of a corresponding instantiation.
The design can be represented by an abstract model (a 'blueprint') that still has to be realized, implemented and set into operation OR the design can be represented by a direct realization that still has to be implemented and set into operation.

design-based research
Research aimed to improve understanding of learning, based on primarily empirical studies of realized and implemented designs of instruction and instructional environments, and on comparative studies.

design goal(s)
A design goal is a set of points defined by a set of constraints in a multidimensional design space. A design goal is an outcome of DOR. Often ' design goal' or 'design goals' will refer to a short description in natural language of the intended result of a design process. Here, 'short' means 'in a few sentences'. Sometimes it is more natural to refer to a singular overall goal, sometimes it is more natural to refer to a few goals that make up the overall design goal. A design goal that is not further articulated into objectives, requirements and operational definitions is seldom very clear. A DOR project may start with an initial short description of the 'design goal', but the design goal that is outcome of the project will usually not be represented adequately by this initial description.

design guideline
A requirement on variables that describe the design process. A design guideline is intended to guide this process. Design guidelines imply constraints on process variables and restrict the set of possible design decisions.

design-oriented research (DOR)
Research that primarily aims to produce an innovative design and applies typical concepts of design methodology.

This implies the following research questions:

1. what are, in a specific real life context, goals that make sense and why
2. how can these goals be articulated in terms of measurable quantities
3. is it possible to achieve these goals
4. if so how

design pattern

A reusable configuration of basic components or activities (including parameter settings), which fits a partial design problem.

design process

See design (verb)

design proper

See design (noun)

design space

The design space is the space of 'candidate' solutions for the design problem. When the design problem is represented as a constraint satisfaction problem this implies that the space defined as the Cartesian product of the domains of the variables of all relevant constraints is a design space.

digital learning material

Learning material that is fully based on the use of digital computers: computer based learning material.

distance learning

Learning in an instructional setting that does not require the students to be present in buildings of the university.

domain (of a variable in a constraint model)

The set of possible values of the variable.

domain (knowledge domain)

= a body of knowledge on a subject or collection of subjects.

eBook

A body of presentational materials that provide a fairly integrated learning experience and have to be accessed via a computer screen. The eBook may include not only text and figures but also multimedia components such as audio, video clips and animations. In addition the eBook is delivered by an environment that supports searching by browsing, hierarchical search, keyword search, cross reference search and free text search. The definition of what constitutes an eBook is likely to evolve over time. For instance it will also imply increasingly advanced levels of adaptivity.

eBook Plus

An eBook that also includes interactive materials.

eLearning

Learning that relies on learning activities supported by ICT.

elaboration

"A category of learning processes by which learners connect information elements to each other and to knowledge they already have available in memory." [14]

external architecture

The set of architectural aspects that are directly related to the experience of the user.

flexibility of a design

Indicates to what degree changes in requirements or changes in the outer environment require changes in the design.

function of a system

A description of what a system has to do or what output a system should produce given a certain input.

functional requirement

A design requirement that is part of the set of requirements that define a function of a system. In practice it is often not possible to make a clear distinction between functional requirements and non-functional requirements.

functional specification

= functional requirement.

functionality

A set of functions of a system.

gap management

Management of the dynamics of the gap between current perceptions that each separate stakeholder has of the current design goal.

granularity of a set of learning objects in a module

The reciprocal of the average size of learning objects in a module.

guideline

See design guideline.

implementation

Fitting the realized design to its intended context. Because some of the initial assumptions about the intended context of a design may turn out not to hold by the time the design has been realized, implementation may require adjustments of the context.

inner environment

The system as it is distinguished from its (outer) environment.

integrated

The meaning of integrated is not defined generically.

Usually it means that separate systems, functions, aspects et cetera cannot be distinguished.

integrated learning experience

A learning experience is called integrated if it does not involve any form of cognitive load due to switching between different tasks or due to switching between the use of different media and if any other effort needed for such switches is negligibly small.

interaction

A combination of a user gesture and a system response.

interface

The *interface* of a system describes a set of assumptions about the environment of the system and a definition of the function(s) of the system.

interoperability (in learning management)

The condition that any learning object, package, course cartridge, question or test can be functional in any LMS as long as both conform to the same specification.

instructional design

Design of instructional processes and learning materials.

learning goal

An abstract formulation of the intended outcome of a learning process. See also learning objective: goals are "more general and inclusive than objectives".
An 'intended learning outcome' can be an objective. For reasons of readability, the use of the term learning goal in this thesis can also be in reference to 'target competency' or include 'target competency'.

learning material

Material that is specifically designed and developed in order to support learning processes of a learner, and aimed towards a specific learning goal or set of learning goals.

learning management system

An information system that at least (1) supports authorized management of digital learning materials including quizzes and trial exams; (2) provides authorized access to these materials; (3) supports electronic communication.

learning object

A reusable piece of digital learning material aimed to support the achievement of ‘atomic’ learning objectives. Many other terms with more or less the same meaning such as *reusable content object* (RCO) or *unit of learning material* (ULM) are used in literature. (see Box 2-1 Learning Objects).

Thus, a learning object is a module of digital learning material with one or only a few functions and few assumptions with a small scope.

presentational learning object

A learning object that *presents* information for instance in the form of texts, diagrams, screen-recordings or animations. Currently, well known presentational learning objects are MS Powerpoint ™ slides.

learning object repository

A searchable store of digital learning objects that can be accessed over the internet.

learning objective

In this thesis this term is used to indicate a sub goal of a learning goal. The term learning objective refers to a smaller scope and is more specific than learning goal, but is not yet an operational definition. When the learning goal is: "the student has adequate knowledge of fluid flows in living systems" a corresponding learning objective might be: "the student understands the Reynolds number". This learning objective might be further articulated in terms of operationally defined requirements of student behavior. An example of such a requirement might be: "the student can determine the characteristic length in the Reynolds number, for a hollow cylinder".

learning path (suggested or followed)

A sequence of actions or activities suggested to the student connecting a prior knowledge state with a learning goal.

learning scenario

A design of a process in which roles are defined for all actors and for all learning materials; in addition the scenario may define which activities and actions are planned and in what order.

load (of a resource)

Effort or work allocated to this resource. May be expressed in terms of the percentage of the capacity that is required for a process or task.

meta-level developmental research

The intended output of the research is not primarily the result of the development itself but the output **of the study of the results**, or the output of **studies of the developments.**

meta-data

Data about data; data that describe data, learning objects or resources.

For instance: attributes that describe a photograph, what it represents, who has the copyrights et cetera.

meta-<object -or-process>

Indicates recursive application of the meaning of the object-or-process to a (class of <object-or-process> . For instance, meta-analysis: analysis of analyses. Meta-model: model of model; meta-cognition: cognition of cognition. For instance, learning-to-learn is a meta-cognitive process.

minimal

What is left after removal of all that is unnecessary, that is after removal of everything for which no arguments can be provided.

minimal dependency

Any dependency that is left after removal of all unnecessary dependencies.

minimal interface

An interface that consists of a minimal set of assumptions with minimal scope or assumptions with respect to a standard and one or few functions with a minimal scope.

model

A representation of a system.

module

A module is a system that can be replaced by another system without inducing the need to impose new requirements on the outer environment. Thus, the interface of the replacing module will be the same as the interface of the original module.

modular system

A system that is composed of modules.

modularity of a system

The modularity of a system is a qualitative indicator based on the number and scope of functions and assumptions in the interface of any of the modules in the system. The higher these numbers and their scope, the lower the modularity.

modeling formalism

The term representational formalism can almost always be interpreted as 'modeling formalism' . However, some formalisms (such as natural language) are in many disciplines not linked to the term 'model'.

normative study effort of a specific body of learning material

The number of hours of study that the average student has to invest in order to achieve the learning goal defined by the learning material using the learning material. This definition assumes self study as learning scenario. In this thesis it has been assumed that any other learning scenario will induce a lower study effort.

object system

A model is itself a system. Often we want to stress the distinction between the system that is the model and the system that is the counterpart in reality of this model. For this reason we call the latter ‘object system’. Thus, the term 'object system' refers to the system under consideration in which a situation occurs that induces the demand for a new subsystem to be designed and realized. The new object system will ultimately consist of the realized design and its outer environment (context).

objective function

In an optimization problem the objective function is the function that we want to optimize given a set of constraints.

open source

Software is considered *open source* if

(1) the binary version and the source code are available to anyone for any purpose,

(2) these can be used and changed by anyone for any purpose,

(3) both the original and the changed version can be (re)published under the same license as the original version.

operational definition (of a variable or requirement)

A definition that implies a procedure for measuring the value of the variable or for determining if the requirement is satisfied.

operationally defined design requirement

A requirement that specifies how to determine if the realized design satisfies the requirement.

operational design requirement

= operationally defined design requirement.

operationalization of a (design) goal

= formulating operationally defined design requirements.

optimization problem

A mathematical or logical problem for which a definition of optimality has been defined such that it makes sense to speak of an optimal solution. Well-known examples of optimization problems are linear programming problems.

optimizing

Making progress in the process of solving an optimization problem.

outer environment of a system

everything that is not part of the system

= context of the system

paradigm

"A paradigm is what the members of a scientific community share, and, conversely, a scientific community consists of men who share a paradigm." [1] Kuhn acknowledges the circularity in the definition. The concept is intended to cover more than the concept 'scientific theory'. In addition to a theory, a paradigm implies a shared set of examples based on exemplary past achievements that all novices in the scientific community learn in introductory courses and textbooks. Students within the discipline thus construct the same tacit knowledge with respect to the meaning of symbols in a range of situations. A paradigm includes 'a way of seeing' (see also [21]).

paradigm example

For digital learning materials, we define paradigm examples as examples of realized designs for which a proof of concept has been provided. A paradigm example illustrates and clarifies in particular the meaning of design requirements, guidelines and patterns for a specific part of the subject matter or a specific learning objective.

parameter

An auxiliary quantity in a mathematical expression which quantity is not used to classify the expression [22].

pedagogical content knowledge

Subject matter related knowledge on learning and instruction that we cannot yet decompose into (1) distinct instructional design knowledge, (2) subject matter knowledge and (3) knowledge of target populations without loss of meaning.

performance objective

A performance objective describes an action, activity or task that a learner should be able to do in order to demonstrate intended performance, as well as one or more corresponding contexts. The description will refer to qualitative and/or quantitative variables and constraints on those variables.

plan

A plan is a model of a process that still has to be executed.

A plan is the result of a design process.

prerequisite knowledge

Assumptions with respect to the prior knowledge that are part of the interface of a module of learning material or a course.

Some authors prefer to speak of prerequisite information and skills [23].

Merriënboer and Kirschner [14] promote analyis of prerequisite knowledge in terms of concepts, principles and plans.

proof of concept

(1) one or a few explicitly described concepts, (2) a proof of feasibility, (3) an explanation that links the proof of feasibility to the concept(s) in a satisfactory way and (4) the absence of alternative plausible explanations.

proof of feasibility

Evaluation of the operation of a realized and implemented design may prove that it is possible to satisfy the design requirements. This is called a proof of feasibility. A proof of feasibility shows it is possible, at least once, to satisfy in reality the constraints in design space. A proof of feasibility falsifies the hypothesis that it is not possible to satisfy in reality the constraints in the design space.

qualitative

Not quantitative.

quality of a decision

The degree to which the process of making this decision has been conform generally accepted guidelines for making a decision of this type.

quantitative

Based on values with which it makes sense to do calculations.

quiz

A set of closed questions which are not very highly related and which are presented to stimulate active learning.

rapid prototyping

This means that a series of prototypes is developed in short design & development cycles, primarily aimed to help users and other stakeholders to understand the functionality of the system to be delivered and to help developers to understand what users really want. The intermediate prototypes are 'mock-ups', rather than functioning intermediate forms of the realized design. Rapid prototyping supports evolutionary development. For many types of systems in information systems development, rapid prototyping is supported with adequate tools.

refactoring

Improving the internal structure of the code of a software system without introducing inconsistencies with the interface of the system.

rendering

The process of realizing a visual or audio-visual presentation that satisfies a set of formal constraints.

representational formalism

Any formalism to represent a system or a process. Examples of representational formalisms are: natural language, the language of differential calculus, any computer language, the formalism used for blueprints of buildings, the formalism used for geographic maps et cetera.

requirement

Requirements articulate a goal. Design requirements can be requirements as to the design proper, the realized design, the realized design operating in its context or the new state of affairs at which the design is aimed.

resource

Any entity that contributes to the realization of a goal.

resource capacity

The maximum load that can be allocated to the resource.

return on investment

The aggregate benefits that can directly be attributed to an investment in a system or process and that are realized within a specified time span after the start of the operation of the system or process.

robustness

Insensitivity to variance of values of variables of its intended context.

screen-recording

A recording of what happens on the computer screen. The recording can contain annotations in text balloons and can also be enhanced with an audio track. The result is often called ‘a movie’ and looks very much like a video take. [24].

seamless access

Access that does not by itself impose any cognitive load.

size of a learning object (see granularity)

The time necessary for the average student in the target population to achieve the learning objective(s) of the learning object with the support of the learning object.

split-attention effect

This effect refers to extraneous cognitive load due to the need to combine information that is presented distributed over physically independent entities. For instance when part of information is presented at one time and complementary information is presented a little later, or part of the information is presented in one location and complementary information is presented in another location. An example of the latter occurs often in diagrams in books that are accompanied by descriptive text just below the diagram.

storyboard

A diagram consisting of screens, interactions and corresponding references to screens. The diagram shows all the possible paths along screens that a student can follow (see Figure 2-1 An impression of a story board).

subject matter

A body of disciplinary knowledge and skills that students are supposed to acquire or construct in a course (also: course content or part of the course content).

state (of a system)

A vector in the state space of the system

state space (of a system)

The space defined by the state variables of a system.

state variable

A variable that represents a property of a system and has a unique value for a certain state of the system.

tacit knowledge

Any knowledge that is not yet expressed in some form such that this expression is adequate to support sharing this knowledge among different people (see [13, 15, 25]).

target competency

A competency that serves as a component of a learning goal.

trial exam

A set of questions or an assignment including a scoring and marking model, and a set of correct answers and solutions, intended to clarify to the student what to expect from the exam.

validation of a model

The process of demonstrating that a model is a useful representation of an object system.

vignette

A brief verbal outline of a short process or simple situation.

verification of a model

The process of demonstrating that the translation of one model (represented in a specific representational formalism A), into another model (represented in another representational formalism B) is correct with respect to the purpose of the latter model.

web-based learning material

Digital learning material that relies on web-technology.

web service

A web service is a software system designed to support interoperable machine-to-machine interaction over a network. It has an interface described in a machine-processable format (Web Services Description Language (WSDL) [26]).

whole task competency

The competency to carry out a specific task completely, i.e. including all its details.

working memory

Central in Cognitive Load Theory is the concept of a "severely limited working memory" with "partial independent processing units for visual/spatial and auditory/verbal information, which interacts with comparatively unlimited long-term memory." [14]

List of Acronyms

ALM	Activating learning Material
CAD	Computer-Assisted Design
CBA	Computer-Based Assessment
CMS	Content Management System
CPD	Continuous Professional Development
DBR	Design-Based Research
DSE	Discipline Specific Education
DOR	Design-Oriented Research
ET	Educational Technologist
FBT	Food and BioTechnology
ICT	Information – and Communication Technology
ID	Instructional Design
IKM	Information and Knowledge Management
ITS	Intelligent Tutoring System
LMS	Learning Management System
LO	Learning Object
LOM	Learning Object Metadata
LOR	Learning Object Repository
PCK	Pedagogical Content Knowledge

PLM	Presentational Learning Material
ROI	Return On Investment
SME	Subject Matter Expert
SMK	Subject Matter Knowledge
QTI	Question and Test Interoperability
WSDL	Web Services Description Language
WU	Wageningen University

Chapter 1 Introduction

Abstract

The primary context of the research described in this thesis has been a range of projects on design, realization, implementation, use and evaluation of digital learning materials at Wageningen University. These projects are shortly described in this chapter.

A view on digital learning materials and on design of digital materials for and within university education is presented. It is argued that design of digital learning materials in university education requires knowledge from a variety of knowledge domains. A short overview of relevant knowledge domains is given. In addition, it is argued that design, realization, implementation, use and evaluation of digital learning materials in university education and corresponding publication in peer reviewed journals needs an approach that crosses disciplinary boundaries.

Given this context, the scope of this thesis is described and related to methodological issues that emerged during the WU projects.

Finally, the results of this thesis, the intended audience and a short overview of the remaining chapters are described.

1.1 Digital learning materials and their design in university education

Learning materials are specifically designed and realized in order to support learning processes of a learner and intended to support the achievement of learning goals. Learning materials are resources for teachers and for learners. Examples of learning materials are textbooks, handouts of presentations, instructional movies and interactive computer-based instructional materials. *Digital learning materials* are fully based on the use of digital computers. For this reason, they can also be called *computer-based learning materials.*

The design of learning materials for university courses requires an approach that does not fit within disciplinary boundaries. On the one hand, given the current situation in most research areas, the central discipline will be the discipline that covers the *subject matter* to be taught. We define subject matter as a body of knowledge and skills that students are supposed to acquire or construct as part of their activities in a course. A course in higher education often incorporates more than one such body of knowledge and skills. This knowledge will often be considered disciplinary knowledge. The subject matter that was initially relevant for this thesis falls mostly within the fields of bioprocess engineering, molecular biology, food chemistry, epidemiology, human nutrition, nutrition behavior and nutrigenomics. In later WU projects many more subject matter fields came in focus. In literature on education, experts on such subjects are usually called Subject Matter Experts (SME's). In practice, most actual design of processes of learning and instruction in university education is carried out by SME's. This is because the subject matter in university education has a high level of abstraction. Many design decisions cannot be made without mastering this high level of abstraction. In university education, most SME's are professors, including assistant and associate professors.

On the other hand, actual design of learning materials will rely on knowledge from many different overlapping knowledge domains. For design of digital learning material for specific subject matter within a course in a university some of the relevant knowledge domains are : the subject matter, pedagogy/education/cognition, curriculum, design (including instructional design, multimedia design, interaction design, information systems design, software engineering, …), assessment in education, knowledge and information management (KIM), information and communication technology (ICT), learning technology standards, learning management systems, content management systems and learning object repositories, direct contextual knowledge, national and international developments in higher education, prior knowledge of university students in a globalized world and research methodology.

With 'direct contextual knowledge' we refer to knowledge of the possible contexts in which learning materials will have to be realized and might be implemented. This implies knowledge of typical forms of university organization, course scheduling, available infrastructural and other facilities, usual learning scenarios, et cetera. Part of the knowledge of the curricula in which the learning material might fit well can also be considered to be direct contextual knowledge.

A *Learning Management System* (LMS) is an information system that at least (1) supports authorized management of digital learning materials including quizzes and trial exams; (2) provides authorized access to these materials; (3) supports electronic communication. A *Content Management System* (CMS) is an information system that is primarily dedicated to manage a large volume of hypermedia content. LMS and CMS functionalities largely overlap. In practice, an important difference is often that the terminology and user interface of an LMS is geared to its intended use for education. A *Learning Object* is a reusable piece of digital learning material aimed to support the achievement of 'atomic' learning objectives. A *Learning Object Repository* (LOR) is a searchable store of digital learning objects that can be accessed over the internet.

As to the other terms in the list of relevant knowledge domains we assume that the reader feels to have some understanding of their meaning. We do not claim that the list is complete. Rather, we have provided this list to highlight that there are many different domains of knowledge that will be relevant. In Chapters 2 and 3 we clarify in more detail that designers will not be able to make use of all available knowledge and also that much of the knowledge that designers would like to have is not available. In this introduction we only give a first impression of the limitations of the most obvious resources.

1.1.1 primary knowledge domains

Design, realization, implementation, use and evaluation of digital learning materials in higher education require knowledge of many variables and constraints in a wide variety of knowledge domains. In this subsection we only provide a short overview of the few main knowledge domains. We indicate these as pedagogical knowledge, subject matter knowledge, knowledge of the educational context within the university, knowledge of information science and knowledge management and pedagogical content knowledge.

Scientific articles and textbooks on education, university education, learning and instructional design are mostly texts that only very little refer to subject matter [9, 12, 14, 23, 27-34]. Apart from a few exceptions, this literature does not intend to provide input for design of learning materials at the level where details of subject matter domain knowledge in higher education become important, nor actual outputs of such design activities for universities. Rather, the intention is to provide generic guidelines and to link or underpin these guidelines with theory and with empirical evidence. Insofar this literature provides examples on the use of concepts and guidelines, it is for many SME's not easy to find examples that demonstrate application of these concepts and guidelines on 'their' subject matter at the level of university education.

1.1.2 Where to publish

SME's can find and publish such examples in journals such as in Table 1-1. In such journals, scientists share knowledge and experience with respect to design of their teaching, their learning materials and their evaluations. Journals like these are sometimes called 'Discipline Specific Education' (DSE) journals. However individual articles in such journals are almost always about specific subject matter. This specific subject matter will usually fit mainly within the knowledge domain or domains that are covered by the pertinent discipline.

Table 1-1 Some examples of journals that in this thesis are labeled as 'DSE journals'

Advances in Physiology Education
American Journal of Pharmaceutical Education
Biochemistry and Molecular Biology Education,
CBE Life Sciences Education
Chemistry Education Research and Practice
Computer Science Education
Computer Applications in Engineering Education
International Journal of Electrical Engineering Education
Journal of Biological Education
Journal of Chemical Education
Journal of Engineering Education

Most articles from our WU projects on design, realization, implementation, use and evaluation of digital learning materials were published in such DSE journals.

Apart from these journals, also journals that are primarily aimed to provide a platform for disciplinary research may provide room for publications on design, realization, implementation, use and evaluation of learning materials and teaching/learning scenarios. For instance, some WU project results were published in Trends of Biotechnology [3], PLOS [35], e-SPEN [36], The American Statistician [37].

Because these journals are read by peers of the SME's (i.e. by molecular biologists, epidemiologists, process engineering et cetera) chances are that they also read these papers. Primarily, this might inform these peers of the materials that have been made available and of other results as well. Moreover, this may raise awareness of a way in which researchers in the fields of process engineering, molecular biology, nutrigenomics, food chemistry et cetera have been able to combine their research with their educational obligations. We believe that it can be important for the development of university-level education in all research fields if professors can realize synergy between their work as teacher with their work as researcher. In many universities these publications will fit well in the academic reward system. Also, sharing knowledge by publications in research journals of their own discipline may be good for education in the field.

1.2 "… a special amalgam of content and pedagogy …"

Design of digital learning materials in higher education requires the capture of subject matter related knowledge on learning and instruction that we cannot yet decompose into distinct instructional design knowledge, subject matter knowledge and knowledge of target populations without loss of meaning.

This is for instance knowledge on:

- how to structure subject matter (or content) for the purpose of learning and instruction for specific target groups and specific overall learning goals; note that disciplinary knowledge can often be structured from different viewpoints; examples are: the viewpoint of mathematical congruence between models in the field, the viewpoint of variables that can be measured, and the viewpoint of making connections with prior knowledge of learners in a specific target population;
- in what order to present subject matter (or content) in such a way that the knowledge can be acquired by the layman, novice or student with relatively little effort;

- which sequences of activities, questions and tasks are likely to be adequate to guide students in a specific target population during their learning process and why;
- what representations and which examples are likely to be adequate to clarify a certain law or concept for a specific target population and why;
- which misconceptions have in the past been most common in specific target groups including the sources of these misconceptions
- what subject matter (or content) related learning objectives fit a specific curriculum and target population and how to define these objectives operationally
  (this implies knowledge of how to assess achievement of these objectives)
- how to operationalize learning objectives

Much more examples can be given but for the purpose of this thesis the examples in this list should be sufficient to provide an idea of the type of knowledge that we mean.

In our view, this type of knowledge nicely fits the description that was given to the term *Pedagogical Content Knowledge (PCK)* by Shulman [38] as "that special amalgam of content and pedagogy that is uniquely the province of teachers, their own special form of professional understanding;". Since then, many authors have linked their own definition and articulation to the term Pedagogical Content Knowledge. This might have lead to remarks such as "the slippery and elusive, yet seductive term PCK" [39] or "... the PCK literature is not nearly as tidy as the SMK[1] literature" [40]. Against the background of a wide body of descriptions and definitions of the term PCK we will refer to subject matter- (or content-) knowledge on learning and instruction that we cannot yet decompose into distinct instructional design knowledge,

[1] SMK = Subject Matter Knowledge

subject matter (or content-) knowledge and knowledge of target populations without loss of meaning as '*PCK'*.

Thus, our research in WU projects implies among other things structuring subject matter, finding sequential orderings, collecting and organizing ideal examples, questions and activities that are adequate for learning and testing, including corresponding argumentation. It also implies research into conceptions and misconceptions of students with specific backgrounds and possibilities to capture attention of students in different target groups. To some extent, argumentation in our research is application of guidelines that come from knowledge domains such as cognitive science and educational research [14, 41]. To a large extent, the argumentation in such research includes argumentation about a mapping from general concepts and instructional design guidelines to course structures, learning materials and learning paths that are specific for specific subject matter and specific target groups. Here we define a *learning path* as a sequence of actions or activities suggested to the student connecting a prior knowledge state with a learning goal.

Thus, one task in the type of projects that gave rise to this thesis is actual design and realization of courses and learning materials. Next such projects produce scientific output with respect to design, realization, implementation, use and evaluation of courses and learning materials.

There are quite a number of journals in which outputs of such research can be published. Nevertheless, the volume of such research output at the level of higher education is disproportional to the need to provide high quality education. In most universities little or no publication efforts are aimed at the growth of PCK. Two barriers seem important. Firstly, for the individual lecturers and professors who develop courses and learning materials it is more rewarding to publish within their own discipline than to invest heavily in education. Secondly, we have not found a widely accepted research *paradigm* that fits the combined teaching and researching context of SME's in higher education. The traditionally available research paradigms stem from research in Learning and Instruction or Educational Research, and from related fields such as Cognitive Science. Publishing research output based on design, realization, implementation, use and evaluation research requires a new research paradigm that we will call 'design-oriented'. Therefore, this thesis articulates a framework for design-oriented research in higher education with focus on digital learning materials.

Because of this focus on digital learning materials and because Information and Communication Technology (ICT) is omnipresent in higher education, we provide in the next section, a short review of three ICT related developments that are important for this thesis.

1.3 Information and communication technology in higher education

Firstly, over the last decade, ICT has become an integral aspect of the disciplinary science in many disciplines to such a degree that much research can only be done using ICT. Examples are sciences focusing on the design of processes, sciences such as bioinformatics and geo-information systems, and sciences that rely heavily on large simulation models or large databases. For these scientific disciplines, teachers in higher education who design learning materials and courses necessarily have to take into account an important ICT component.

Secondly, ICT has become important in everyday life of practically every student and teacher as well. In many countries, almost every student who now enters a university has already been using ICT intensively. Consequently, ICT has become a part of the mindset of almost every student. As such, this mindset has to be taken into account in learning and instruction. For instance, some years ago a common advice given by educational advisors was that a ‘search task’ was a good learning task. Now it is more likely that many students expect the computer to carry out search tasks. Also for many students, the effort to walk to the ‘physical’ library and to walk along shelves with journals is now much bigger than accessing an electronic library. Many students expect learning resources to be ‘electronic’.

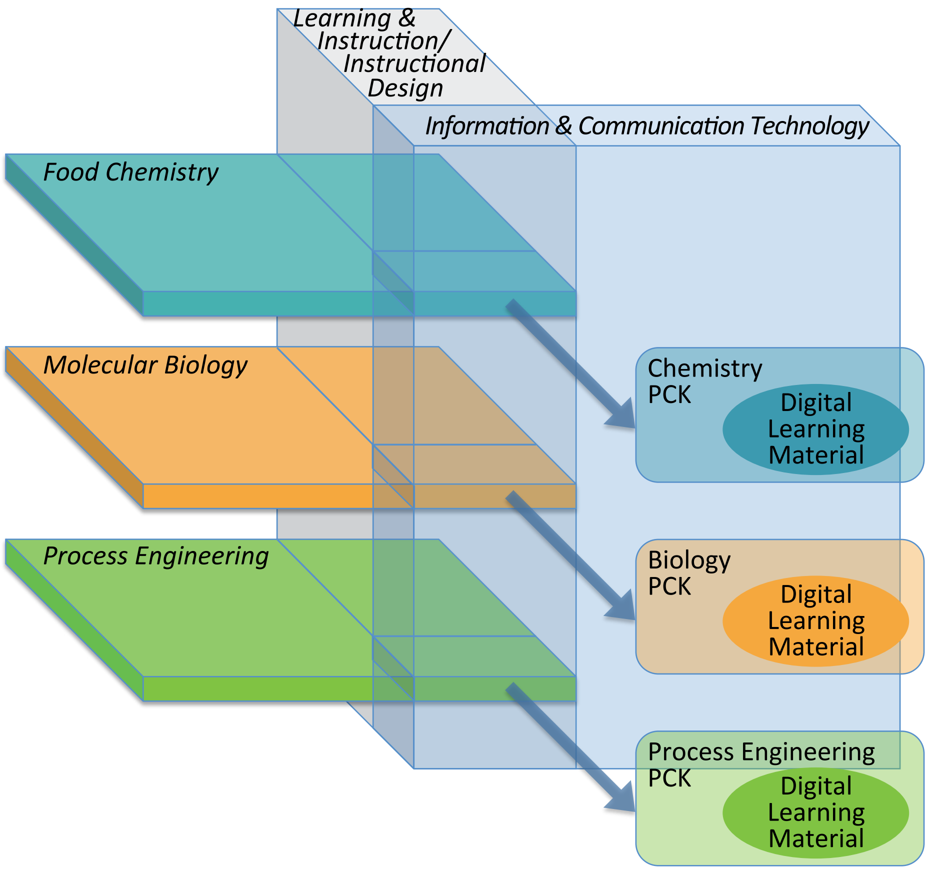


Figure 1-1 The relationship between the primary knowledge domains in the WU projects

Finally, almost every university now supports education with facilities and services such as Learning Management Systems (LMS's), local area networks, internet services and computer rooms or laptop lease programs for students. These facilities are used in

more and more courses in higher education.
An LMS is an information system that at least:

- supports authorized management of digital learning materials;
- provides authorized access to these materials;
- supports electronic communication with e-mail and electronic discussions.

Most web-based learning support is based on learning management systems such as Blackboard [42], Moodle [43] or Sakai [44]. In practice, most instructors in higher education by now at least make their presentations, lecture notes and course information available 'on' an LMS. Actually most LMS's have much more functionality. For instance, most LMS's support authoring and management of quizzes consisting of closed questions such as multiple choice questions, ordering questions et cetera. Since many years also more advanced web-based learning support systems enable enhancement of web-based learning materials based on electronic communications [45-47].

The availability of an advanced infrastructure that supports managing large quantities of digital learning materials and enables easy access to these learning materials is one of the conditions for large-scale use of digital learning materials. The other most obvious condition is that the LMS provides access to digital learning materials.

1.4 Learning materials

Instruction implies supporting learning processes by allocating resources to these learning processes. The most obvious and well-known resources at university level are students, teachers, laboratories, learning materials, publications in scientific journals and resources on the worldwide web.

The tools that are used in research and academic professions form a special class of resources. Examples are spectrometers, autoclaves, pumps, heating and cooling devices, computer algebra software, spreadsheet software, computer simulations and project-management software. Sometimes these tools can be used to develop dedicated learning materials. Students naturally use tools like these because they are necessary in the professional task execution for which they are being prepared. Moreover, students are often required to learn the basic concepts underlying these tools. As such, these tools are resources that are used in learning processes and in general, their use by students will support their learning process. Note however, that these tools are not learning materials in the sense that they are developed for a specific learning goal or class of learning goals.

At a university, learning materials are usually tools that leverage teaching. Good learning materials often reduce teaching load. For instance, certain textbooks are for a large part of the target population sufficient support to achieve the learning goals.

We distinguish two major classes of learning materials.

(1) *Presentational Learning Materials* (PLM), such as videos of lectures, instructional movies, animations, lecture notes, textbooks, and (digital) slide presentations (see for instance [24, 48]).

(2) *Activating Learning Materials* (ALM), such as instructional laboratory set-ups and digital learning materials that require the student to make decisions, enter data and make combinations.

For some students all learning material and all learning resources induce active behavior. However, in this thesis the term ‘activating’ is reserved for learning materials that require the student to take conscious action in order to progress through the material. A textbook with assignments is not labeled as ‘activating’ even though turning a page can be considered an action.

1.5 Digital learning materials

In comparison to traditional learning materials such as lecture notes, textbooks, audios, videos and animations, digital learning materials provide **additional possibilities** to:

- gain and capture attention of students (for instance in digital games);
- stimulate active behavior during learning; in particular learning material that requires the student to think and make a selection or provide some input in order to make something happen, is likely to stimulate activity;
- adapt access and presentation of learning material to specific characteristics of the individual student [49];
- evaluate user actions and provide corresponding advanced feedback on actions of the user;
- enhance the palette of learning materials and learning activities in a university;
- reduce any cognitive load that is not directly necessary for achieving the learning goal;
- realize learning activities such as designing, developing, adjusting and running qualitative and quantitative computer simulations [35, 48, 50, 51];
- provide experiences that would otherwise be too dangerous, time consuming , expensive or impractical and to enable virtual experiments; in the appendix we will explain the benefits of *virtual experiments* in detail;

With respect to costs of design, realization, management and use of learning materials, ICT and digital learning materials provide additional possibilities to:

- reduce the staff/student ratio [52];
- manage and make available multimedia materials at lower costs than non-digital multimedia material;
  in particular the costs of just in time (JIT) presentation of just enough information in just the right form and the costs of providing worldwide access to learning materials have been reduced tremendously;

- design and develop digital multimedia material at much lower costs than non digital multimedia material; a good instance of low cost realization of digital multimedia is the realization of screen-recordings [24, 53, 54];
- enable very short update cycles for learning material;
- relieve academic staff from boring, repetitive and labor-intensive aspects of their teaching task;
- collect data about students and student behavior;
  this offers opportunities for educational research as well as for incremental improvement of learning materials.

In the light of these additional possibilities provided by digital learning materials it is to be expected that output of research with respect to digital learning materials dedicated to specific subject matter will increasingly involve the use of these possibilities. This implies that design and realization of digital learning materials will increasingly contribute to PCK.

1.6 Myths and misunderstandings about digital learning materials

Many authors compare digital interactive learning materials implicitly with a human teacher. Such literature often lists the following 'advantages' of such materials.

Use of digital interactive learning materials is independent of

- time (the student can choose his own time)
- place (the student can work anywhere)
- pace (the student can work at 'his own pace').

Such a presentation obscures many of the actual possibilities of digital learning materials. Until the advent of mobile devices such as tablets it also obscured the fact that learning materials such as textbooks are excellent candidates for learning support when it comes to time and place independence or the possibility to study 'at your own pace'.

Some literature lists 'high costs' as a major characteristic of digital learning materials. The first problem with this statement is that no distinction is made between different cost categories such as realization costs per unit, logistics costs per unit and costs for bringing the material to the attention of the students.

The second problem is that 'high' is seldom defined quantitatively. Indeed when 'high' is defined quantitatively, it usually turns out that the qualification 'high' is based on an implicit comparison of realization costs of digital learning materials with the costs of acquiring other learning materials such as textbooks. The costs of making multiple copies of a piece of learning material and providing access to this material are many times lower than the costs of making multiple hardcopies of textbooks and distributing these to end-users. In other words, at least the variable production and logistics costs are relatively low for digital learning materials. This makes it likely that for sufficiently large numbers of students the costs of digital learning material per student will be much lower than the costs of other learning material per student.

The third problem is that we do not yet have a satisfactory unit to which we can relate the costs such that they also represent a value. Of course, the closest to this might be the price on a transparent market. Ultimately, when we have different learning materials that support achieving the same learning objectives it would also be interesting to compare their production costs. Currently the costs of digital learning materials are reported as costs per screen, per interaction, or per webpage. For instance, Hasebrook & Maurer [46] report that the budget for developing a set of 30 to 50 web pages for one hour of learning is about 30 k€. Hartog et. al. [55] report the costs of a closed question that is based on one *interaction,* in terms of hours of design and realization. According this report, one such digital closed question requires on average 2 hours of design and realization effort. Alternatively, one could measure the design and realization costs per hour of study effort. This is more or less in line with the way courses are defined in higher education: costs are related to lecture hours or hours of study effort and not to achieved learning objectives [56] . The problem with measures like these is that they are not related to the function and the value of the product. In particular, the aim to realize a specific number of hours of study effort does not promote the design and realization of learning material that generates highly efficient learning processes.

1.7 The Food and Biotechnology (FBT) program

In this section, we present a short historical overview of the Food and BioTechnology (FBT) program [57], the main WU projects and their context. The preliminary activities of the FBT program started early in 1997. The actual start workshop was held on January 27 and 28 1998. Shortly after that, the author of this thesis was invited to become FBT program manager. Later he became also project manager in a range of other projects on design, realization, implementation, use and evaluation of digital learning materials.

The direct stakeholders in the FBT program were

- the board of directors of Wageningen University
- successive boards of successive schools of technology, food and nutrition
- those chair holders who were directly responsible for education under auspices of this School
- the chair of applied computer science

Initially an important intention of these stakeholders was to explore new opportunities that ICT might offer for education in a range of courses in fields related to food and biotechnology at Wageningen University. Thus, ICT was from the start an important component of any FBT project.

Soon it was decided that the centre of mass of the investments should be in the design, realization, implementation, use and evaluation of digital learning materials. One reason was that we wanted to capture and embed the pedagogical content knowledge of professors and instructors in digital interactive learning materials, making use of the particular representational possibilities of the latter. This was believed to be useful for use within the university as well as for use at other universities. In practice, part of the teaching and lab instruction experience disappeared for instance due to the fact that the allocation of instructional tasks to staff tends to change over time. At the detailed level, certain instructional tasks are sometimes reallocated to new PhD students and new staff. Embedding PCK in instructions on learning scenarios and in slide presentations or wet-lab assignments seldom provides satisfactory condensation. Such embedding at least requires capturing many subject matter details in some descriptive form such as 'instructions for instructors'. However, 'instructions for instructors' tend to be influenced quite strongly by university-specific variables such as the capacity of laboratories and lecture rooms, specific equipment, the details of the available ICT infrastructure and also the university-wide teaching schedule. As a result, 'instructions for instructors' that are sufficiently detailed tend to be less interesting for the staff of other universities.

1.7.1 Focus on the design of learning materials

Therefore, it was decided to require condensation of PCK in the form of learning materials and publications. This decision was in line with calls for an engineering approach in the context of more basic research in learning and instruction [58]. The decision was also in line with existing practice to develop learning materials within the university. The decision to connect design and realization of learning materials within the faculty tightly with the duty of the university lecturer to produce publications was indirectly inspired by [59] and defended by Johannes Tramper at our university. The initial aim was to embed the details of PCK in learning materials and to present the more generic rules and findings in publications.

1.7.2 Aspects of quality assurance for learning materials

A second position taken, at least initially, was that learning materials should not be university specific. Lecture notes that never leave the university and are never scrutinized by staff of other universities, are likely to be of lower quality than textbooks that are available worldwide and exposed to external reviews. A lecturer investing time in lecture notes for a number of students in the range of thirty to fifty students per year – will tend to invest less time in raising the quality of these notes, than a lecturer who is co-author of a textbook. In case of the latter he will usually also work in a team supported by a publisher aiming at an audience of thousands of students a year. Moreover, a publisher will raise the quality of the process and product by making available considerable knowledge and resources that would otherwise not be available.

The capacity of resources that can be allocated to the design and realization of learning material is clearly related to the number of users of this material. In other words, also for learning materials economics of scale are important. Sustainable high quality of learning materials requires large-scale use.

For the student, it is important that learning efforts are supported by high quality learning materials and bear a tight relation to the exam as well. On the one hand, a course that is based on a textbook authored by experts in the field is to be preferred to a course that revolves around lecture notes of a local university lecturer. This is because such lecture notes are seldom reviewed. On the other hand, in our experience, the relation between the learning material and the corresponding exams is the primary determinant of the student perspective. Thus, many students tend to relate the costs of a textbook or lecture notes with their perception of the contribution of these learning materials to their exam results. If a local exam matches the lecture notes very well, it is unlikely that students will ask their lecturer to prescribe a more expensive textbook. Furthermore, while textbooks often contain considerable redundancy because more redundancy allows for a larger audience, students generally dislike redundancy.

The more specialized a subject is, the less likely it is that the ideal of having every university level course based on high quality learning materials authored by renown experts will be realized. However, the globalization trend of the last decades and the new possibilities of modern ICT lower the barriers to match students anywhere in the world with authors and experts anywhere in the world. Every year it becomes easier to make worldwide connections between students and experts with a common interest. The worldwide web enables us to link more potential students to an SME in a specialized field. The number of biotechnology or food-technology students in the Netherlands may be small, but the number of students interested in these fields worldwide is much larger.

This view on globalization in higher education led to four related propositions in the FBT program defining an ideal situation as follows.

(1) Whenever a university has specific core competences and claims to be leading in specific fields, this university should be the source or at least a major contributor to learning materials in this field.

(2) These learning materials should be exposed to reviews worldwide.

(3) A student anywhere in the world who wants to study this field should want to prefer reviewed learning materials from such a renowned source.

(4) Lecturers from universities anywhere in the world should at least consider using those learning materials in their courses that have been created by experts and have been reviewed with satisfactory results.

Furthermore, not only learning materials should be exposed to criticism from experts, but also conclusions that are drawn from experience with the design, realization, implementation, use and evaluation. This experience should be captured in requirements, guidelines and illustrative learning materials, and be exposed to peers worldwide. This requires us to isolate aspects that are specific for the temporary situation at one university from aspects that are generic for a whole class of situations in many universities and for many years.

1.7.3 Focus on digital Learning Materials

For several reasons it was decided in the FBT program and successive WU projects to focus on **digital** learning materials.

The university administration had already decided to invest specifically in gaining experience with 'ICT in education' and to stimulate that a substantial number of courses would be more 'ICT based'. This decision was partly influenced by the success of LMS's in terms of growth of installed base and by the ubiquitous presence of digital technologies in society. Firstly, given a desire to innovate in a knowledge and information intensive domain and to develop knowledge and information intensive products, it is indeed natural to investigate the possibilities of ICT for higher education. LMS's allow authorized users to make digital learning materials available worldwide and to access digital learning materials worldwide. Thus, an LMS matches the four propositions formulated above. Secondly, digital learning materials can satisfy the requirement that we want to capture learning activities and actions and stimulate the student to make and motivate decisions. Thirdly, ICT had become so integrated in many research areas (think of the importance of DNA - and protein databases), that for several learning goals it is only natural to try to design and develop digital learning materials.

The FBT program resulted in five PhD theses [60-64], 40 small to very small educational innovation projects and a number of related activities. Next, several other projects on web-based learning support and assessment ensued [47, 65, 66]. Currently, much of the learning materials is still in use.

Inspired by a call for design-oriented theories [34], we had often called our work on design, realization, implementation, use and evaluation of digital learning materials 'design-oriented research'. Moreover, since the beginning of the '90-ties of the twentieth century, many researchers began to give attention to design-related research approaches [16, 18, 67-84]. It is against the background sketched in the previous paragraphs that we can now define the scope of this thesis.

1.8 The scope of this thesis

This thesis has a 'double' scope:

- it defines a model for research in design, realization, implementation, use and evaluation of digital learning materials and web-based learning support in higher education and contributes to its methodology;
- it relates the methodological framework to actual research that has been carried out in innovative design, realization, implementation, use and evaluation of digital learning material and web-based learning support in higher education.

During all our WU projects, relatively much time had to be dedicated to discussions on the methodology of design-oriented research both in general as well as in direct relation to design, realization, implementation, use and evaluation of (digital) learning materials in higher education in particular (see also [85]). This involved discussions within the projects and with researchers from other universities as well as correspondence with reviewers of journals. These discussions unveiled a need for a coherent methodological framework. An approach based on just taking a framework from one of the well-known engineering disciplines such as chemical engineering, civil engineering, electrical engineering, information and communication technology or from architecture could not be realized.

It turned out that seminal works like that of [86], [87], [88], [89] were unknown to most of the researchers involved in the WU projects and not easy to understand by these researchers. There was a strong need to clarify the semantics of terms like 'object system', ' inner environment', 'outer environment', 'form', 'context', 'interface', 'modularity', 'component', 'minimal dependency', 'design space', 'constraint', 'requirement', 'guideline'.

The fact that much literature on design does not consistently use design terminology was one of the factors behind this need. While concepts in design literature are quite generally accepted, terminology and definitions are not used consistently across many disciplines. For instance, what one author coins as ‘guidelines’ is by another author coined as ‘ requirements’ and by yet another author as ‘criteria‘ and by yet another author as ‘principles’ . What one author calls ‘context’ is called by another author ‘outer environment’.

As a consequence, much attention had to be dedicated to terminology. Often one term pointed to different meanings and also often one meaning was referred to by different terms. These many-many relationships between terms and meanings were experienced as a serious difficulty in the process of arriving at publishable articles.

Finally, the question what outputs can reasonably be expected from the type of research in our education that we called 'design-oriented' raised many discussions. These discussions and the importance of that question are more recently also reflected in a "Memorandum on design-oriented information systems research"[16].

1.8.1 Results

Firstly, this thesis defines a type of research that we call 'design-oriented research' and presents a unified methodological framework for design-oriented research in higher education. The unification is the result of a systems-oriented theoretical discussion of relevant literature on learning and instruction, knowledge and information systems research, engineering design and a range of other knowledge domains. The concepts and terminology are illustrated with examples from a range of WU projects on design and realization of innovative digital learning materials in higher education and are adjusted based on experience in those projects. In accordance with [16] this implies that the glossary is an essential result. Rather than giving for each glossary term a range of definitions found in literature, the glossary presents for each term one definition or explanation such that the aggregated glossary terms provide a coherent and ‘workable’ set. In research that crosses disciplinary boundaries one must accept that an explanation of a glossary term seldom will convey the many possible subtleties that some disciplinary researchers might feel to be essential aspects of the meaning that they themselves want to attribute to that term. Such subtleties are not always essential and useful in boundary crossing research. The glossary pages provide for a 'language' that allows to describe and discuss methodological issues with respect to design, realization, implementation, use and evaluation of digital learning materials in higher education with reduced ambiguity. Secondly, this thesis presents a view on the characteristics and possibilities of digital learning materials in higher education. Thirdly, this thesis proposes handles for evaluation in design-oriented research in education and reflects on evaluation in several WU projects.

1.8.2 The intended audience

The primary audience of this thesis consists of those people who are directly involved in the design and realization of advanced digital learning materials for higher education, in particular but not exclusively, in natural and engineering sciences. By being more conscious of the essence of design and design-oriented research, members of this audience should be able to share their knowledge in journals such as those that were listed in subsection 1.1.2. It is likely that combining the design and realization of (digital) learning materials with additional efforts to present the results to a scientific community will improve the quality of the designs. Moreover, combining educational duties with research efforts in a way as described in this thesis allows the individual staff member to increase the benefits of these educational duties. A natural next step will be to increase the growth rate of PCK in design-oriented PhD projects.

The intended audience includes educational support teams who support faculty in educational innovation projects. In more and more universities, these teams are also extended to include educational technologists (ET's) who support the realization of digital learning materials. Transforming the work of these teams and team members into design-oriented research as described in the subsequent chapters and publishing the results can raise the quality of the work.

A second audience of this thesis includes researchers in other fields, in particular in information systems research, management sciences and more general those sciences that aim to propose changes in social systems. For them this thesis aims to provide new insights in design-oriented research.

The third audience of this thesis consists of those researchers who agree with the calls for an engineering approach to complement current research paradigms in cognitive science and research on learning and instruction. For them this thesis elaborates a meaning of such an engineering approach.

1.8.3 Overview of the remaining chapters

Chapter 2 explains design concepts without requiring extensive knowledge of other disciplines except for some basic mathematical knowledge. The chapter is partly based on literature research into design in many disciplines such as software engineering, electrical engineering, aircraft design, architecture and instructional design. Here we propose a set of explicit definitions. In this way, it aims to contribute to a shared vocabulary. This vocabulary may improve the ability of team members to work and communicate without much ambiguity. The chapter also provides examples of applying design concepts in the design of digital learning material or in instructional design. It aims to make the required engineering approach explicit in a way that we could not find in literature on educational research nor in DSE journals. The chapter covers relevant concepts and their use in instructional design and design of learning material.

Chapter 3 compares basic research with design-oriented research. Many SME's are primarily familiar with empirical research. Chapter 3 intends to clarify the essence of our model of design-oriented research by articulating both what basic research and design-oriented research should have in common as well as what makes them different.

Research implies publication. In design-oriented research like in all research, any candidate for publication should be subject to peer review. Currently, most of the types of output that one may expect from design-oriented research in education are not sufficiently visible in peer reviewed journals. For this reason Chapter 4 articulates what types of output are potentially valuable and should provide candidates for publication.

The next two chapters discuss strategic decisions in design-oriented research projects that aim to deliver digital learning materials in higher education. These strategic decisions have consequences for the definition of the design goal. In Chapter 5 we propose to distinguish six types of projects based on the focus of innovation and the type of learning goals. In Chapter 6, we discuss the necessity of defining a goal that fits one or more sustainable large-scale use scenarios. We describe scenarios that are supposed to have potential for large-scale use of digital learning materials in higher education. Finally, Chapter 7 discusses evaluation in design-oriented research, reflects on evaluation in the FBT program and identifies opportunities for additional and/or improved evaluation in design-oriented research projects that aim to deliver digital learning materials in higher education.

Appendix: Virtual experiments[2]

Virtual experiments form a subclass of digital learning materials that deserve extra attention. Many investments in laboratory practice are scarcely supported by evidence of their effectiveness and efficiency with respect to several of the intended learning outcomes and based on faulty motives (see for instance [91]). At the same time, the advance of ICT has raised the sophistication of virtual experiments. These can help the student to attain learning objectives for which, until recently, only laboratory practice seemed to be the right option. Moreover, a recent development is that many real laboratories are now much more “computerized” than a decade ago. This means that a student working in a laboratory often is looking at a computer screen and adjusting experimental settings on the computer. In many of these cases, the learning experience of the student would not be different if the feedback that the computer gives came from a virtual phenomenon instead of from a real phenomenon. Providing the same learning experience with virtual experiments would not require expensive laboratory space, equipment and materials. Thus, there are learning objectives, for which virtual experiments enable active involvement of the students with experiments that would otherwise be impossible in the practice of higher education. This is because real experiments would be too expensive, time-consuming [92], dangerous [93] or unethical (such as for instance certain experiments with animals). Finally, in virtual experimental situations, we can in principle control all conditions from an instructional

---

[2] This section has been elaborated in
Hartog, R.J.M., H. van der Schaaf, A.J.M. Beulens, and J. Tramper, *Virtual Experiments in University Education* in *Looking Toward the Future of Technology Enhanced Education: Ubiquitous Learning and the Digital Native*, M. Ebner and M. Schiefner, Editors. 2009, IGI Global: Hershey New York. p. 373 - 393.

design point of view, while in real experimental situations some conditions may be impossible to control. Thus, for more and more institutions, the time becomes ripe to start gaining experience with virtual experiments.

For a first introduction to virtual experiments and their possibilities, a comparison with the flight simulator is helpful. Most readers will probably have a conception of flight simulators: advanced systems consisting of hard- and software that simulate the behavior of a specific type of aircraft in a wide range of conditions to a high degree of perfection.

The flight simulator enables and supports the student to:

- become familiar with the behavior of the aircraft in many conditions;
- practice specific maneuvers in a wide range of conditions many times over at costs that are much lower than the costs of practice in a real aircraft; in particular, the student can practice maneuvers in dangerous conditions or maneuvers that are themselves dangerous without the penalty of irreversible damage on errors;
- link theoretical concepts to practical experience at costs that are much lower than the costs of these learning experiences during real flights;
- learn planning ahead during simulated flights and to make decisions in situations that would create cognitive overload as long as they are unfamiliar.

Finally, the flight simulator allows the instructor to design a special sequence of learning experiences and to adjust this sequence to the ability of the student and to repeat learning experiences without having to wait for the conditions necessary for these learning experiences. This is possible because instructors have full control over the virtual flight conditions.

Most of the learning objectives that can be achieved with virtual experiments in higher education are in one of the categories listed above. In addition, in many disciplines in higher education, we want students to construct mental models of a system and of classes of systems. Finally, we want students to learn the characteristics of the behavior of certain systems and to design experiments. We will now shortly elaborate the relationship of virtual experiments with these four learning goals.

Constructing mental models

In higher education a common learning goal is that the student is able to construct one or more mental models of a specific system that are congruent with accepted scientific models of this system. For instance in case of a system described by a mathematical model, we want the student to construct a mental model that has the same characteristics and includes the same definitions as the mathematical model.

Both virtual experiments as well as real life experiments can provide the student with information that (s)he can use to construct a mental model and check its congruence with a corresponding scientifically validated model. In practice, the process of acquiring information by experimenting must usually be guided for instance by a sequence of questions and/or hints [94]. Theory suggests that the model that drives the computer simulation, which enables the virtual experiments, should be congruent to the intended mental model [50]. For instance, if the intended mental model is a set of differential equations, the same set of differential equations should drive the computer simulation. If the intended mental model is a set of productions (if …then … rules), the same set of productions should drive the computer simulation. This is in line with the classical way of assisting the student in acquiring a mental model of a system based on interaction with the corresponding mental model of the instructor.

Learning to construct mental models within a discipline

In many disciplines in higher education, we want students to construct mental models of specific **classes** of systems. Again the three most obvious interactive options to support this learning are (1) interaction in real experiments, (2) interaction in virtual experiments and (3) interaction with (a mental model of) an SME.

A student who has learned to construct mental models within a specific field is able to construct such a model given a real or a virtual environment. For instance, in many fields in higher education the most essential component of the model is a set of differential equations. For those models, we can test a relevant modeling ability of the student by offering a number of situations in different environments. For these situations we can ask the student to define systems and for each system to define a model in the form of a set of differential equations and to explain the meaning of the symbols [95, 96].

Learning the characteristics of the behavior of a system

An important learning objective for students in the flight simulator is to learn the typical behavior of the airplane in a wide range of conditions and to learn to apply this knowledge adequately. In much the same way, many learning objectives in higher education essentially boil down to learning the typical behavior of a system and learning to apply this knowledge. Focus on this learning objective often implies that we also want to focus on *retention*. For instance, we want that students retain the knowledge about typical Monod kinetics with respect to microbial growth. To assist the student in learning this we can choose to let the student read a textbook, we can tell the student but we can also require the student to do specific experiments in the laboratory or to do specific virtual experiments.

In this case, the role of the experiments is primarily related with the necessity to make an impact that promotes retention. For instance, for some students it may be sufficient to tell them a number of times that they should always be aware of the internal resistance of a voltmeter. For many other students it is believed that the actual experience of being in a situation where ignoring the internal resistance of the voltmeter results in unintelligible outcomes, will have a more lasting effect. This is partly based on the belief that in such a situation the student is actively engaged [30, 97, 98], partly based on the belief that experiences like this gain the attention of the student [98].

Learning to design experiments

Students who learn to design experiments will also have to evaluate their designs. Experience gained by carrying out their self-designed experiments can often contribute to evaluation of these designs. Thus, we should provide students not only with the possibility to design experiments but in particular also with the possibility to carry out their own experiments.. Some virtual experiment environments that allow a student to run the experiments (s)he has designed are described in [2, 50, 92, 99, 100].

Chapter 2 Design concepts for design-related research approaches in higher education

Abstract

In the last two decades attention for design-related research approaches in fields of education and information systems has been growing. This chapter aims to link these approaches to design concepts that are available in disciplines other than cognitive science and educational research. These design concepts are usually defined implicitly or they are defined in a way that assumes experience within the field of a specific engineering discipline. Consequently, these concepts are not directly useful for the educational researcher. Furthermore, design terminology is not used consistently across different disciplines. This chapter presents and illustrates a unified set of basic design concepts that are important for the design component of design-related research in higher education. Finally, this chapter presents a definition of 'design-oriented research' (DOR).

2.1 Introduction

In literature on instructional design and design-related research approaches in education we have not been able to find descriptions or conceptions of design that are clearly linked to the considerable body of literature on design in many other disciplines, in particular engineering disciplines. Explicit calls for an engineering attitude are barely accompanied by references to design literature in other knowledge domains (see also [101]). One reason might be that design literature is often highly integrated with a specific discipline such as architecture, mechanical engineering, electrical engineering, software engineering, user interface design and yacht design. Moreover, examples that are used to illustrate concepts in design methodology are directly related to the field in question. As such, the examples are often difficult to understand for readers outside the field.

As a result, a number of valuable design concepts and corresponding lines of reasoning may be unknown to many instructional designers, educational researchers and to those who design learning materials. This became clear in a research program aimed at the design and realization of digital learning materials for Food and Biotechnology (The FBT program, [57]). Even when the designers of the learning materials endeavored to acquire basic design concepts, their efforts were hampered by the fact that the terminology and the scope of definitions are not the same across a wide range of disciplines.

For this reason, this chapter defines 'design' and highlights and explains the most relevant concepts. In addition, this chapter provides examples that can be understood without extensive knowledge of other disciplines. Furthermore, it applies general design concepts to the design of learning materials.

2.2 Design

2.2.1 Terminology

Design can be a verb that denotes the design process as well as a noun that denotes the result of a design process. The result of a design process is a new entity that is often called ‘*a design*’. This chapter describes both 'what it is *to design*' as well as 'what is *a design*'. For reasons of clarity, in this chapter the term *design* will be used to denote the output of the design process and the term *design process* to denote the process itself.

Essentially, a design is a model of an *intended instantiation* that still has to be realized, implemented and used. Examples of a design are the blueprint for a building or a ‘storyboard’ (see Figure 2-1) for a piece of interactive digital learning material.

A design can also be a *plan*. A plan is a model of a process that still has to be executed. A plan for a course includes a course schedule and a description of the learning activities. Realization of a model (blueprint) of a building means that the building is physically constructed (built). Realization of a plan of a course means that the plan is executed, i.e. the course schedule and the activities are actually executed. The result of a realization is an instance. When the distinction between the model and its corresponding realized instance, such as the blueprint and the building, should be highlighted, the term *design proper* will be used to denote the model (‘the blueprint’) or the plan.

Implementation is: fitting the realized design to its intended context. Because some of the initial assumptions about the intended context of a design may turn out not to hold by the time the design has been realized, implementation may require adjustments of the context. For instance, implementation of an information system often requires training of users. Implementation of digital learning materials in a regular course in a university may require adjustment of course schedules and adjustment of the content of lectures.

2.2.2 Some definitions of 'design' as verb

Literature on design provides a large set of definitions for the verb 'design'. A close look shows that many definitions should be characterized as complementary rather than as inconsistent. A few definitions are sufficient to give the gist of the various meanings of design and to select the most important characteristics of design activities.

Smith and Ragan [23] define instructional design as "… the systematic process of translating principles of learning and instruction into plans for instructional materials and activities. An instructional designer is somewhat like an engineer." From the viewpoint of knowledge and information systems Brown & Chandrasekaran [102] define design as: "… complete specification of a set of 'primitive' components and their relations so as to meet a set of constraints." Dasgupta clearly defines the open problem character of design problems [103]: "In order that a design can be shown to satisfy a set of requirements, all requirements - including those that are conceptual or imprecisely stated - must eventually be transformed into a set of requirements that are entirely empirical....The elaboration, expansion or refinement of the requirements is thus an integral part of the development of a design." Alexander [104] in relation to architecture and spatial planning, describes the process of design as "…. inventing physical things which display new physical order, organization, form in response to function." Dym & Little [20] define engineering design as follows: "Engineering design is the systematic, intelligent generation and evaluation of specifications for artifacts whose form and function achieve stated objectives and satisfy specified constraints." Finally, Vincenti [105] pinpoints decision-making as an essential element in the design process: "Engineers frequently have to make decisions of great practical consequence in the face of incomplete and uncertain knowledge."

Many more definitions can be found in a wealth of literature on design. Most of these definitions point to 'constraints', 'decisions', 'synthesis', 'goals' and 'requirements' as being crucial during design processes. Furthermore, the output of a design process, i.e. *a design*, will have to function within a context. This distinction between (the model of) the realized instance and (the model of) its context is one aspect of what is called a systems science view. Different authors use different terms to distinguish the realized instance of the artifact and the environment in which the instance performs its function. Alexander [104] speaks of 'form and context', Simon [86] of 'inner and outer environment' and Asimow [87] of 'the system and its environment'. Other concepts that are typical for a systems science view such as the concept of 'interface' and the idea that a system can be described as a configuration of other systems will be discussed below. In instructional design the systems science view is particularly explicit in [31].

For the remainder of this chapter, we propose to use a generalization of a definition of Dym & Little [20]:

> A design process is the systematic, intelligent generation and evaluation of specifications for artifacts whose form and function achieve objectives and satisfy specified constraints. In this definition, generation should be understood to include decision-making as well as formulating the specifications for the artifact[3]. The design is the model of the intended instantiation. When evaluation has to imply tests of its concrete manifestation in operation, the design process includes the realization, implementation and use of that concrete manifestation. Finally, the process of stating objectives is also often considered to be part of the design process [83, 85, 103].

[3] Digital learning materials and resources, in particular interactive materials are information systems in their own right. For this reason we define 'artifact' according to Österle, H., J. Becker, U. Frank, T. Hess, D. Karagiannis, H. Krcmar, P. Loos, P. Mertens, A. Oberweis, and E.J. Sinz, *Memorandum on design-oriented information systems research.* European Journal of Information Systems, 2011. **20**(1): p. 7-10. as "constructs (e.g., concepts, terminologies, and languages), models, methods, and instantiations (i.e. concrete solutions implemented as prototypes or production systems). Concrete manifestations of such artifacts can be axioms, guidelines, frameworks, norms, patents, software (with open source code), business models, enterprise start-ups, and much more."

2.2.3 The inputs for the design process

The first major input for a design process is a situation involving a perceived problem or opportunity. This situation is one of the inputs to a design process. For example, there is a worldwide demand for education in bioprocess engineering and food technology and a limited local availability of experts who can provide the required education. In this case, we call the world 'the *object system*', i.e. the system that is the object of our consideration. A perceived problem in this object system is the discrepancy between the demand for education and the insufficient availability of local expert capacity. In this situation, an opportunity is the fact that web technology enables us to make relevant learning material widely available. This in turn, reduces the need for locally available experts and supports the work of the experts that are locally available.

The second major input for a design process is knowledge in the form of a *paradigm*, theories, factual knowledge and *tacit knowledge*.

Box 2-1 Learning Objects

An important category in the literature about digital learning materials is the category of learning objects. *Learning objects* are reusable pieces of digital learning material aimed to support the achievement of 'atomic' learning objectives.
Many other terms with more or less the same meaning such as *reusable content object* (RCO) or *unit of learning material* (ULM) are used in literature.
One of the ideas underlying this concept is that the same 'piece' of knowledge is often necessary and useful in different contexts. For instance, the concept of exponential growth is relevant in the context of compound interest, in the context of cell growth and in many more contexts. It seems a waste of resources to redesign learning material for all these courses. One could imagine that there might be just one or a few pieces of learning material that can be used by anyone who teaches or learns the concept of exponential growth. In case of digital pieces of learning material we use the term *learning object*.
Of course, this ideal is partly inspired by the success of reusable components in almost any hard- and software industry. At the same time the need for something like *learning objects* becomes visible in the common practice of many university teachers who compose slide presentations from tables, diagrams, quotes et cetera and so-called 'readers' from boxes, sections and chapters of different textbooks or articles.
If we were successful in realizing a learning object that adequately helps students to acquire the concept of exponential growth, millions of teachers and students could use this learning object in different contexts. Such large-scale use would free considerable resources for raising the quality of such a learning object.
An important condition for such large-scale use is that any teacher or student can find the right learning object at the right time. This requires metadata. Consequently much effort has been invested in metadata standards .

An example of one of the relevant paradigms with respect to web technology in education is primarily based on 'computer supported collaborative learning' [106, 107]. Another paradigm aims primarily to deliver the right learning object (see Box 2-1) at the right time to the right student or teacher. This paradigm is based on the concept of learning object and learning object repository (LOR) [108]. A learning

object repository is a searchable store of digital learning objects that can be accessed over the internet.
A third paradigm is relatively new and supports an *integrated learning experience* based on balanced roles for learning materials, experts and collaborating learners [46]. Here, a learning experience is called integrated if it does not involve any form of cognitive load due to switching between different tasks or due to switching between the use of different media and if any other effort needed for such switches is negligibly small.

An example of one of the relevant theories could be cognitive load theory [14, 109, 110]. A major assumption in this theory is that humans' working memory capacity is limited. In general, three forms of cognitive load are distinguished: intrinsic cognitive load, which is supposed to have its roots in the nature of what has to be learned, germane cognitive load, which refers to allocation of working memory capacity to construction and automation of schemata, and extraneous cognitive load, which does not contribute to learning.

The knowledge embedded in theories, a shared paradigm and any available factual knowledge is seldom sufficient for making the necessary decisions in design. This implies that the design process will almost always be directed to a large extent by *tacit knowledge* including *intuition*. Several authors attribute an important role in science and design to tacit knowledge [1, 13, 15, 25]. We define *tacit knowledge* as any knowledge that is not yet expressed in some form such that this expression enables sharing this knowledge among different people. According to Simon [86] "most intuitive leaps are acts of recognition". Different experiences imply different possibilities for recognition; therefore, the design process will be partially based on experience of the decision makers who take part in the design process. In instructional design, both experience with teaching for specific target groups as well as experience in the knowledge domain of the subject matter will determine many decisions at the detailed level.

2.3 Models and representational formalisms

Models are at the core of design processes. In this section, we shortly discuss models of the intended context of the instantiation of the artifact, models of the intended instantiation itself and models of the future object system.

2.3.1 Modeling the intended context (i.e. outer environment) of the artifact

The designer of a yacht might decide to describe the yacht's intended outer environment (the water and the air, for instance), in terms of differential equations. Informally we could say that the representational formalism is the 'language' of differential equations. Using this language in order to describe the outer environment of the yacht also implies that the designer has to define the symbols used in these differential equations.

A designer of a course may model the target population by specifying the knowledge, skills and attitudes that members of the target population are assumed to have developed prior to the course. Such a specification will usually be a specification in a structured semi-natural language, e.g. following a pre-defined template.

2.3.2 Modeling the intended instance

2.3.2.1 Blueprints

As the design is a model of what has to be realized, a design must also be represented by some formalism. Well-known formalisms are blueprint formalisms that are used in construction of buildings. A blueprint of a building is subject to a set of representational rules. These rules define what symbols in the blueprint are allowed and what the meaning of each symbol is. The meaning of a symbol in a blueprint is usually explained in a legend.

2.3.2.2 Direct realizations

Alternatively, a design can be directly represented by its realization (i.e. an instantiation that can actually be used in its intended context). In such a case, the design proper is represented by its realization. In other words, no blueprint or plan description is made. In particular, in innovative design, an adequate representational formalism for abstract modeling may not yet have been developed. Much innovative digital learning material is directly realized. This blurs the distinction between design and realization. Such a distinction is desirable because in evaluation it helps us to match specific aspects of the design with specific areas of expertise. Moreover, a clear distinction between 'blueprint' and realization helps to prevent waste of efforts in realizing (parts of) a design that do not make sense. In section 4.3, we will elaborate the benefits of a clear distinction between design and realization.

2.3.3 Modeling the future object system

The future object system consists of the artifact implemented and in use in its intended outer environment. For a yacht, this may be the yacht sailing on the North Sea. The behavior of the yacht depends on the waves and the air and the behavior of the waves and the air depends on the behavior of the yacht.

For a learning object the outer environment consists of the learning management system, the network, the desktop configuration, one or more courses, students and teachers.

Box 2-2 Obvious candidates for representing the design of digital learning materials

In instructional design, traditional visual representational formalisms are for instance flowcharts [23] or concept maps [111, 112] for parts of the design. Merrienboer et al. present a visual design language to support the Ten Steps method [14]. Recently, attention for visual design languages for instructional design has been growing [113, 114]. Alternatively, a ten steps training blueprint is represented as one table per task class and structured text in the table cells.

At a more detailed level, storyboards (see Figure 2-1) are often used as representational formalism for the design of digital learning materials, [60, 115, 116]. For activating digital learning materials, a storyboard is usually a large set of sketches of screen appearances. Such a sketch suggests how a screen dump in a certain state would look. In other words, every sketch represents a state of the system. Arrows connect each sketch to successive sketches. Thus, an arrow represents a possible transition from one state of the system to another state of the system.

Unified Modeling Language [117] is by definition a candidate for representing the design of digital learning materials [118, 119]. UML is both used during the design of information systems as well as for documentation of information systems. Because digital learning materials are essentially information systems, it is actually remarkable that UML representations of digital learning materials could not be found.

Less visual representational formalisms are vignettes, which are brief verbal outlines of a short process or a situation [120]. Many instructional designers rely on short natural language descriptions of planned activities in tables.

The most well-known formal representational formalism at a lower level is Learning Design [121]. Furthermore, the QTI 2.0 specification [122] specifies a representational formalism for modeling interactions between student and computer.

Again, the behavior of the learning object depends on what components of the outer environment do and vice versa. For practical application, modeling the future object system often involves combining different representational formalisms or combining models with widely different scales of time or geometry.

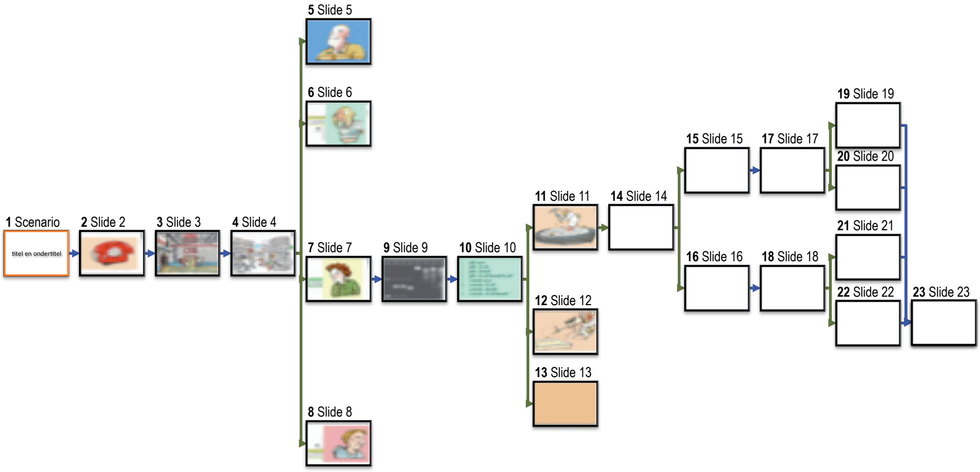


Figure 2-1 An impression of a story board

This storyboard is made with Captivate ™. In this case it is a piece of reverse engineering of web-based learning material in one of the first FBT projects. The SME was J. Verver. The drawings are made by H.Ruitenbeek. [see also 7, 8]. Captivate was not yet available during the FBT projects. Each rectangle represents a sketch of a screen view. The reader should 'read' the storyboard from left to right. A connection between two rectangles presents a transition between two screen presentations. A transition is the response of the system on a user gesture. For instance, in the fourth screen the user can select one of four virtual students. This fourth screen implements a simple interaction with four hotspots, one for each virtual student. Interactions define the relation between the user gesture and the system response.

2.4 Thinking in terms of variables, constraints and a design space

Within the scope of this chapter, it is important that designers will have to make a selection from available representational formalisms for their design. Sometimes, the choice for a specific representational formalism may be obvious and not fit for debate. Sometimes the choice is much less obvious and calls for thorough argumentation. Sometimes it is necessary to develop a new representational formalism. A formalism that is both generic and practically applicable in every possible knowledge domain is not likely to be developed.

Shared understanding of a specific formalism that is directly applicable for parts of the design process requires much effort of all team members. This is a problem in research that crosses disciplinary boundaries. On the one hand, in practice, different disciplines and different design challenges need dedicated theoretical constructs and corresponding representational formalisms. On the other hand, in order to generate at least an abstract shared understanding of design we need a generic abstract formalism. A formalism based on constraints and variables seems suitable for this role for two reasons. Firstly, most definitions of 'design' refer implicitly or explicitly to constraints. Secondly, a formalism based on variables and constraints encompasses many mathematical and logical modeling formalisms that are broadly known and used in science.

2.4.1 Variables and Values

A system is often represented by a set of variable-value pairs in a space defined by the variables. The set of possible values of a variable is called its domain. The initial set that we define for a domain is in general rather arbitrary.

Formulating the characteristics of the artifact to be designed in terms of variables implies defining the dimensions of a multidimensional space. Such multidimensional spaces are well known in engineering disciplines. For instance, Quality Function Deployment [123] is a long standing design management approach using the 'House of Quality' as basic design tool. This 'House of Quality' is a matrix-like structure that relates engineering variables to variables describing user perceptions. For the door of a car, examples of engineering variables are: the force $\mathcal{F}_{\texttt{door}}$ necessary to close the door and the force $\mathcal{F}_{\texttt{door-10}}$ necessary to close the door when the car is on a 10 ° slope facing downwards. Another variable is the road noise reduction due to closing the door: $\mathcal{L}_{\texttt{door}}$. A variable describing a related user perception is the perception of road noise $\mathcal{R}_{\texttt{in\_car}}$. Each of these variables has a domain. For the domain of $\mathcal{L}_{\texttt{door}}$ we might choose a range from 1 dB to 12 dB in steps of 1 dB (i.e. a domain of 12 possible values). A very different quantitative variable is the rating of $\mathcal{R}_{\texttt{in\_car}}$ on a scale of 1 to

5 as given by a test team. The domain of this variable could be defined as {1, …,5}. For the car as a whole, thousands of such variables are defined. They can be clustered into variables that are attributes of a subsystem (such as the door), variables that describe an aspect of a subsystem (such as isolation aspects) and so on. In addition, there will be qualitative (i.e. a non-quantitative) variables. For a car, the material for the door $\mathcal{M}_{door}$ is such a variable. The domain of $\mathcal{M}_{door}$ might be {'steel', 'aluminum', 'magnesium', 'wood fiber'}. The variables $\mathcal{F}_{door}$, $\mathcal{F}_{door-10}$, $\mathcal{L}_{door}$, $\mathcal{M}_{door}$ $\mathcal{R}_{in\_car}$ and thousands of other variables make up a *design space* of thousands dimensions.

Requirements can be expressed as constraints. For instance,

$\mathcal{F}_{door}$ < 1 Newton and

$\mathcal{F}_{door-10}$ < 12 Newton

are unary (i.e. 'one-dimensional') constraints.

Laws of physics and attributes of materials can also be expressed as constraints. For instance, $\mathcal{F}_{door-10}$ will be a mathematical function of the total mass of the door, the center of mass and the friction of the hinges. A mathematical function is just a special type of constraint. Finally, there will be a relation between the total mass of the door and the proportion of each material in the door. Again, this relation can be represented as a constraint.

2.4.2 Examples of variables of digital learning objects

Like the door of a car, a learning object such as an instructional movie is just a small component in a larger system. Both for the door of a car as well as for an instructional movie we can define many variables and domains. Figure 2.2 and Figure 2.3 show some of the variables that describe an instructional movie and their domains.

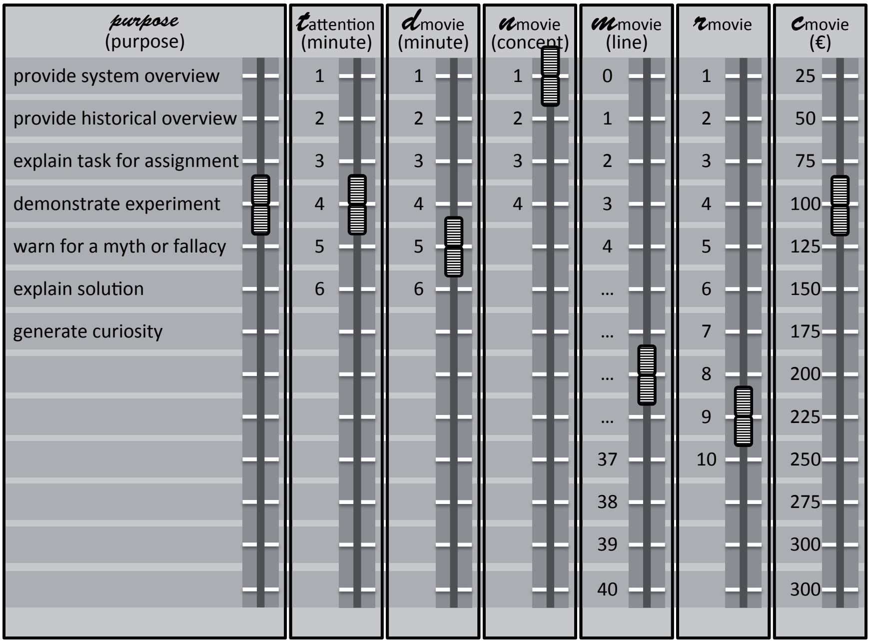


Figure 2-2 A few of the many variables and domains describing an instructional movie

An instructional movie in its context can be described by many variables. Seven of these variables are shown in this figure. Another subset of variables is displayed in a different way in Figure 2.3 A more complete description of an instructional movie would be far too large to fit one page. The variable ***purpose*** is represented in both figures. This variable has a domain of seven predefined values such as 'provide system overview', 'provide historical overview', et cetera. Currently the first slider has selected the value ‘demonstrate experiment’. The variable $t_{attention}$ describes in minutes how long the average student can direct full attention to the movie. Note that $t_{attention}$ must be determined experimentally, or found in literature or estimated. Both knowledge as well as requirements will be represented as constraints. For instance, the value of the variable d_{movie} should not be much larger than the value of $t_{attention}$. This can be represented as a constraint between $t_{attention}$ and d_{movie}. Part of the design task implies finding a combination of values (i.e. settings of sliders) that satisfies all constraints.

Figure 2-3 More variables and constraints describing an instructional movie

	purpose		***line-1***		***line-2***		***line-3***		***et cetera***		***line-30***
domain of ***purpose***	provide a system overview	domain of ***line-1***	"On the screen you see …"	domain of ***line-2***	"…"	domain of ***line-3***	"…"	domain of ***line-4***	"…"	domain of ***line-30***	"From now on you should remember …"
	provide an historical overview		"In the middle of the screen you see …"		…		…		…		"In this movie we have given you an overview of …"
	explain task for assignment		"Please recall that, …"		…		…		…		"This experiment has shown you …"
	demonstrate experiment		"At the end of this movie we will give you an assignment"		…		…		…		"As we have seen, the main differences between <…> and <…> are <…>"
	demonstrate procedure		"This movie will show you how to carry out < procedure >"		…		…		…		"Now start carrying out this assignment."
	warn for a myth or fallacy		"In this movie we will introduce the concept <…>"		…		…		…		"From now on you must always be careful not to
	explain solution		"Please look in the upper left corner"		…		…		…		"This movie was your first encounter with the concept <…>"
	surprise and generate curiosity		"…"		"…"		"…"		"…"		"The main conclusion of this movie is <…>"

Suppose an instructional movie has 30 lines of spoken text. Some of the variables and domains describing such a movie are presented in this table. The first variable is ***purpose***. Possible values of ***purpose*** are 'provide system overview', 'provide an historical overview' et cetera. These values make up the domain of the variable ***purpose***. Each line of spoken movie text is also a variable. A possible value of such a variable is a templates (i.e. an incomplete phrase). For ***line-1*** and for ***line-30*** some of the templates are shown. For instance, one possible value of ***line-1*** is "On the screen you see…". The last column shows some possible values of ***line-30***. In practice, the sets of templates (i.e. the domains) will be larger.

Constraints describe which combinations of values make sense. One set of combinations (i.e. one constraint) is indicated by double cell borders.

Most combinations are not allowed by the double cell border constraint. The double cell border constraint defines only two valid tuples:

{ 'provide a system overview', "On the screen you see …", "In this movie we have given you an overview of …"} and

{ 'provide a system overview', "In the middle of the screen you see …", "In this movie we have given you an overview of …"}.

A second constraint defines the set indicated by bold cell borders. This bold cell border constraint defines three valid tuples.

A typical variable in describing an instructional movie is its duration in minutes: d_{movie}. Initially we might define the domain of d_{movie} as: {1, …,60}
A typical variable of the context of the instructional movie is $t_{\text{attention}}$. This indicates how long we expect that a student in our target population can direct his/her full attention to an instructional movie. For lack of data from literature we might base the domain for the span of attention in a lecture on personal experience and define this domain by {1,…,20}. Here we take one minute as unit as well.

Some additional examples of variables are: the number of concepts that are used in the movie: n_{movie}, the number of lines of spoken text in a movie: m_{movie}, and the rating given by a test team of students: r_{movie}. The latter could be an average of ratings on a scale of 1 to 5 received from members of the test team. The test team could be a sample of students from the target group for which the learning object is being designed. An important variable is c_{movie} : the total production costs of the movie. For each of the variables we will have to define a domain. For n_{movie} a reasonable domain is {1, …,4}. For m_{movie} a reasonable domain is {1, …,50}. For the text of a line, we might also define a variable with a domain. The values in such a domain might be standard templates. In particular, for the first few lines and the last few lines it makes sense to choose from a set of templates. One of the templates for $line_1$ in a movie in the FBT program was: 'On the screen you see …'. Another often used template was 'In this movie we will show you …'. Thus, these two templates would be two respective values in the domain of $line_1$. Experience in WU projects suggests that for c_{movie} a reasonable domain in Europe would be {1, …,300} in unit of 1 €. Note that the definition of many of these domains is rather arbitrary. One designer might decide to define domains that are very much larger and extend the design space in this way, thus allowing for many more potential designs in theory. Another designer might decide to define smaller domains. For instance this designer might limit the number of lines of text to 30 lines, the number of concepts to 3 concepts and distinguish only 12 values for c_{movie} based on a unit of 25€ instead of 1€. Only for some of the domains, it will be possible to provide a definition based on good arguments. Below we will explain that this is an essential aspect of design. For the design of an instructional movie, many other variables will be more difficult to describe than its duration. First, there will be very fundamental variables such as the instructional approach that can be followed in the movie: $\mathcal{I}_{\mathtt{movie}}$, the type of learning objective of the movie: $\mathcal{O}_{\mathtt{movie}}$, and the techniques that are used to direct the student's attention: $\mathcal{A}_{\mathtt{movie}}$.

2.4.3 The design space and constraints in the design space

The *design space* is the space of 'candidate' solutions for the design problem. If we model the system that we want to design and its context in terms of variable-value pairs, then the design space is the Cartesian product of the domains. Let us suppose that the variables for the challenge to design an instructional movie (i.e. $d_{\mathtt{movie}}$, $t_{\mathtt{attention}}$, $n_{\mathtt{movie}}$, $m_{\mathtt{movie}}$, $r_{\mathtt{movie}}$ and $c_{\mathtt{movie}}$) completely describe the design problem (which they certainly do not!). In such a case, the design space defined initially would contain 6 dimensions and 60 * 20 * 4 * 50 *5* 300 = $3.6 * 10^8$ candidate solutions. Domains defined in another way, for instance based on different units, might imply a much larger domain or a much smaller domain. Given the example in the previous section, the design space based on smaller domains would contain 60 * 20 * 3 * 30 *5* 12 = $6.5 * 10^6$ candidate solutions. Many of these candidate solutions however will not satisfy all relevant constraints.

A constraint is a simple representation of a requirement or law or some other piece of knowledge. A constraint defines a *relation* in the mathematical sense of the word. A constraint defines a subset of the points in a space. Such a set of points can be defined by the mathematical description of a line or a plane, by a set of differential equations, by mathematical inequalities or by a table, to name only a few.

Newton's laws can be represented as constraints on forces, masses and accelerations in any engineering design space. The laws that govern human information processing are constraints in any instructional design space. For instance, limits on attention directed to the learning objectives [124] and on working memory capacity [109, 110, 125] might be expressed as constraints. For instructional movies, an experimental estimate for the average attention time in the target population might be 5 minutes with a standard deviation of 1 minute.

Next, a design requirement might be that the duration of the movie must be smaller than the 'average attention time' minus the standard deviation. In other words: $\boldsymbol{d}$movie < $\boldsymbol{t}_{\mathrm{attention}}$ – 1 minute.

Furthermore, we might require that the average number of spoken lines of text per minute is less than 4. In other words: $\boldsymbol{m}_{\mathrm{movie}}/\boldsymbol{d}_{\mathrm{movie}}$ < 4 lines/minute. These representations of knowledge and requirements are just a few of the many constraints that together determine which of all possible movies are satisfactory and which are not.

2.4.4 Finding a solution: constraint satisfaction

Now the core of the design process can be described as a constraint satisfaction problem [126]. To satisfy a constraint implies to find a point that is included in the set of points defined by the constraint. To satisfy a set of constraints implies to find a point that satisfies every constraint in the set of constraints.

2.4.5 Extending the design space

The design process involves not only searching within the design space for points ('candidate solutions') that satisfy all constraints. In practice, we often take actions to restrict the design space as will be discussed below. Designing also involves extending the design space by adding variables, extending the domains of the variables, and adding knowledge in the form of constraints. Thus, we see design as a process not only of constraint satisfaction but also of constraint exploration (see also [127]). In particular, extending the design space is a sub task of the design process.

New technology extends the design space. For instance, information and communication technology (ICT) has extended the design space for learning materials with several new variables in order to describe interactivity and connectivity. An example of such a new variable is the binary variable that describes a dialogue as modal (input necessary in order to proceed) or non-modal (input not strictly necessary in order to proceed). The ability to extend the design space by adding additional variables and/or by extending domains of variables is an important component of creativity. For instance, adding variables that describe the drives of students and adding variables that describe the essence of interactivity extend the design space of learning materials towards educational computer games.

2.4.6 The design space changes with the design goal

Design goals tend to change over time. Consequently, the design space may change in the sense that some dimensions (variables) become irrelevant and some new dimensions (variables) become relevant. In a course on software engineering, a traditional goal is 'the student must be able to design a sorting algorithm given specific design requirements'. Nowadays this goal might be replaced by the following goal: 'given a specific sorting problem, the student must be able to find, select and use an adequate sorting component from one of the available software component libraries'. The former goal will imply proficiency in the use of control structures as a learning goal and certain variables to describe this goal. The latter goal will imply proficiency in the use of component libraries and certain other variables to describe this goal.

2.4.7 Traveling through design space

It is often helpful to describe part of the design process as a process of moving through the design space. Part of the design process can be described as moving from intermediate design to intermediate design until one arrives at a point in design space that satisfies all constraints. This is analogous to part of a voyage. This can be described as a process of traveling from waypoint to waypoint or milestone to milestone until one arrives within the geographic boundaries of a harbor or a camping that was chosen as goal. The boundaries of the goal of a voyage are specified in terms of geographic coordinates in two or three-dimensional space. In the same way the 'boundaries' of the goal of a design and realization project can often be specified in terms of a set of coordinates (i.e. variable – value pairs) in multidimensional space.

2.4.8 One typical output of a design process is a path in design space

A design process produces a path through design space towards a point that satisfies all constraints. This path or parts of this path contribute to the growth of knowledge. In the past, discovering geographical connections was an activity that contributed to our knowledge of this world. In a similar way, other scientists are now mapping metabolic pathways. Design has much in common with such ventures in the sense that it is a search for connections between points in a space. Thus, a course designer tries to find learning paths for students and a designer of a digital learning object tries to find a satisfactory sequence of human-computer interactions.

2.4.9 Design patterns and synthesizing a journey

A journey can often be the result of composing different ‘legs’ and we can compose the solution of a constraint satisfaction problem by synthesizing partial solutions.

Thus, different compositions of train-, ferry- and airplane connections can be candidate solutions to the problem of designing a journey to our destination. A combination of for instance train, flight and taxi ‘legs’ for parts of the journey might be a recurrent pattern in many journeys. A recurrent pattern might be taxi-train-flight-taxi-train-taxi. This pattern might be a solution for the challenge to connect two distant locations while guaranteeing a satisfactory balance between comfort and costs. Combinations of learning activities might form recurrent *design patterns* in a course.

A recurrent pattern in a bioprocess or software engineering course could be the following pattern of learning activities: view-instructional-movie, view-movie-with-assignment, make-assignment, select-a-solution-to-assignment-from-a-set-of-presented-solutions, view-selected-solution, compare-selected-solution-with-own-solution [24]. This pattern is one answer for the challenge to realize an adequate balance between freedom, learning-by-doing and learning-by-example. Patterns like this are called design patterns. Design patterns have been defined a few decades ago by Alexander [88] and since then are being used more and more in other engineering disciplines such as in software engineering [89] and in instructional design [88, 106, 128-131]. A *design pattern* is a reusable configuration of basic components or activities (including parameter settings), which fits a partial design problem.

2.4.10 Other metaphors for designing

The metaphor of the traveler through the design space should not be the only metaphor that supports us in thinking about the design process. In particular, a design requirement or a law of nature in the form of a constraint can also be regarded as a 'rule' along which we can cut or remove part of the design space in order to reduce our search efforts [132].

Note furthermore that a boundary can be a constraint but a constraint does not have to be a boundary. This is because the points that satisfy a certain combination of constraints might be scattered in the design space. This implies that the analogy between the design process and a voyage is not always applicable. In particular, when solutions are scattered we will rather view a constraint as a filter that removes specific points from the design space. Successive applications of different constraints finally leave us with those points in design space that satisfy all constraints.

2.4.11 Reducing the design space

The metaphors of the rule along which we cut off part of the design space and the filter that removes points from the design space, are particularly adequate for focusing on the order in which to apply constraints. Searching for solutions in the design space also involves reducing the design space in order to limit our search. We already suggested that in the design of instructional movies one might assume that it is good if the first line of spoken text just tells what the student sees on the screen. Working with such an assumption is a design decision. In this case it is a generic decision for a whole set of movies. Another generic decision might be to state that the first or second line in every instructional movie should be a line that tells the student what the main learning objective of the movie is. Yet another generic decision might be the decision that the last line of spoken text in the movie always should just recapitulate what the student is supposed to have learned from the movie. These decisions might be based on assumptions with respect to the average student. All in all, many decisions in a design process can be viewed as steps to reducing the design space.

2.5 Design requirements

2.5.1 Defining design requirements

The goals of a design project are specified by design requirements. Ultimately, design requirements should be specified in such a way that it is clear how we can determine if the realized design meets its requirements. Such requirements are called *operationally defined requirements* or *operational requirements* for short. Operational requirements are important for evaluation of design and design processes [83], but they are also needed for a shared understanding of the design goals.

The activity of finding operational definitions itself is an activity that we recognize in other disciplines. For instance, in physics much effort has been invested in operational definitions of 'temperature' and in cognitive sciences much effort has been invested in operational definitions of 'intelligence'.

One of the typical outputs of a design process is a set of operational definitions of design requirements. For the design of courses and learning materials, the learning goals and learning objectives determine to a large extent the design goal. Clearly, learning objectives must be mapped onto operational requirements. In higher education these operational requirements are often based on exam questions.

Design requirements can refer to the design proper. Alternatively, design requirements can refer to the realized and implemented instantiation in operation, or to the new system that is composed of the realized and implemented instantiation and its environment.

2.5.2 Design requirements as to the design proper

Many design requirements will refer to the design proper. The question if the design proper satisfies these design requirements can only be evaluated by inspection. For a house, a simple requirement to the design proper could be that the blueprint shows two bathrooms. Inspection should confirm that the design proper - in this case the blueprint - implies two bathrooms. An example of a requirement to the design proper of a body of digital learning material might be that it should have built-in scheduled moments of active behavior for the student, at least once every 5 minutes. Yet another example is a requirement that the design of an instructional movie focuses on just one learning objective. In some designs, the check if a design proper satisfies its design requirements can be done by anyone. However, in most designs such a check requires one or more experts.

In design, inspection by experts can be much more valuable than experimental tests of the instantiated artifact in operation. Researchers who have been raised in an environment with an experimental focus may not be aware of this. Good examples to illustrate the value of theoretical checks on a design are the evaluation of the design of a nuclear waste facility or of the design of a bioreactor. We are much less interested in empirical evaluation of the realized and implemented design in operation than in an expert evaluation of the design proper. One reason is that waiting for the outcomes of the empirical evaluation will take too long. The same applies for science education aimed to develop secondary school students into life long learners [133]. Moreover, we usually need the evaluation of experts in order to decide if the risk that something goes wrong is acceptable. Finally, there are many other requirements, for which it makes sense to evaluate the design proper. For instance, requirements with respect to the duration of an instructional movie, the number of interactions per unit of time, conformance with a specific instructional theory, adequate coverage of a certain topic, correct chains of inference, et cetera can be checked by inspection. A positive outcome of a check against a set of requirements provided by the subject matter discipline, and by the discipline of learning and instruction has distinct value and is in principle possible without empirical evaluation of the realized design. In literature on (digital) learning materials in higher education we seldom encounter a publication of a detailed expert evaluation of a complete design proper. Even in DSE journals, the level of detail is seldom sufficient for readers to evaluate by themselves a published design proper. Moreover, it is not yet customary to document every design decision including its links with theory or with generally accepted and well documented rules of good practice.

2.5.3 Design requirements that refer to the new object system

Once the design is realized, implemented in a context and set into operation, we have a new object system. Very often operational design requirements refer to variables of this total system. An example of such an operational requirement could be the requirement that the number of customer complaints about a certain electronic payment system must not exceed 0.005 percent of all payments. Another example is the requirement that a student evaluation of certain learning material is awarded a minimal average overall rating of 4.0 on a scale of 1.0 to 5.0 under specified conditions [52, 99, 134, 135]. For learning material that aims to support understanding of a specific concept, we can define 'understanding' in terms of assignments that students have to carry out satisfactory [136]. Examples of such assignments as used in WU projects can be found in [36, 51, 99, 134, 137, 138]. In addition, we can ask students if, in their perception, the learning object supported them in understanding the concept. This is what we did in the following WU projects that aimed to support achievement of such understanding [2, 36, 49, 52, 92, 100, 135, 137-139].

2.6 Design guidelines

2.6.1 Describing the design versus describing the design process

Certain variables that are necessary to describe the design **process** are not in the design space. For instance, during design we will allocate specific resources with specified capacity to design tasks (see for instance [55]). In a project on the design of digital learning materials, one could assign the task of defining learning objectives to a Subject Matter Expert (SME) and the task of designing the human computer interaction to an interaction designer. The design space does not have variables to describe design tasks and resources because these are themselves not part of the solution.

Another typical design process variable is the rate at which the budget for the design project is consumed. Furthermore, there will usually be a set of process variables describing the *load* (i.e. effort divided by time) that is allocated to each resource. Yet another set of process variables describes how much information processing capacity is directed to a specific design variable during the design process. An example of such a design variable is *modularity* (see 2.8.3). A very important set of process variables describes the marginal increases of design variables resulting from one unit of additional effort. Thus, we can imagine an additional increase in the quality of the layout of a screen design by additional efforts invested in improving this layout. Every variable that is part of the description of the quality of learning material, gives rise to a relationship between additional efforts and a change of the value of this variable. Finally, the sequential order of design activities is described by one or more process variables. None of these variables are in the design space but some do have a direct counterpart in design space. For example, while the quality of the screen layout might be defined in terms of variables in design space the additional effort invested in improving this quality is not.

We define *design guidelines* as requirements on the design process variables that secure and guide the design process. Design guidelines can in principle be represented as constraints on process variables and restrict the set of possible decisions during the process. Design guidelines tell us what to do in which stage, what tools and methods to use, how to plan (parts of) the design project, how to communicate, et cetera.

We have seen that design requirements can be compared to coordinates of a berth in a specified harbor that define the goal of a boat trip. Design guidelines can be compared to directions to follow a certain compass heading and advice how to make use of wind, streams and waves along the boat trip. Design guidelines may prevent waste of efforts and errors during the design process. Moreover, design guidelines may be an alternative for design requirements in case it has not yet been possible to define these requirements operationally. Finally, design guidelines allow evaluation of the design process, which may complement and guide evaluation of the final product.

2.6.2 Examples of general design guidelines

There are many general design guidelines that are often assumed to be valuable for any design task. A few examples are:

- distinguish function and technology (distinguish what and how),
- in an early stage look for suitable design patterns (see 2.4.9)
- use adequate computer assisted design environments and tools

- think in terms of systems that are composed of systems
- avoid efforts in the saturation stage of a variable
- identify and assess risks and balance risk-reducing efforts.

Below, we will come back to the 'systems-are-composed-of-systems' view and we will define modularity, flexibility and the concept of interface. We will first shortly discuss the last two guidelines.

2.6.2.1 Avoid efforts in the saturation stage of a variable

For many design variables, the relation between increase in invested efforts and increase in ultimate benefits will reach some saturation stage. If saturation becomes manifest before the corresponding design requirements are satisfied we should seriously consider to redefine these requirements. In practice, it also happens that for a certain variable the requirements are already satisfied while the designers are still investing efforts. Finally, designers are sometimes investing considerable efforts trying to improve values of saturated variables for which a target not even has been defined.

The need to signal saturation implies that the design team at least monitors the consumption of resources and the corresponding progress. It also requires that the design team can estimate how much progress can be expected with additional consumption of resources. Indeed, in software engineering, collection and use of process metrics is considered normal [140]. For various classes of digital learning materials it makes sense to collect data with respect to the amount of different types of resources required for design and realization. For example, in a large project on the design and realization of digital closed questions for natural and engineering sciences in higher education we found that it was virtually impossible to reduce the average design and realization effort for a closed question to less than two hours [55]. We believe that information of this type is essential for improving the process of design and realization of digital learning materials. However, currently it is not usual to publish metrics of instructional design and realization processes and to use such metrics in order to evaluate design and realization processes in education.

2.6.2.2 Identify and assess risks and balance risk-reducing efforts

Implementing a design and setting it in operation always involves a certain risk. In case of possible disasters, this legitimates substantial investments to prevent this risk. For instance, the costs of failure of a long-term national educational policy are very high, and therefore, it pays to invest substantially against failure during the process of redesigning any national educational policy. But also, when we would aim at large-scale use of an instructional movie, it is essential that such a movie does not induce a misconception. Thus, we should guard against wrong presentation and/or wrong application of a certain concept.

In practice, it is important to define guidelines that support us in finding a sound balance between the risk and the costs of reducing this risk. This means that we need practical guidelines for risk identification, quantification and prevention. While such guidelines are standard in many other design environments such as in food technology or software engineering we could not find any such guidelines with respect to the design of learning materials.

2.6.3 Examples of guidelines specific for the design of learning materials

A typical example of a design guideline for the design of learning materials in the FBT program was that we would take a typical professional task from the corresponding field as a starting point for the design of learning material (see [14]). Typical tasks in WU projects were the design of an industrial process, the design of an experiment, the design of a model and the design of an epidemiological study. Another design guideline often used in WU projects was: 'Pay during the whole design process attention to the degree in which the material will sustain the motivation of the student'. An awareness guideline for a SME is 'Be always aware that novices have no idea of typical orders of magnitude of the relevant variables in your field'.

2.7 Requirements and guidelines

2.7.1 Many guidelines in literature are suggestions for requirements

In instructional design literature, many 'guidelines' are formulated that essentially suggest a requirement. Many of these 'guidelines' were applied in the FBT projects. Examples of 'guidelines' for the design of exercises in [141] are

"use prior knowledge of the student as an entry point" (for a learning object);

"include a motivational element";

"force students to take action".

An example for the design of cases in [2] is

"start each case with an interaction that aims at gaining the attention of the student".

Some of the projects also provided 'guidelines' that typically reflected the vision of the primary SME on pedagogy or didactics for subject matter within his field [51, 134] e.g.:

"when students learn to build models of processes at the molecular level, force them to evaluate the biological implications of any modification they make during model building".

These examples of 'guidelines' are actually pointers to requirements. Indeed, a requirement will give direction to the design process and as such will guide the design process. However, why do we not just require that 'the entry point of a learning object must be certain specific prior knowledge of the student' and that 'every exercise must have a motivating element' and so on?

We can think of three implicit intentions. The first is to suggest that the implied requirements are nice but not necessary. The second is to suggest that the implied requirements should be taken into account in early stages of the design process. The reason for this would be that satisfying these requirements in a late stage of the design would involve a lot of backtracking. The third is to indicate that during the whole design process a certain percentage of the design capacity should be allocated to the implied requirements.

These 'guidelines' are different from a guideline such as

> "If you design learning tasks, you need to take real-life tasks as a starting point for design." [14]

From now on, we will restrict the term guidelines to those guidelines that are essentially constraints on design process variables and distinguish them from 'guidelines' as described above.

Our distinction does not mean that guidelines and requirements are independent. On a journey at sea, waypoints and compass headings are not mutually independent and on a journey through design space, design requirements and design guidelines are not mutually independent. Often, there is a well-defined relation between a set of requirements and a design guideline. In such a case, to provide and use both the requirements as well as the guideline implies providing and using redundant information. This type of redundancy, however, has its value. Over centuries of scientific endeavor, redundancy has proven its value in tracking computational or inferential errors. Thus, redundancy in the combined definition of guidelines and requirements will help to prevent the designer from making mistakes. For the same reason a bookkeeper applies crosschecks and a scientist carries out dimension analysis.

2.7.2 Guidelines are sometimes more valuable than requirements

Some design requirements may require very expensive tests. In these cases, evaluation based on design guidelines, if possible, is more appropriate. Guidelines are also valuable because they enable us to prevent 'going of course'. This can prevent waste of design and realization effort but it can also prevent that the final product contains errors. For instance, in software engineering there is a strong belief that following guidelines aimed to prevent errors during realization can be cheaper than the detection and reparation of errors in the final product [142] (see also [143]). In the design of learning materials, it is probably better to prevent phrases that may induce misconceptions in all communication during the whole design process. More generally, for each subject matter field, it is important to have guidelines that help the instructional design team to prevent misconceptions [144].

2.7.3 Some information cannot be included in requirements.

Sometimes it may be difficult to provide an operational definition of each relevant requirement variable. For instance, a well-known guideline tells us that we should take advantage of affordances during the design process [145]. The term *affordance* "refers to the perceived and actual properties that determine just how a thing could possibly be used”. The guideline to take advantage of affordances is very relevant for the design of digital learning materials, but we have not been able to find a satisfactory translation of this guideline into one or more operational requirements.

In conclusion, evaluation of a design process should take into account both design guidelines as well as design requirements for two reasons. If there is redundancy, this redundancy will provide an extra check against inferential or computational errors. If there is missing information along the line of requirements, this information can be complemented by information along the line of guidelines and vice versa.

2.8 Composing systems of systems with interfaces

Many generic design guidelines are based on a ‘systems-consist-of-systems’ view in which every system has an interface with its environment.

2.8.1 Interfaces

The *interface* of a system describes a set of assumptions about the environment of the system and a definition of the function(s) of the system. A function of a system is a description of what a system has to do or what output a system should produce.

For instance, assumptions in the interface of the headlights of a car might include the availability of a positive and a negative terminal with a specific voltage of 12 volts and a socket with a diameter of 23 mm. Furthermore, the interface will specify the light intensity. In practice, the interface of the headlight will be much more detailed. Interface definitions are typical outputs of design processes.

A *module* is a system that can be replaced by another system without inducing the need to impose new requirements on the outer environment. Thus, the interface of the replacing module will be the same as the interface of the original module. A learning object (see Box 2-1) is a module of digital learning material with an interface that has one or only a few functions and few assumptions. Moreover, the assumptions will have a small scope. For a learning object or a paragraph in a textbook, the most important assumptions refer to the prior knowledge of members of the target population. In other words, it is necessary to make explicit assumptions about the prior knowledge of the student. In instructional design, the knowledge described in these assumptions is referred to as *prerequisite*. Furthermore, the most important function of a learning object will usually be to support students in achieving a learning objective.

Given one interface, there is in general more than one way to define the internals of the module such that its realization can execute its functions. This allows us to redesign the internals of the module without having to change the outer environment of the module. For instance, we can improve the design of a coffee machine as long as it makes coffee and we do not change assumptions with respect to an electric power outlet and a supporting surface et cetera. Alternatively, we can improve the design of a lecture as long as it supports the same learning objectives and we assume the same prior knowledge et cetera.

2.8.2 Minimizing dependence on outer environment

The interface of the headlight will not include assumptions about tire pressure. Thus, if we want to change the light bulb, we do not have to worry about the pressure in the tires. Reversely, if the tire pressure changes we do not have to worry about the functioning of the headlights.

More generally: an interface should assume not more than strictly necessary. In other words, the dependence of a system on its outer environment should be minimal. The less a system assumes about its outer environment the less changes in the outer environment will influence its operation. The less all subsystems that make up a design depend on each other the less a change in any of the subsystems will influence the operation of the system as a whole. Thus, the interface of a learning object that must support students to learn when to use gel filtration in a product purification process, will contain the assumption that the prior knowledge of the student includes a mental model of the gel filtration mechanism. The interface of this learning object does not have to contain, and therefore should not contain, the assumption that the student knows about specific properties of different gels [146]. The concept of interfaces and requirements with respect to interfaces provide excellent handles for expert reviews.

2.8.3 Modularity

The number of functions in the interface should result from balancing the costs and the benefits of modularity. Designing a separate module for each function implies defining many interfaces. Designing modules for larger clusters of functions or functions with a large scope, will often require less design effort because fewer interfaces have to be defined [147, 148].

The modularity of a system increases with a decrease in any of the following factors: the number and scope of functions and the number and scope of assumptions in the interface of any of the subsystems. Maximum modularity of a system would require that each subsystem's interface with other subsystems has not more than one ‘basic’ function and has no assumptions. Given a certain design we can increase its modularity by reducing the number of assumptions in some interface(s), by reducing the scope of assumptions in some interface(s) and by designing more subsystems with less functionality per subsystem.

In practice, searching for a highly modular design often requires considerable effort. Thus, the benefits of modularity must be balanced against the costs of designing modular. For most designs, including the designs of learning materials, the benefits of a higher modularity are increased flexibility, increased transparency and increased number of potential users. Furthermore, in case of a failure to satisfy a specific requirement, modularity may help to localize the source of that failure (i.e. to attribute that failure to a specific subsystem).

Not all authors in instructional design place heavy weight on modular design. For instance in [23] the concept of modularity is practically absent. When the concept of modular design is more explicit [31, 149], the level of detail is still not defined precisely. Furthermore, arguments that are important in finding a balance between costs and flexibility or costs and modularity do not receive much attention. The specification of interfaces of digital learning materials and learning objects as well as the specification of the necessary modularity requires deep involvement of SME's. We consider it to be a typical aspect of the development of pedagogical content knowledge, but also an aspect of budgets and economics of scale.

2.8.4 Granularity

Granularity of a module is loosely defined as the average size of its learning objects. Granularity is not the same as modularity. Granularity refers to the size of modules respectively learning objects and modularity refers to the scope and number of the functions and assumptions in the interfaces. Reducing the scope of the interface of a learning object by removing some prerequisite knowledge might make it necessary to include information in that learning object. This would be necessary in order to help the student to acquire or construct the knowledge that has been removed from the prerequisites description. Consequently, reducing the scope of the interface might involve increasing the size of the module. The size of a module could be defined in terms of the time needed by the average student in the target population to achieve the learning objectives of the module. A unit for this is one hour of study.

In short, the discussion about size of learning objects (the 'granularity issue', see for instance [85]) is related to the discussion on modularity, but granularity should not be confused with modularity.

2.9 Design and decision-making

2.9.1 Decision-making is an integral element of any design process

At many points during the design process, a set of options will have to be generated among which to choose. In terms of the abstract model of the design space discussed above, options may be generated by defining variables and by defining their domains. In practice, the definition of the design space is usually rather implicit. Furthermore, progress through the design space or successive application of constraints as filters suffers almost always from lack of resources such as time, data, knowledge and computational capacity. If we know for sure that every instructional movie of more than four minutes is too long to be useful, this knowledge constrains our design space. Actually, we do not have such knowledge. It might well be that the average maximum time span of attention for students anno 2010 in Amsterdam differs from that in Wageningen, Graz or St Petersburg or from that anno 2005 in Amsterdam. Nevertheless, we can still decide not to make any movie longer than four minutes. Alternatively, we might distinguish different types of movies and decide on a maximum duration for each type. Any design process, including instructional design and the design of learning materials, requires many decisions.

Implicit definition of sets of options (variables and their domains) and insufficient resources raise the problem of evaluating such decisions. Anyone who has to learn how to make decisions, will have to learn what distinguishes one decision from the other in terms of which decision is better and why.

2.9.2 The quality of a decision

We assume that the quality of design processes is at least partially determined by the quality of the underlying decisions or decision-making processes. There seems to be no general agreement on the question how to judge the quality of a decision, but the majority of researchers would prefer to derive the quality of a decision from the quality of the decision making process [150]. Following this line of reasoning, we should derive the quality of a design from the quality of the design process. In other words, evaluation should include evaluation of the design process. This would include measurements of process variables and comparing measurement results with design process guidelines. In this line of reasoning, a failure of the realized design in operation does not necessarily imply that the design was bad. Moreover, in this line of reasoning, a positive answer to the question if the goal can be achieved at all, does not necessarily imply that the design was good.

2.9.3 Allocation of resources and capacity involves decision-making

One of the criteria for a good decision-making process involves the question: has this decision been taken with the use of adequate resources and the right amount of these resources? In this question, the 'right amount', i.e. the right capacity, implies sufficient and not too much in relation to the possible consequences of the decision's execution. For instance, the decision-making process with respect to what background color should be used for a digital learning object should not justify the same investment as the decision which learning management system to implement. On the other hand, when the same operational decision has to be made daily or when a decision is a decision with consequences for all learning objects to be developed over the next several years, the consequences of this decision can merit a large investment. Both decision-making processes as well as design processes are by definition constrained in time and budget. Therefore, applying too much resource capacity to one decision implies applying insufficient resource capacity to another part of the process.

2.9.4 Working with incomplete information

In many typical design disciplines, such as management sciences, engineering disciplines and in the design of learning materials, researchers work in a predominantly goal-driven manner. The designer of new learning material, a new course, a new organization, a new business process or a new boat will have to make many assumptions as a substitute for missing information. A design process that involves completing information does not exist. **A designer who tries to make information complete, will never produce a design**.

2.9.5 Uncertain information

It would be useless to restrict design processes to processes that use only information of which we can be certain. Theory and practice on how to handle uncertain information in decision-making and design are an integral part of design. For example, there are uncertainties with respect to the future context of the learning materials that we design. We do not know when the administration of each university involved will select another learning management system. We also are not certain about the stability of governmental policies with respect to accreditation of eLearning in higher education. The most important uncertainties for the designer of learning materials are related to the prior knowledge of the students in our target population. This holds in particular for learning materials aimed at a worldwide target population in higher education.

2.9.5.1 Uncertain information about prior knowledge

Any design that has to be used by people will imply assumptions with respect to prior knowledge of the intended user. This holds in particular for digital learning materials.

Every instructor is familiar with consequences of assumptions with respect to students' prior knowledge that turn out to be wrong. In a live tutorial, a good instructor can usually make some adjustments. In textbooks so-called 'boxes' are often used to support those members of the target population who may lack some prerequisite knowledge. For digital learning objects in learning management systems, there are comparable options to hedge against the consequences of incorrect assumptions about the students' mastery of prerequisite knowledge and skills.

Instead of hedging against consequences of incorrect assumptions, we can also build adaptive systems. For instance a system might be adapted in such a way that it offers to a student with little relevant prior knowledge more presentations and interactions than to a student who has already mastered all or most prerequisite knowledge [49].

The design and realization of all learning materials rests heavily on assumptions with respect to the prior knowledge of the students in the target population. One way to get this information is to teach a sample of the target population and try to find out the prior knowledge of the students in this sample. This is what good textbook authors do and this is what many developers of learning materials do. Another way is to use online reactions and questions of students in order to adjust and extend the material 'on the fly' [46, 151].

These approaches to handling uncertain information should give our design a certain degree of *robustness*. A robust design is a design that is insensitive to variance of values of variables of its intended context.

2.10 Innovative design and routine design

Designs as well as design processes can be characterized along a dimension that runs from ‘rather routine’ to ‘highly innovative’. Vincenti calls the latter 'radical design' [105] and Dasgupta [152] calls it ‘inventive design’. In contrast, Dym [153] distinguishes 'creative design', 'variant design' and 'routine design'.

In general, innovative design can be regarded as a precursor to routine design. Innovative design involves much more intelligence of unknown territory. As such, innovative design is also much more a learning process. This becomes especially apparent in the role of prototypes and the characteristics of prototypes.

2.10.1 Prototypes

In routine design, a prototype is the first realization of a well-described design proper that is the output of the design process. In routine design, testing a prototype is not expected to lead to major changes in the design. In innovative design, the design proper is often not so well described or even absent. Innovative design often suffers from lack of an adequate representational formalism. Consequently, prototypes are relevant much earlier in the design process. In innovative design, prototypes support stakeholders in understanding new functionality and new concepts. In innovative information systems design, methods in which a design and a realization project are regarded as a learning process is often called *evolutionary development* or *dynamic development* or *agile methods* [154-156] . The design of information systems usually involves opportunities to define very abstract goals, which translate into very abstract requirements. Making these abstract goals and the corresponding requirements more tangible by demonstrating the prototypes, has proven to be very important.

An aspect of evolutionary development/agile methods is *rapid prototyping*. This means that a series of prototypes is developed in short design & realization cycles. This helps the users to understand the functionality of the system to be delivered and it helps the developers to understand what users really want. For many types of systems in information systems development, rapid prototyping is supported with adequate tools. Moreover, in contrast to a prototype of a ship, an airplane or a machine, a prototype of an information system does not consist of expensive material constructs. Several authors have suggested that design and realization of digital learning materials is likely to benefit from an evolutionary approach [85]. However, for the specific area of web-based digital learning materials in higher education, adequate tools for rapid prototyping were not available within the period of the FBT program.

2.10.2 Proofs of feasibility

Evaluation of the operation of a realized and implemented design might prove that it is possible to arrive at the design goal. We call this a *proof of feasibility*. A proof of feasibility shows that it is possible, at least once, to realize a point in design space that satisfies all constraints, which define the design goal. This set of constraints represents the union of operational design requirements, laws of physics, laws of human behavior and laws of society. In terms of a classical empirical research approach [157], this means that one counter-instance has been found for the hypothesis that it is not possible to realize a point which satisfies all constraints. Therefore, the 'impossibility' hypothesis should be rejected.

2.10.3 Successive tests in innovative design involve changes in many variables

During an innovative design process, designers will seldom move step by step along one dimension at a time. Rather, just like a traveler in three dimensions, designers tend to move in several dimensions 'at the same time'. Most design challenges involve a design space of hundreds or more dimensions. Budgetary and time constraints on the innovative design process do not allow designers to realize and test a candidate solution after each one-dimensional step. In practice, designers have to change the value of many variables between tests of successive realizations of candidate solutions.

For instance, at a university, redesign of a course will often involve many variables. We do not have the budget and the time for stepwise changes and tests. Moreover, in particular in innovative design, certain variables cannot be changed without changing other variables at the same time. For instance, the difference between a textbook presentation of a concept and a video presentation of a concept cannot be expressed in terms of just one variable. It is seldom possible to attribute a change in properties of the new object system (i.e. the realized design including its context) to a change of the value of a specific variable. Russell [149] suggests that for a small module with only one function it might be possible to attribute a specific effect to that module. However, the example of the instructional movies in section 2.4.2 makes clear that such a specific small module will still be based on decisions with respect to many variables. Apart from this, we would like one decision to be applicable to a whole set of screen-recordings.

2.10.4 Usually the definition of a goal implies many solutions

Any position within the boundaries of the intended campsite, or any berth in the intended harbor, satisfies the goal definition of a journey in our physical environment. In the same way, any combination of values of the variables of the design requirements that satisfies all requirements defines an endpoint of a journey in design space. Although the one best place at the campsite or the one best berth in the harbor can sometimes be defined, very often the goal of our journey will not be defined as one position at the campsite or one berth in the harbor. Typical for innovative design is that during the design the requirements will become more articulate and that many solutions will satisfy the requirements. Thus, the specific choice for the campsite or the harbor can change during the design process or can be specified later in the design process. Again, this is in line with the concept of design as constraint exploration (see [127]).

An author who designs learning materials – say with the objective to support acquisition of quantitative modeling skills in molecular biology [35, 51] or the acquisition of design skills in bioprocess engineering [146] – tries to satisfy a number of constraints. In general, it is likely that many approaches to support students in acquiring quantitative modeling skills or design competencies satisfy the constraints. At the same time, it is not likely that the stakeholders will succeed in defining an object function that defines the direction in which to optimize the learning material.

Innovative design problems are not defined as optimization problems. Problems that can satisfactorily be described as optimization problems have to include an objective function that is agreed upon by all actors. In innovative design, the choice of one particular design that satisfies all constraints is not based on a formal proof of optimality. Usually, many candidate solutions are discarded without any formal reason to discard them.

2.10.5 Design is not theory building

We regard theories as coherent constructs of hypotheses about facts and relations. The function of most design activities is neither theory building nor hypothesis testing. A failure of a design can seldom falsify a hypothesis.

A designer uses knowledge from many fields and combines 'pieces of knowledge' from many fields. If a design does not meet the requirements, this failure can rarely be attributed to one specific 'piece of knowledge' among all other 'pieces of knowledge'. In particular, this failure can seldom be attributed to one constraint or to one variable.

2.10.6 Design and comparison

2.10.6.1 Comparison with an initial state

Design is inherently related to comparison. When we define a goal, we define a design space and we implicitly state that the design process is not yet in a goal state. This suggests two sets of measurements. One set of measurements should ascertain that the goal has not yet been reached before the realized design has been set into operation. Another set of measurements should tell us if the goals have been reached after the realized design has been operating for a certain time. For instance, in education it is often necessary to do a pre-test and a post-test [158].

In practice, it may occur that we know already were we are before we start and decide that the costs of the first set of measurements (for instance the pre-test) will not be in balance with the resulting information. For instance, in most FBT projects we thought it extremely unlikely that the group of students already had achieved the learning objectives before they started to use our learning materials.

Apart from the fact that measurement of the initial state sometimes will provide little information, such measurements may also change the object system. For instance, students might learn from a pre-test.

Often, part of the definition of the goal state will be based on variables that do not yet have a value before the operation of the design. For instance, it does not make sense to ask students if they appreciate certain learning material before they have encountered the learning material.

Thus, there can be case studies in which there are good reasons not to measure certain variables in the initial state. Nevertheless, we do compare two states: one implicitly measured and one explicitly measured.

2.10.6.2 Comparison with a norm

When the stakeholders require as part of the goal definition that a certain variable should at least have a certain minimum value this will often imply an **indirect** comparison with values that have been achieved before. For instance, when stakeholders require that 80 % of the students in a course will be able to carry out a specific task with success, they essentially say that this norm makes sense. This is part of the process of defining design requirements (i.e. requirements engineering). Depending on the importance of this variable and this norm, we will invest more or less effort in supporting this norm with arguments. Often, the stakeholders will have made an implicit or explicit comparison with other data.

For transparency, it is important to separate the argumentation that underpins the norm from the process of measuring the values of the variable and interpreting the measurement results.

2.10.6.3 Optimization

In a design process, after the innovative stage, we may arrive in a stage where it makes sense to change or add requirements in a way that implies incremental optimization. We have arrived at the camping and on successive nights we sleep at different locations looking for the location that is most quiet during the night. Or, we have an instructional movie that has been proven to explain a certain concept very well and we now select different voices in order to find out which voice is most appreciated by the students. Such a question related to one aspect of the design sometimes may induce a series of measurements based on changes of one or a few variables. Next, we can compare the resulting data.

2.10.6.4 Comparing design goals

In practice, it may occur, that different designer/developer teams have aimed at the same goal and produced different designs and corresponding instantiated artifacts that satisfy the same overall goal. Comparing these designs may reveal that it makes sense to redefine the goals for these designs as different goals. This might induce a discussion to distinguish the redefined goals in terms of one goal being better than the other goal. Ultimately, this implies that the stakeholders will have to compare the design goals.

2.10.6.5 Different designs and instantiated artifacts that both do not satisfy the goal

Alternatively, different designer/developer teams aiming at the same goal may have produced different designs and corresponding instantiated artifacts that do **not** satisfy the constraints that define the goal. If the teams agree on an operational definition of the distance of each of the realized designs to the goal, this usually induces the desire to find out for which of the designs the distance is smallest. A study to this end might be called ‘a comparative study’.

2.10.6.6 Excluding alternative explanations for goal achievement

In some cases, it may be necessary to exclude the possibility that the goal is achieved by the internal dynamics of the object system or by outside influences. For instance, some patients return to a state of health without treatment or some kids acquire certain knowledge without schooling. In such cases, we want to do measurements on two object systems, one that includes the realized and implemented design in operation and one that does not. We can call this 'a comparative study' as well. In this case, the comparative study is a way to ascertain that the achievement of the design goal can be attributed to the use of the realized and implemented design.

When alternative explanations are extremely unlikely, a comparative study for this purpose would just be a waste of resources. For instance, in case of using a rocket to put a man on the moon, we cannot think of the man arriving at the moon without the rocket. In a case like this, we do not carry out a comparative study. Thus, there can be situations in which a comparative study for answering the attribution question would not be balanced by an urgent need to exclude alternative explanations. For this reason, we did not carry out measurements in a control group in our WU projects. Of course, any statement that there is no alternative plausible explanation for the achievement of the design goal is theoretical and requires expert evaluation.

2.10.6.7Implementation in different contexts

The interface of a design includes a set of explicit assumptions with respect to the outer environment or context of the design. Once there has been a proof of principle in a case study of implementing the realized design in a specific context we will want to implement the realized design in other contexts as well.

When the realized design fails in a context for which all explicit assumptions hold we must conclude that there are implicit assumptions that do not hold.

When the realized design is successful in a context in which a certain explicit assumption doesn't hold we must find out if we can drop that assumption from the interface or if the assumption should be adjusted.

Because we are in fact comparing case studies of the operation of the realized design in different contexts we can call such studies 'comparative studies'.

2.11 Design-related research methodology

Over the last 20 years the amount of literature on 'design-based research', 'research-based design', 'design-oriented research', 'design research', 'evidence-based design', 'design experiments' et cetera has been growing, in particular in relation to education and to information systems. In section 1.8 we listed a set of references with titles and abstracts that refer both to 'research' and 'design', which at first sight seem to indicate a relevant relationship with the type of publications that we produced in our WU projects. However, based on a review of this literature [159] we found that such a relationship is very weak. With one exception, which will be explained in subsection 2.11.2 , we decided that incorporating paragraphs on such research in this thesis would be confusing. Moreover, it is our belief that this literature is not in the field of interest of the intended audience of this thesis.

The type of research that we want to define and articulate is oriented towards the delivery of a design. We have called this type of research 'design-oriented research'[4].

[4] We already indicated earlier that we encountered numerous many-many relationships between terms and their meanings in various bodies of literature. Thus the reader should be warned that the term 'design-oriented research' does not have the same meaning in every publication.

2.11.1 Design-oriented research

We define *design-oriented research (DOR)* as research that primarily aims to produce an innovative design and applies typical concepts from design methodology.

The primary research questions in design-oriented research are:

1. what are, in a specific real life context, goals that make sense and why;
2. how can these goals be articulated in terms of measurable quantities;
3. is it possible to achieve these goals;
4. if so, how.

One of the typical outcomes of a design-oriented research project can be a set of requirements that define a goal that makes sense. Another example of a typical type of outcome of design-oriented research is a proof of feasibility. In Chapter 4, we will explain in more detail what design-oriented research in education should mainly produce as its output.

2.11.2 Design-oriented research and design-based research

Design-**oriented** research in education as defined in this chapter is not the same as design-**based** research in education [18, 160]. Bell, Hoadley and Linn define *design-based research* as follows: "Design-based research in education includes testing theoretical accounts of learning, cognition and development by using theory to design or engineer instruction and conducting empirical studies of the instruction in a naturalistic setting." [18]. Contrary to this, our definition of design-oriented research does not include testing theoretical accounts of learning, cognition and development. This is one essential difference between design-based research and design-oriented research. Another difference is that de facto "design-based research is surprisingly quiet about design" Goodyear [101], while articulation of design in design-oriented research as defined by us is very central.

Furthermore, Bell, Hoadley and Linn indicate that “design research includes *compelling comparisons* in which two forms of innovation are enacted under otherwise similar conditions." [18] (Note that the authors introduce a new term ‘design research’ which seems however to have the same meaning as design-based research.) The authors refer to several chapters in [133] for examples of compelling comparisons. For instance, E.A.. Davis [161] compared the effects of generic prompts and the effects of directed prompts with respect to promotion of productive learning outcomes within a given digital environment [162]. This digital learning environment was already designed and developed and in operation (see also [163]).

Finally, design-based research also includes comparative studies in which the same design is implemented in different contexts.

When the costs of a comparative study are well balanced by the information provided by this study a comparative study might be included in a design-oriented research project. However, the inclusion of comparative studies is not a requirement for design-oriented research projects.

2.12 Concluding remarks

Gardner [58] called more than two decades ago for a coupling of inquiry to development of resources in education. About ten years later Robert Glaser [164] called for a consciously chosen engineering attitude.

An engineering attitude requires linking to engineering concepts. At the same time such linking is not a trivial operation. On the one hand desk research of existing design literature has unveiled an implicit conceptual framework common to most disciplinary design-oriented research. On the other hand, the implicit conceptual framework for design-oriented research in engineering disciplines is obscured by many terms and definitions that are not consistently used across disciplines or even within disciplines.

This chapter aims to provide a more explicit and more consistent set of concepts, definitions and examples. Furthermore, this chapter has provided an initial definition of design-oriented research and explained what can be expected from design-oriented research and what not to expect. This is important because wrong expectations with respect to the implications of an engineering attitude will direct researchers to a goal that cannot be attained.

Chapter 3 Articulating design-oriented research against the background of basic research

Abstract[5]

Design, realization, implementation, use and evaluation of innovative learning materials in higher education require a design-oriented research (DOR) paradigm. In higher education, such DOR will often cross disciplinary boundaries. In particular, such research will require contributions from subject matter experts. Many subject matter experts are researchers who have been raised implicitly with 'a basic research or hypothesis-testing research paradigm'. This chapter aims to clarify the DOR paradigm by discussing characteristics that DOR should have in common with basic research and by highlighting in which way DOR will be different from basic research. The characteristics and differences are illustrated with examples from cognitive science, research in education and projects on design of digital learning materials.

[5] A version of this chapter has been published as
Hartog, R.J.M. and A.J.M. Beulens, Digital Learning Resources, Design-Oriented Research and Basic Research in Proceedings of World Conference on Educational Multimedia, Hypermedia and Telecommunications 2012, T. Amiel and B. Wilson, Editors. 2012, AACE: Chesapeake, VA. p. 2941-2950.
This led to an invitation to submit an expanded version to International Journal on E-Learning (IJEL) - Corporate, Government, Healthcare, & Higher Education. This expanded version has been sumbitted as: Hartog, Rob.J.M., Johannes Tramper and Adrie.J.M. Beulens, Digital Learning Resources, Design-Oriented Research and Basic Research

3.1 Introduction

3.1.1 The need for a design-oriented research approach .

Design, realization, implementation, use and evaluation of innovative learning materials imply a type of research that we have called design-oriented research (DOR). In Chapter 2, we have defined DOR as research that primarily aims to produce an innovative design and applies typical concepts of design methodology.

Designing innovative learning materials will require a knowledge from a specific subject matter knowledge domain but also from many knowledge domains that are not covered by one or a few disciplines. Many researchers involved in the design of such learning materials in higher education will have been raised with a basic research - or hypothesis-testing paradigm. These researchers constitute the primary intended audience for this chapter. Clarification of a research paradigm needs to be more than just providing definitions of its constituting concepts. Firstly, not every concept can be defined unambiguously. Secondly, it is important to highlight differences with other research paradigms. In this chapter, the differences between DOR and basic research are highlighted and illustrated.

3.1.2 Experience with design of innovative digital learning materials

This chapter is based on design literature in many disciplines, including instructional design, and on experience in the Food and Biotechnology (FBT) program [57]. The FBT program was aimed at design, realization, implementation, use and evaluation of innovative digital learning materials. The experience gained in this program is particularly relevant because the program was carried out in an environment in which the empirical approach is normative. Many of the methodological questions that surfaced in this program were implicit requests for a comparison of the reality of basic research and the reality of DOR.

In order to clarify the essence of DOR, this chapter describes a number of characteristics or dimensions that DOR and basic research have in common and a number of characteristics or dimensions in which DOR might be distinguished from basic research. Basic research and DOR are not mutually exclusive. In some projects, research can be classified along a continuum ranging from basic research to DOR. The focus in this chapter is on the endpoints of this continuum. However, the reader should keep in mind that actual research projects often include different research modes. In particular, basic research often involves design activities such as the design of experiments, the design of representational formalisms and the design of models. The differences between research with several design-oriented sub tasks and research with primarily basic research oriented sub tasks that we discuss in this chapter are related to the different goals and corresponding different outcomes.

3.2 Preliminary remarks on the term 'basic research'

Many scientists and institutions attach different meanings to the label ‘basic research’ [166-171]. The intention of this chapter is not to discuss all the subtle variations and differences in these characterizations. The primary intention of this chapter is to highlight challenges encountered by faculty with a basic research (BR) background who become deeply involved in DOR of digital learning resources and to clarify DOR for those faculty. For this purpose, we assume that the definition of the Frascati Manual [172] will suffice: "Basic research is experimental or theoretical work undertaken primarily to acquire new knowledge of the underlying foundation of phenomena and observable facts, without any particular application or use in view." In addition, the manual explains: "Basic research analyses properties, structures and relationships with a view to formulating and testing hypotheses, theories or laws."

Different disciplines have different norms and values with respect to methodology and evaluation criteria and define ‘scientific’ (often implicitly) in different ways. We do not know general principles that demarcate scientific from non-scientific knowledge (e.g. [169, 171]). As this paper aims toward further articulation of the term DOR, we describe now certain characteristics that DOR will have in common with BR.

3.3 What basic research and DOR have in common

If we want to apply the term research both to basic research as well as to DOR, we should acknowledge that both have essential characteristics in common.

3.3.1 Both in basic research and in DOR knowledge must be shared

First, the term ‘research’ implies that its outcome is shared knowledge. This implies for experiments and observations that they should be independent from the experimenter/observer. Both an experimental result from basic research as well a proof of feasibility of a design should in principle be reproducible. For theoretical results, it means that these results and their argumentation will be open to critique by experts in the field. In this respect a theory, or contribution to a theory, in mathematics, economics, physics, cognitive science, educational research, information systems development and other disciplines is not different from the design of a nuclear waste facility, a national system of education, or a public health system.

3.3.2 Both types of research ignore the inner workings of many systems

Both basic research as well as DOR could not exist in the way it exists now, without ignoring inner workings of certain systems. Simon [86] expresses this as follows: "... This skyhook-skyscraper construction of science from the roof down to the yet unconstructed foundations was possible because the behavior of the system at each level depended on only an approximate, simplified abstracted characterization of the system at the level next beneath." Thus, it was centuries ago already possible to make cheese, long before the biotechnological process was understood at a microbiological and molecular level. In other words, all research allows us to ignore inner workings of a system as long as all state variables are within their operational limits. Thus, we view the existing as well as an imagined world in terms of systems, which have interfaces defining assumptions with respect to their outer environment.

3.3.3 Both types of research aim at minimizing number and scope of assumptions

The set and scope of assumptions about a system's environment should be minimal. This means that any assumption about the system's environment that is not necessary for the system to function properly should not be part of its interface. In basic research this is the counterpart of the 'law of parsimony', which is often attributed to a 14th century Franciscan Friar William of Ockham and called Ockham's razor [173]. Essentially, it says: remove any assumption from your line of reasoning that is not necessary to arrive at your conclusion. In addition, it suggests that the larger the set and scope of assumptions in a particular line of reasoning, the less confidence we should have in that line of reasoning. In design, this principle is more directly based on the desire to minimize the dependence of one component on any other component. In design, this principle is the core of what is called 'modular' design.

For the design of learning materials or courses, 'minimizing dependency' implies that prerequisite knowledge should be minimal and that the learning material or the course should be composed of modules and that each module should assume only minimal prerequisite knowledge. It should be stressed that 'minimal' is not the same as 'little' or 'few' or 'small'. For course modules 'not more than necessary' forces us to demonstrate the necessity of any piece of prerequisite knowledge.

The principle of modularization is well known in every scientific field and presented in most textbooks on information systems development and instructional design.

3.3.4 Both types of research imply modeling

Apart from research in disciplines such as pure mathematics and logic, both types of research have in common that models are produced of systems that exist in reality. In addition, DOR is also strongly engaged in the design of models of systems, which do not yet exist in reality. These models will be discussed below when the differences between basic research and DOR are listed and clarified.

A model is itself a system. Sometimes we need to be aware of the distinction between this system that is the model and the other system that is the counterpart in reality of this model. For this reason we call the latter 'object system'.

3.3.5 For both types of models representational formalisms are essential

Modeling implies representation and therefore both types of research need representational formalisms. Many models can be expressed in the formalism or language of differential equations. Other models may be expressed in some computer simulation language. In cognitive science many models have been represented in the form of 'if – then' rules, called 'productions' [174]. In instructional design we sometimes use flow charts in order to describe a plan for a course or a lesson [23, 32, 175]. Some examples of models in information systems research are data models [176, 177] for database design, UML use-case, activity and sequence diagrams [117, 178] for process design. Each of these models requires a different representational formalism.

It may happen that we need to translate one model into another model. The process of demonstrating that the translation of one model, represented in a specific representational formalism A, into another model represented in another representational formalism B, is correct with respect to the purpose of the latter model is called *verification*. As with design, there are several definitions of verification, each dedicated to a specific discipline (see for instance [75, 154, 179]).

Within a number of disciplines, we can rely on a widely accepted rigid modeling formalism. The most rigid formalisms are based on a branch of mathematics such as set theory, logic or algebra. Such a formalism, allows long chains of inference between experiments, allows very little ambiguity and supports prevention of errors as well as crosschecks for errors. The more we can rely on a widely accepted rigid formalism the more contributions to the growth of knowledge can be realized with a data set. Also, a more rigid formalism allows us to rely more on expert reviews [180]. A rigid formalism will help to reduce subjectivity: i.e. the extent to which the outcome of an expert review depends on the individual expert. Furthermore, rigid representational formalisms tend to make it easier to enable computer support both for basic research as well as for design. Finally, rigid formalisms help to reduce uncertainties both in design as well as in basic research.

3.3.6 Both types of research still require the formalism of natural language

Both in design as well as in basic research, rigid representational formalisms are severely limited as to what can be represented. One formalism can only represent specific types of knowledge. Moreover, a specific formalism is usually dedicated to communication between or within specific populations. In basic research, a formalism is primarily a vehicle for communication within the discipline. In design-oriented research, we need formalisms that also support communication between different disciplines and between disciplines and stakeholders and users. For instance, a data modeling formalism is powerful for representing the structure of a database and can be used as a vehicle for communication with users and stakeholders as well. A blueprint of a house or a building can be read by users, stakeholders and constructors as well.

In practice, both basic research as well as design-oriented research, still require the ambiguous formalism of natural language as an additional vehicle for communication within the discipline or within the design field. In fact, in educational research, this ambiguous natural language formalism is by far the primary vehicle of communication.

3.3.7 Both involve bounded rationality and decision-making

Humans as well as communities face practical limits of rational choice due to limited data collection resources, limited information management and limited information processing resources. This is characterized as bounded rationality [181, 182]. Both the designer as well as the basic researcher make decisions and defend their choices. Decision-making involves handling uncertainty, lack of information and insufficient computational and inferential resources. Decision-making involves taking responsibility. Thus, a selection made by a machine is not a decision. This is particularly relevant when we evaluate a design with respect to ethics.

In basic research, the formulation of basic hypotheses and definitions involve decision-making. In design, the articulation of a design goal into design requirements requires decision-making. All design requirements as defined in the FBT projects [2, 5, 49, 52, 96, 99, 135, 137, 138, 141, 146] were results of decision-making processes.

3.3.8 Both take into account the cost/benefit ratio of research activities

All research takes into account in some way the ratio of costs and benefits of research activities. Costs can in principle be measured in the same way in both types of research, but benefits are usually implicitly defined. The implicit definition of benefits will be based on how much the outcomes of the activities bring us closer to the goal. Most researchers will not invest in research activities that will cost much research time unless they expect sufficient returns in relation to the goals. However, because both types of research aim at different types of goals, perceptions of the benefits of the same activity are likely to be different.

3.3.9 Both follow widely accepted methodological guidelines

Both types of research are directed by disciplinary requirements with respect to acceptable goals and mostly follow methodological guidelines that are widely accepted within the discipline. New types of goals and large departures from such guidelines raise fundamental discussions within the scientific community. Sometimes such departures and the subsequent discussions result in new guidelines and new requirements. Acceptance of results depends largely on the degree of compliance to generally accepted methodological guidelines.

3.4 Differences between basic research and DOR

3.4.1 Different drives, different questions

Typically, basic research is driven by the desire to create or discover structure, regularity or repeating patterns. We try to discover general principles and laws of nature or laws of human interaction. Alternatively we try to map parts of reality by listing and ordering the possible states of that reality. A classical example is the equation $E = mc^2$ [183] where E is the energy of a particle and m is its mass and c is the velocity of light in vacuo. In cognitive science an example is the equation $T = \alpha P^{-\beta}$, where T = a measure of performance, such as the recognition time or time to generate a geometry proof, P is a measure of the amount of practice and α and β are parameters. This relationship is recognized as a ubiquitous phenomenon in learning [41] and usually referred to as the 'power law of practice' or the 'power law of learning'. A classical example of a less compact mapping of parts of reality is the periodic system of elements. Other examples are geographical maps, genetic maps, taxonomies, and metabolic pathways.

Both compact mappings as well as less compact mappings represent relationships. Typical instances of basic research questions with respect to learning are: 'what is the relationship between performance and the amount of practice?' and 'what is the relationship between the type of learning task and the cognitive load that task imposes on the student?' [125].

Typically, DOR is driven by the expectation that it is possible to realize (instantiate) an artifact that will satisfy a group of stakeholders willing to provide a certain combination of resources.

The focus on the creation of artifacts implies a focus on how the world should be. An example of a DOR question is: 'is it possible to design learning material, which can be developed within certain budgetary constraints and which supports students in learning to design laboratory experiments in chemistry?' [100]. Another example is 'is it possible to realize an adaptive tutoring system for tutoring heterogeneous groups, which is appreciated both by a specific population of lecturers as well as a specific population of students?' [49]. Many more of these 'is it possible' questions are answered by proofs of feasibility as described in design-oriented publications.

3.4.2 Discipline-oriented versus stakeholder-oriented

Definition of goals in basic research takes into account the state of knowledge within the discipline. Explicit references to the state of knowledge outside the discipline will be scarce. The question about cognitive load in the previous section is not likely to come from university lecturers but rather from within the cognitive science discipline.

Definition of goals in DOR attributes much weight to demands of stakeholders and therefore to the state of knowledge outside any scientific discipline. The need for digital learning materials that support a specific learning goal in a specific context can come from many stakeholders. For instance, the university may want to reduce wet lab instruction in order to cut budgets, lecturers may want to attain learning goals that are not yet adequately supported by learning materials, students may feel motivated by interactivity, publishers may see an opportunity for a new market and educational researchers may see an opportunity to track student behavior. The definition of a goal that is related to digital learning material, will to a large extent be determined by the state of knowledge of university administrators, lecturers, students and publishers. Their state of knowledge with respect to digital learning materials is likely to be very different from the state of knowledge of the disciplines of educational research, cognitive science, information and knowledge management and eLearning.

3.4.3 Structuring within a discipline versus structuring across disciplines

The demands from a variety of stakeholders are seldom disciplinary demands. Mapping the structure of the demands of stakeholders from outside the discipline to the structure of knowledge within and across the required disciplines is seldom a trivial task. This task usually places considerable load on the resources, which are available for the design project while in pure basic research this load will be zero.

In DOR, structure must be mainly instrumental to our attempts to reach the goals. In basic research, structure is a primary characteristic of the goal.

3.4.4 Difference in selection of goals

In basic research, goals are selected because attaining these goals implies the building of new pillars, which will support new theoretical developments. In basic research, a goal is an intended new element of the theory or a new piece of knowledge that fits an existing mapping structure and the people who must be satisfied are primarily the members of the basic research community. Basic research focuses on theoretical progress. In basic research, selection of a problem tends to be based on the expected contribution of solutions of the problem to theoretical knowledge or to extending the knowledge of a system.

In DOR, goals are selected because it is expected that attaining the goal will satisfy the stakeholders. DOR focuses on practical problems or opportunities in real life that got high priority on our agenda because of a balance between urgency, available resources and the expectation that a satisfactory solution can be realized.

Thus, the knowledge domains of many problems formulated in DOR in education are much less restricted than the knowledge domains of the highly artificial problems referred to above. Most challenges in the design of courses and learning materials involve real educational settings. It is necessary to take these real educational settings into account during the design process in order to prevent implicit assumptions about the environment in which the design will be realized and implemented.

3.4.5 Differences in process management

An important practical consequence of the differences listed above is that in a process of DOR considerable communication efforts are needed in order to keep the perceptions of diverse stakeholders in line. These communication efforts should be part of dedicated 'gap management' over the time of the project. Here the term *gap management* refers to management of the dynamics of any gap between current perceptions of the current design goal that separate stakeholders have. The DOR process is a learning process both for designers as well as for stakeholders. However, in general, designers invest much more time in that learning process than other stakeholders do. In addition, the role of each stakeholder is different. Furthermore, the role of stakeholders may evolve over time. Finally, different persons represent stakeholders at different times. Overall, it requires effort to keep the learning processes of designers and stakeholders sufficiently aligned. In a pure basic research project, researchers can allocate almost all communication efforts to the normal publication process, which is required anyhow.

3.4.6 Including budgetary constraints as components of the goal

Budgetary constraints are often explicitly part of the design question. For instance a design question in the FBT program was: Is it possible to design digital learning materials, which satisfy a certain set of requirements, such that design and realization would cost less than k€ 1.5 for one hour of study effort. Satisfying this constraint would also imply an improvement in terms of cost-efficient realization in comparison to some other figures. For instance, Hasebrook et. al. [46] report a figure of almost k€ 30 for one hour of study effort. Burkhard [184] reports figures between 5 k$ and 30 k$ for redevelopment of one classroom hour of mathematics instruction.

In basic research, such cost constraints would seldom be expressed explicitly nor be part of the definition of the research goal. However, also in basic research, budgetary constraints do determine if a research group can invest successfully in answering a specific research question.

With respect to improving education or improving information and knowledge management in society it is clearly necessary to take into account the costs of adopting and implementing new designs at a large-scale. Thus, in DOR it is necessary to take into account cost constraints on the large-scale realization of the design goals. Not considering cost constraints in DOR severely limits its value.

3.4.7 Falsification of a hypothesis

Benefits of research activities are, often implicitly, measured in terms of an estimate of a distance to the goal. In a DOR project, stakeholders will not perceive empirical research aiming at falsification of a hypothesis as an activity that brings the design closer to the goal. In other words, the expected benefits of efforts to falsify a hypothesis are, in terms of reducing the distance to the design goals, almost never in balance with the costs.

The one exception is the falsification of a hypothesis stating that some design goal cannot be realized. Such falsification results in a proof of feasibility or contributes to a proof of concept and is usually valuable to the stakeholders. For stakeholders in a DOR project, the fact that a specific goal has been achieved at least once, contributes to evaluation of their requirements and implies that their support has not been in vain. A proof of feasibility is likely to be a requirement for their continued support for research aimed at a more general class of goals that are similar as the achieved goal (see also [83]).

3.4.8 Most DOR projects cannot allocate resources to basic research questions

In DOR, allocating resources to basic research questions is usually in conflict with the need to achieve the design goal. Allocating resources and investing efforts to basic research questions in the context of a design project may have the consequence that the design goal is never reached. For researchers who have a background in basic research it is often difficult to accept this. Pure DOR would not allocate resources to theory building and corresponding experimental testing of laws of nature or human behavior. Managing a DOR project involving researchers with a basic research background, in practice often involves restraining the project from side slipping into a basic research mode. Although DOR does in general result in a shift of the design goal(s), large changes in the general direction of the project are often not acceptable for the stakeholders. The context of DOR will generally not allow researchers to allocate project resources to the follow-up of a surprise, unless such a follow-up matches the general direction of the project.

In contrast to DOR, pure basic research would not allocate resources to the application of basic knowledge to problems that come primarily from outside the discipline. Finally, in pure basic research it is more likely that a surprise will deserve a follow up **within** the project.

3.4.9 Shared understanding within a discipline versus shared understanding within the community of stakeholders

In basic research, 'increased shared understanding **within the discipline**' is a primary goal. In DOR such 'increased shared understanding' is not the primary goal, although shared understanding **within the community of stakeholders** is likely to support confidence in the design. Basic research is more driven by a desire to understand, while DOR is more driven by a desire to create. It is likely that creations benefit from understanding, both shared understanding of the stakeholders as well as a deeper understanding of individuals in the project. However, deeper understanding of individuals will have to be supportive and it will not be the primary goal.

3.4.10 DOR involves more models based on different representational formalisms

DOR will use models and corresponding representational formalisms that are provided by basic research. Often, these models stem from different disciplines. However, contrary to basic research, most resources and most capacity will be allocated to synthesis activities. The intended output of DOR is not a set of models or mappings of parts of existing reality. The intended output is, a model of a system, which still has to be created. Because of the synthetic nature of design and because design challenges mostly require an approach that crosses disciplinary boundaries, such a model is more likely to be a configuration of several models based on diverse representational formalisms. For instance, the design of digital learning materials requires among other things a configuration of pedagogical (didactical) models, cognitive science models, information systems models and models in domains of the subject matter such as chemistry, molecular biology, epidemiology and bioprocess engineering.

3.4.11 DOR requires more efforts in connecting theories

A discipline generally focuses on defining and articulating a conceptual framework and in discovering relations between a limited set of variables or acquiring and structuring knowledge within the scope of the disciplinary field. For stakeholders in a design project who have different disciplinary backgrounds, concepts and variables are only instrumental to attaining a goal. Consequently, DOR will have to invest much more in making connections between theories than basic research.

3.4.12 DOR requires applying concepts across disciplines

In the design of learning materials, application of concepts from different disciplines is essential. In particular, the synthesis of subject matter knowledge such as knowledge of a subject within chemistry or within biology with pedagogical (didactical) knowledge into PCK is essential. To a certain extent, the design of the digital learning materials will imply capturing or elicitation of PCK from SME's. The design of learning material will very much rely on such PCK and at the same time contribute to PCK. Furthermore, the design of learning material must imply the application of medium related knowledge. Examples are knowledge about layout of texts for different purposes in textbooks or knowledge about human computer interaction for digital learning materials.

3.4.13 DOR requires connecting modules across disciplines

A bioreactor is composed of several modules or components. Some examples are a stirrer, an electromotor and the vessel. The stirrer and the electromotor are designed by different disciplines. For digital learning material diagrams, style sheets, definitions of concepts and variables et cetera require different disciplines. The design of style sheets is typically the task of the graphics design discipline, while definition of concepts and variables is typically the task of the subject matter discipline. Other modules or components such as learning objects involve intrinsically knowledge from a variety of knowledge domains. In Chapter 1, we have defined learning objects as modules of digital learning material aimed to support the achievement of an ‘atomic’ learning objective.

Many modules or components have a wealth of embedded knowledge and are products of a long history of design and realization in different disciplines. Learning to use modules that are made available by other disciplines requires a specific attitude. In particular the designer must learn to search for potentially useful components and to evaluate them. The better the worldwide availability of modules or components the more important it becomes to be able to find and connect them. For digital learning materials this issue will be discussed in 6.2.3.

3.4.14 DOR needs to invest more time in knowledge acquisition across disciplines

Many basic disciplines are still in a stage of development in which they cannot deliver much knowledge in a form that can easily be used by anyone outside the discipline. Ideally, a basic discipline would deliver results in the form of well-defined concepts, laws or conclusions illustrated with paradigm examples and fit to be used outside the discipline. Most outputs of cognitive science and research on learning and instruction in instructional design or the design of learning materials do not yet take such a form [185].

For example a lecturer in chemistry cannot directly apply a conclusion like "Elaborate processing facilitates explicit memories but not implicit memories" [41]. Also the use of guidelines on the basis of cognitive load theory [186] requires more than just applying guidelines [37]. In most cases, the lecturer will have to find out the meaning of each of the terms in such a conclusion or advice. At first sight, it might seem that this can be solved in a straightforward way by presenting the definition of each of the terms. However, importing a definition from another discipline may imply the import of more other terms and set in motion a recursive process of importing definitions and terms. In practice, it even occurs that several of the definitions used within the 'supplying' discipline do not convey exactly the same meaning or are inconsistent. Finally, it is fundamentally impossible to define every word in a natural language. Thus, many concepts in any discipline are not explicitly defined. For many concepts, the meaning is more or less conveyed in the way the concept is being used within the discipline. A person who wants to learn to understand such a concept will have to invest time in becoming involved in the use of the concept.

The bottom line for DOR is that goals cannot be achieved by one discipline and therefore force the designer(s) to allocate considerable resources to acquisition of knowledge of other disciplines. These resources cannot be allocated to the growth of knowledge within one discipline. Thus measuring progress in design cannot be equated with measuring growth of disciplinary knowledge. An attempt to compare a DOR project and a basic research project, even when the resources allocated to each project would be equal, in terms of disciplinary output does not make sense.

Anyone who decides that part of research in education should be design-oriented, should also stop comparing this part of the research with basic research in terms of disciplinary output.

3.4.15 Academic reward for investing in knowledge acquisition outside own discipline

In a team in which various members have a background in different knowledge domains, both the application of concepts as well as communication still requires a shared body of well-understood concepts. Even in the simple case of two team members each with their own discipline, at least one member needs to acquire knowledge from domains that is typically covered by the discipline of the other. Chances that this investment in acquiring knowledge of another discipline will be balanced by academic rewards are relatively low. For one thing, there are few journals that cross disciplinary boundaries and have a high impact factor. Secondly, relatively few scientists succeed in acquiring so much knowledge and expertise in a second discipline that they can publish in high impact factor journals in the second discipline as well. In practice more than two disciplines will be involved in a DOR project. This makes it less attractive to be involved in DOR for those members of a community who mostly publish in journals at the basic research end of the continuum.

In research published in journals such as the Journal of Computers in Mathematics and Science Teaching, Computer Applications in Engineering Education, European Journal of Engineering Education, Biochemistry and Molecular Biology Education, Journal of Chemical Education and so on, the SME is often the person who acquires some knowledge of cognitive science and research on learning and instruction. Actual application of outcomes of research in cognitive science and learning and instruction is very limited [185]. For instance, lists of references in publications in Biochemistry and Molecular Biology Education 2007 includes few references to cognitive science journals or journals on learning and instruction. Rather one will refer to an introductory textbook on instructional design, a monograph dedicated to university education like [9] or a report like [27].

At the same time authors of such textbooks (see for instance [14, 23, 30, 32, 175]) refer seldom to disciplinary references about subject matter in those journals that we have referred to as DSE journals. Thus we have not many illustrations of how to apply outcomes of research in cognitive science and learning and instruction in important subject matter areas. In other words, currently both 'sides' do invest little in providing such illustrations.

Thus, this state of affairs in higher education imposes not only a challenge on SME's, but also on cognitive scientists and researchers in learning and instruction and higher education. The challenge for the latter is to work on reduction of the necessary effort to acquire knowledge that they generate.

Alternatively, cognitive scientists, researchers in learning and instruction and instructional designers might acquire knowledge in domains such as chemistry, biology et cetera. The benefit of this would be that it would help them to apply the concepts of their discipline themselves and in this way to increase their understanding of these concepts and their limitations. This might be beneficial for the disciplinary research on cognition and learning and instruction. On the other hand, this requires extensive effort from cognitive scientists, researchers in learning and instruction or instructional designers. In the past this has given rise to so-called knowledge acquisition and knowledge engineering methods [187].

3.4.16 Selecting relevant systems and behaviors

In basic research many efforts are aimed at the construction of models of existing object systems. The function of these models is mainly to support communication, understanding and prediction of certain behaviors of the object system within the discipline. A typical class of models is the class of models of the living cell. Another typical class of models is the class of models of cognitive processes. Relevancy of selected behaviors, for instance the behaviors of cells or cognitive processes, is determined almost exclusively by existing knowledge within the discipline. If these models have been able to predict the outcomes of experiments such correct predictions are usually regarded as evidence supporting their validity.

In DOR many efforts are aimed at modeling the existing object system **and** at constructing models of new object systems, which **yet have to be** realized in the future. Relevancy is primarily determined by the stakeholders and is reflected in the design requirements.

The function of a model of the existing object system in DOR is to support the formulation of requirements for the design and assumptions about the context of the design. Recall that these assumptions are part of the interface of the design. Moreover, when we want to provide evidence for a difference between the existing object system and the new object system we might want to have a model of each.

The function of a model of a new system is primarily to describe a system, which can be realized and which will operate according to well-defined requirements. We will call such a model ‘the design proper’. DOR requires confidence in the design proper in advance. Normally we want to be able to judge the design proper of a nuclear waste facility or the design proper of a public health system or an insulin or aspartame production facility before it is built and set in operation. In the context of architectural design Alexander wrote [104]: “… we certainly need a way of evaluating the fit of a form, which does not rely on the experiment of actually trying the form out in the real world context.” In software engineering, this need is widely recognized and many tests have been developed that can be carried out in laboratory like conditions [154]. In particular there is a host of literature on usability testing based on samples of the end-user population. One suggested ideal for usability testing is based on combining data from thinking aloud sessions, with questionnaire data and unobtrusive observations of users [188].

3.4.17 In DOR risk assessment and hedging against risks is important

In general, stakeholders want experts to have confidence that a realization and implementation according to the design will bring reality closer to the goals [87]. Part of this is also the confidence that any risks are acceptably low. A general perception of shared understanding will support such confidence.

Risk assessment is important in DOR because DOR implies the realization of a new situation. One of the experiences in the FBT program was that the practice of DOR in regular education requires attention to be directed to hedging against risks. A number of the lecturers who were going to use the resulting learning materials expressed concerns about the risk that introducing the new learning materials might create any problems in terms of dissatisfaction of students or unsatisfactory exam results or both.

3.4.18 Allocating resources to measurements and observation

In basic research, the aim to know an object system directs us to hypotheses concerning relations on a few variables. Consequently we want to change one or very few variables at a time and measure changes in other variables. For instance, we may want to change the oxygen concentration in the outer environment of cells in a culture and measure changes in protein concentrations in the cells. Alternatively, we measure for a period of time the values of a set of state variables that change due to the internal dynamics of the system. For instance, we may want to measure the spatial distribution of protein concentrations in an organism such as Drosophila during its early development stage. We assume that, in contrast to design challenges generated by society, the laws of nature and of human behavior that we want to discover are more permanent.

In DOR we design, realize (instantiate) the artifact, implement it in an instance of the class of intended environments, plan and carry out measurements. In DOR we want to change reality before the moment arrives at which the design challenge does not exist any more. For example, within university courses one can expect that many learning objectives have a life cycle of less than a decade. The challenge to design and develop digital learning materials for these learning objectives will have to be satisfied in the beginning of this life cycle. The context for DOR is often determined by a specific organization or infrastructure. This context will not be permanent. This limits the time frame for DOR. In particular, the technical context of digital learning materials, such as desktop and laptop configurations, user interface technology and learning management system, will change rapidly, even to the extent that many of the solutions of today will be outdated in a decade. Thus, in the practice of DOR, it is usually necessary to change many variables between two measurements. Consequently, interpretation of measurement results with the intention to provide evidence for a basic law or to validate a model is seldom feasible. In DOR, the next set of measurements we want to take is not selected for its expected value to validate a model, but for its expected value in relation to an actual practical need.

Experiments in DOR may suggest to us useful and realistic assumptions or increase our confidence in certain assumptions. However, in DOR, the amount of resources and time that is allocated to an experiment must be in balance with the amount of resources and time that is allocated to other design tasks. In pure basic research all capacity is allocated to learning to know the object system and to understand its behavior.

If the primary goal of DOR differs from the primary goal of basic research it is not likely that evaluation of a contribution in DOR means the same as evaluation of a contribution in basic research.

3.4.19 Different forms of generality

Both types of research aim at generality. Generality in basic research mainly implies applicability in many possible problem states or many possible inference chains. Generality in DOR can imply that the same problem – solution pair occurs many times in the real world. Generality in DOR has often the meaning of 'valid in many contexts' or 'valid in contexts that occur often'. In DOR projects on digital learning materials and in particular learning objects this will imply design requirements that should guarantee such generality. Generality in DOR is much more related to return-on-investment-as-perceived-by-the-stakeholders.

In education, a typical design challenge is often: how to support the understanding of a specific concept. In these cases, we have to design a sequence of activities. Some of these activities should be based on situations in which students have to apply the concept and some activities should be based on situations in which students have to identify the most useful concepts in a range of potentially useful concepts. Ideally, such a sequence of activities forms a pattern, which can be reused worldwide in a range of contexts. Such patterns are instances of so-called 'design patterns'.

A design pattern is a specific configuration of components or activities and parameter settings, which fits a type of sub problem. Components, activities, design patterns and designs almost never can be represented as the compact structures mentioned above in the discussion of basic research output. Explicit representation of design patterns was proposed in [88] and taken up in software engineering some twenty years later [89]. Recently more and more publications about design patterns in education have been appearing [60, 106, 128-131, 189].

An example of a design pattern in digital learning material is the 'what do I need pattern'. This pattern consists of the presentation of a set of options and the request to select a subset of these options that is necessary for a specific purpose. Thus we might present a photograph of a laboratory and request the student to select what is necessary and sufficient for a specific measurement or experiment. Or we might present a set of unit operations that can be applied in a purification process and request the student to select which unit operations are applicable for a specific fermentation broth. Or we might present a set of basic definitions and equations and request a subset that is necessary and sufficient for calculating the value of a specific variable from a given data set. In all these examples the pattern focuses the student on what (s)he needs. In most projects on digital learning materials in which we were involved, SME's believed that this is a step that should precede actual calculations or actual detailed processing.

3.4.20 Different types of publications

Research should at least result in shareable knowledge. In its most used and most tangible form, this means publications in refereed journals. Both basic research and DOR deliver such publications.

At the basic research end of the continuum, publications contribute to the growth of disciplinary knowledge in terms of new experimental results, new fundamental concepts, reusable models including (contributions to) maps and taxonomies. Finally many publications also are reports of validation of these knowledge structures.

At the DOR end of the continuum, publications contribute to growth of knowledge that is related to a class of related design challenges. Such publications present goal definitions, design requirements, designs, design decisions, architectures, underlying arguments, components, design patterns and proofs of feasibility or proofs of concept and paradigm examples. Finally, interfaces, proposals for standards and specifications, prototypes and reference models should also be acknowledged as outputs of DOR

3.5 Beyond DOR

Depending on the design goals, the final stage of the design process may be: a large-scale implementation of the same instantiated artifact in many different contexts. Such a large-scale implementation can reveal implicit assumptions about the context that turn out not to hold in some of the contexts. For instance, designers of learning materials are likely to be unaware of their implicit assumptions with respect to prior knowledge of student populations.

Alternatively, large-scale implementation may reveal that certain explicit assumptions are not crucial for satisfying the design requirements. This is because a design will in practice never be consistent with all assumptions. If, in many contexts, the design is successful despite the inconsistencies with explicit assumptions about the context, these assumptions are probably not essential.

Research that is based on large-scale implementations in different contexts implies a transition from design-oriented research towards other forms of design-related research and corresponding methods.

Alternatively, large-scale implementation in the same context, might realize conditions necessary for carrying out studies aimed to provide empirical evidence that corroborates certain hypotheses. In general, these hypotheses will refer to relations between one or a few variables of the design and one or a few context variables. Large-scale implementation enables us to carry out many tests of versions of the design that differ with respect to one variable. When these variables are considered to play a role in a specific theory this would imply a transition towards basic research.

Investments in such large-scale research are costly. It is clear that the necessary resources will only become available, when cheaper case studies at least seem to promise useful results. It does not make sense to invest in a large-scale study on the effectiveness of certain learning materials in many contexts until small-scale tests have delivered promising starting points for such research. It also does not make sense to invest in a large-scale study on the relation between certain design variables and certain variables of the context until we have a theoretical model that we want to validate or at least indications that such relations are likely to exist.

3.6 Concluding

In this Chapter 3 we clarified the concept of design-oriented research by comparing two extremes on a dimension that runs from pure design-oriented to pure basic research.

Table 3-1 provides a comprehensive overview of this chapter. The most important differences are directly related to differences in the type of goals. The goals in DOR are not goals of a discipline. Rather a DOR project has a project goal. This goal is determined by the stakeholders. In practice, achieving the goal in almost any DOR project will require research that crosses boundaries of existing disciplines.

In basic research the structure of the disciplinary knowledge will have been a basis for the definition of the research goal or research question. From a disciplinary viewpoint the research question will be rather focused.

In a DOR project there will usually be many variables and many relevant knowledge domains. From a mono-disciplinary viewpoint the focus will seldom be clear. Moreover, the problems that must be approached in a DOR project will not exist forever. Thus, it is seldom possible to change one or a few variables at a time in order to find relationships between variables.

Ideally, a DOR project should be followed by a large-scale realization and implementation of the design. This might realize the conditions for other types of design-related research and for more basic research as well. Alternatively, it will be difficult to acquire the necessary resources for such other types of research without a DOR project as precursor. In particular, other types of design research will need a good starting point. It is the task of a DOR project to provide such a starting point.

The examples in this chapter were mostly in the domains of the Food- and Biotechnology program. However, we now believe that a design-oriented research paradigm as illustrated in this chapter is relevant for a range of other problems and opportunities that involve both design and a broad range of knowledge domains.

Table 3-1 basic research versus design-oriented research

Both in basic research and in DOR: ▪ aim is **shared** knowledge ▪ ignore the inner workings of many systems ▪ aim to minimize number and scope of assumptions ▪ modeling is an important activity ▪ representational formalisms are essential ▪ formalism of natural language plays important role ▪ rationality is bounded ▪ many decisions have to be made ▪ decisions take into account cost/benefit ratio of research activities ▪ follow widely accepted methodological implicit or explicit guidelines	
BASIC RESEARCH	DESIGN-ORIENTED RESEARCH
▪ primary drive: need to 'understand'	▪ primary drive: need to create
▪ discipline-oriented	▪ stakeholder-oriented
▪ structuring within a discipline	▪ structuring across disciplines
▪ aim: contributions to knowledge construction	▪ aim: solutions problems/opportunities in society
▪ most communication in scientific publications	▪ managing perception of stakeholders w.r.t. goal
▪ budgetary constraints implicit	▪ budgetary constraints part of goal
▪ many hypotheses can be interesting	▪ aimed at falsifying impossibility hypothesis
▪ most resources for basic research questions	▪ no resources for basic research questions
▪ shared understanding within a discipline	▪ shared understanding among stakeholders
▪ models: one/few representational formalisms	▪ models: range of representational formalisms
▪ efforts in unifying theories	▪ efforts in connecting theories
▪ applying concepts within discipline	▪ applying concepts across disciplines
▪ connecting modules within disciplines	▪ connecting modules across disciplines
▪ invest in knowledge acquisition within discipline	▪ invest knowledge acquisition across discipline
▪ academic reward structure clear	▪ academic reward structure absent
▪ modeling existing systems	▪ modeling systems that do not yet exist
▪ allocating resources to measuring/observation	▪ allocating resources to risk assessment/hedging
▪ generality: useful for much new knowledge	▪ generality: useful in many real life occurrences
▪ many high impact factor journals available	▪ few high impact factor journals available

Chapter 4 Output-classes for design-oriented research on digital learning materials in higher education

Abstract[6]

Design, realization, implementation, use and evaluation of digital learning in higher education requires a synthesis of knowledge from different knowledge domains. In addition to knowledge and information management (K&IM) and information and communication technology (ICT), other relevant knowledge domains are in the fields of learning and instruction, cognitive science and various subject matter in specific fields such as biology or chemistry. However, research in these fields is traditionally mainly analytical. In this chapter, it is argued that synthesis requires a design-oriented research paradigm. Such a design-oriented research paradigm can be clarified by means of a typology of its output. This chapter specifies classes of outputs that are potentially valuable as results of design oriented research aimed at (digital) learning materials in higher education. These outputs should be submitted for review within the corresponding scientific communities. The output-classes are illustrated with examples from the design and realization of digital learning materials in higher education.

[6] An expanded version of this chapter has been submitted to InSITE as:
Hartog Rob.J.M., Huub Scholten and Adrie J.M. Beulens
Output-classes for faculty-based design-oriented research on digital learning resources in higher education

4.1 Introduction

In previous chapters we have referred to a range of projects on design, realization, implementation, use and evaluation of digital learning materials for higher education. In these projects several methodological difficulties surfaced for which we could not find a satisfactory answer in literature on design-related research methods. In Chapter 2, we have defined *design-oriented research* (DOR) as research that is primarily aimed to produce an innovative design and applies design methodology. This chapter aims to provide a handle for DOR in higher education and to give direction to corresponding publication efforts. We do this by describing typical output-classes of DOR.

The next sections provide descriptions of the output-classes. Each of these classes is usually well-known in some engineering discipline. However, these classes are currently still difficult to recognize in the type of journals that we have listed in 1.1.2. Outputs in these classes are often not yet recognized as valid contributions to the growth of Pedagogical Content Knowledge (PCK). Some of the output-classes can be matched with result-types as defined in the recently published "Memorandum on design-oriented information systems research" [16]. Insofar an output-class already has been used in a WU project, that output-class will be illustrated with examples from that project. It should be noted that the DOR vision is the result of reflection on our WU projects, rather than input of those projects. This explains why the match between the outputs of WU projects and the classes described in this chapter is not complete for all output-classes.

4.2 Output-class 1: design goals and operational design requirements

Design goals are overall formulations of goals of design processes. Some examples of design goals in the FBT projects were, to realize digital learning material that

- 'promotes active acquisition of Food Chemistry knowledge' [141];

- 'supports students in acquiring a certain level of understanding of a specific concept' [138];
- 'supports students to achieve a specific new learning goal' [35, 51, 92, 99, 134, 139, 146, 190];
- 'supports students in learning a structured approach to specific classes of modeling problems' [95, 96]. This was a sub-category of so-called structured approaches to problem solving (see [14, 191]).

A very different element of the design goal might be a specific return on investment (ROI) of the curriculum, the course, or the material.

At the start of an innovative design process, it is usual to formulate an overall design goal. This is needed to get support of stakeholders and to give some rough indication of the direction in which the process must evolve. Some design goals are quite clear from the start. For instance the goal to 'put a man on the moon' is quite clear. In practice however, the formulation of design goals is often not so clear and needs to be articulated and adjusted later on. One reason is that innovative design processes are learning processes. In particular, the knowledge and terminology to define the intentions of the design process are seldom available at the start of the process. A second reason is that design goals are seldom adequate for determining if these goals have been achieved at the conclusion of the design process. The examples in the beginning of this section make this very clear. Jonassen [127] describes instructional design as a process of constraint exploration, where not only pedagogical constraints are explored but also constraints on available/preferable/accessible technologies, available funds and talents, political and organizational mores and rules, environmental variables, learner characteristics, learner goals and the physical context in which the instruction will be delivered. All these types of constraints are also relevant in design of digital learning materials.

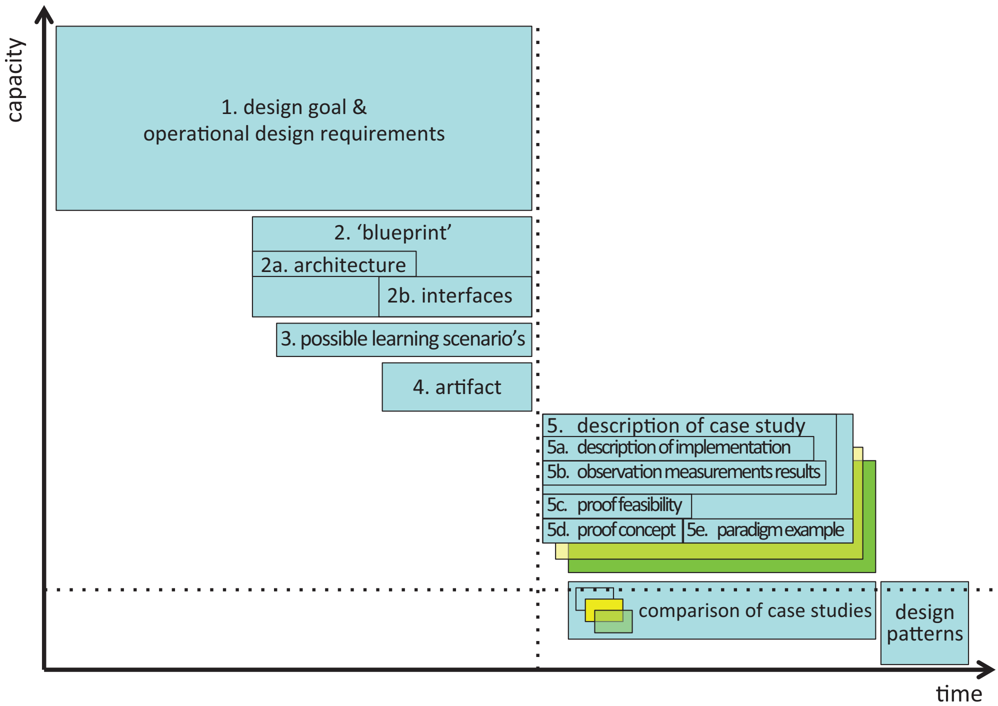


Figure 4-1 A qualitative extended Gantt Chart for DOR aiming at digital learning material

The rectangles denote tasks. The labels in the tasks refer to the output-classes, not the tasks. The structure of this figure is also the structure of this chapter. Each label should be read as 'Produce an instance of <output-class in label >'. The relative area of a task suggests the relative effort required to produce the indicated output. 'Blueprint' refers to the abstract representation of the design. In this figure 'artifact' refers to the realized design. With one realized design several case studies can be carried out. The realized design and one learning scenario must be ready before the first case study (see vertical dashed line). In the figure the blueprint is also ready before the first case study. However, the 'blueprint' can also be delivered much later, or not at all! Only one possible ordering of tasks in time is displayed. The area below the horizontal dashed line could be labeled as design-based research (DBR). The sets of DOR activities and activities that characterize other design-related research approaches are all fuzzy sets. In the area below the horizontal dashed line there will be more activities that fall outside the scope of this thesis.

In order to determine if the goal has been achieved we need *operational design requirements*. Operational design requirements define the actual goal of the design process. Operational design requirements make explicit how we can observe or measure if design goals have been reached. An operational design requirement is often based on a normative value of a variable in design space. A reference to such a value is called a *criterion*. In instructional design this general definition has been interpreted as " the desired level of achievement of a learner on a performance objective" [28]. A criterion can also refer to the desired achievement of an entire cohort of students or target population [28]. This latter type of criterion was used in most WU projects.

Most literature uses phrases like “**deriving** requirements from goals”. However, the mapping from design goals or learning goals to operational design requirements for learning materials **cannot be derived** by pure logic inference. Defining this mapping involves empirical research, decision-making and corresponding argumentation. In the field of information systems design, this is becoming more and more acknowledged. In fact, requirements engineering [154, 192, 193] has become a recognized field of research.

What is called 'requirements engineering' in other design disciplines ranges in traditional instructional design literature from 'needs assessment' to 'defining learning goals' to ‘writing down performance objectives’ [194] or defining 'intended learning outcomes' [9]. Alternatively, a target can be a *competency*. This is according to [19]:"

- a situated element of competence, which can be;
- behaviour-oriented and/or;
- task-oriented; and
- meaningful in a specific context and at a sufficient level of specification."

Here, c*ompetence* is [19]:"

- the integrated set of capabilities (or competencies);
- consisting of clusters of knowledge, skills, and attitudes;
- necessarily conditional for task performance and problem solving;
- and for being able to function effectively (according to certain expectations or
- standards); and
- in a certain profession, organization, job, role and situation."

Thinking in terms of competence and competency has become more relevant in higher education.
At the level of a course, focus on intended learning outcomes is currently policy at Wageningen University [195]. For reasons of readability we will mostly just refer to 'learning goal' or the 'learning goal-component' of the design and imply that this can be or can include target competencies, performance objectives or intended learning outcomes.
Assessment will refer to a set of constraints on performance indicators for whatever has to be learned. In terms of the constraint exploration view on 'design' a *performance objective* describes an action, activity or task that a learner should be able to do in order to demonstrate intended performance, as well as one or more corresponding contexts. The description will refer to qualitative and/or quantitative variables and constraints on those variables.
The operational definition of the learning goal-component should define what actions by the student would demonstrate that the student has achieved that goal and be sufficiently precise in order to come to agree with peers about whether a student has achieved that goal [196]. Thus, ideally, the operational definition is based on a measuring or observation instrument. Such an instrument can be used in assessment. Concept inventories [197-201] can be considered such instruments. These instruments are highly subject matter specific. They require primarily or exclusively input of faculty. They imply an operational definition of understanding certain concepts, but are not sufficient for assessing performance on many tasks that go beyond demonstrating understanding of a certain concept.
Apart from such instruments assessment in higher education is mostly formalized in exams. An exam can take many forms. Students might have to write down answers on open questions, do an oral exam, demonstrate a lab procedure, do an exam based on closed questions - possibly but not necessarily - computer-based, deliver a thesis and so on.
The assessment comprises those operational design requirements that are related to learning objectives, goals or target competency. However, for digital learning material

there will often be other operational design requirements as well. For instance, a range of requirements are related to motivation of students and to efficiency of studying. Those requirements are usually only partly related to the learning goal.

Once again it is important to stress that the design goal and requirements are the result of a constraint exploration process and that many constraints are not part of the definition of the learning goal. An example of an operational design requirement that is not at all related to the learning goal would be the requirement that those stakeholders who have no background and appreciation for the subject matter will give the look and feel of the material a rating of 8 or higher on a scale of 1 to 10. Another example of an operational design requirement that is not related to the learning goal is a restriction on the costs of distributing and maintaining the material. Finally there may be requirements that refer to the technical and organizational context in which the material will have to be managed. For instance a requirement might be that the interface of the material conforms to certain specifications. An example of such a specification is the SCORM 2004 specification [202, 203].

In some design challenges, the mapping from design goal to operational design requirements can be the result of a process that starts with the design goal and further articulates this goal. This is likely to lead to adjustments of the original formulation of the goal. In other design challenges, design goals are actually some form of abstract aggregation of a set of operational design requirements. Thus, in innovative design, the most adequate formulation of the design goal will often appear late in the design process.

The mapping between a goal and its operational design requirements as well as the argumentation for this mapping should be recorded and evaluated by other designers and stakeholders [83]. For the purpose of this chapter, we define *goal - assessment evaluation* as evaluation of the mapping between the learning goal and the operational design requirements (see also [204]). This comes close to traditional content validation as described in [205]. However, content validation usually assumes that the goal has already successfully been articulated in learning objectives described in natural language, or that the learning material already exists and de facto defines the learning objectives. In short, *goal - assessment evaluation* implies answering the question if we really assess what we want to assess. In higher education such evaluation is primarily the task of the SME.

Clearly, in innovative design and design-oriented research both the design requirements and the design goal are output rather than input for the design-oriented research process. At the same time, publication practice with respect to digital learning resources in higher education does not demonstrate widespread awareness of this. What we need is a publication culture that acknowledges the value of sets of operational design requirements. This implies that a set of exam questions or a description of a new type of assessment and its underpinnings are worthy of publication. This allows peer reviewers and readers to contribute to validation of assessments.

The same holds for design requirements (i.e. results of constraint exploration) that are not related to learning goals. These requirements will for instance be published in journals on distance learning, eLearning or 'computers in education'.

4.3 Output-class 2: the abstract representation of the design

4.3.1 Introduction

A second output of a design process should be a 'design'. Ideally that design is a model of what has to be realized. As such that model is one of the result-types listed in [16]. It is a set of specifications. The *instantiated artifact* is the realized design. We will use the term 'instantiated artifact' and 'realized design' interchangeably. Österle et.al. articulate instantiations as "concrete solutions implemented as prototypes or production systems". Note that the class of instantiated artifacts can also include realized processes. An example of a realized process is a lecture.

An abstract representation is described in an adequate representational formalism. For a house or a building the description is a blueprint using a representational formalism based on a set of symbols for components like doors, windows and walls, and relations between components. For airplanes, ships, information systems or music other sets of symbols are available. Of course, the instantiated artifact itself (e.g. the house or the ship or a realized learning scenario) is also a representation of the design.

4.3.2 Acknowledging the abstract representation of a design as output of DOR

Design-oriented research should ultimately result in one or more designs that can be submitted for review to the scientific community. For digital learning materials this will primarily be the community of SME's who are interested in the pedagogy of their teaching. A submitted design will be represented in some abstract form **or** by means of its realization. Both can be valuable products of a design process. In this section we argue that the abstract representation should be acknowledged as such in its own right. In fact, reviewing scientific products that are represented in an abstract formalism is in line with what is common in a range of scientific disciplines, in particular in mathematics but also in many engineering disciplines. It is much less common in DSE journals and in journals in educational research and the field of learning and instruction.

For digital learning materials, this implies that we need a publication culture in which it is normal that an abstractly represented design can and should be submitted for review. Digital learning materials are essentially information systems. In software engineering, the value of abstract models for communication, highlighting aspects, postponing expensive work and for documentation, is generally acknowledged [75, 177, 206, 207]. We plea for a publication culture in DSE and other relevant journals (see 1.1.2) in which it is acknowledged, in much the same way as in engineering sciences, that a blueprint without a corresponding instantiated artifact (e.g. a yacht or a building or a patent that is not yet realized), merits its own reviews. Such a review makes sense from the viewpoint of contributing to the growth of knowledge for that engineering discipline. For such reviews the actual realization (e.g. the building or the information system or the production process example) is not required.

An example of such a publication in bioprocess engineering is the design of the production process for D-(-)-4-hydroxyphenylglycin, which is an intermediate (side chain) in the synthesis of semi-synthetic penicillins and cephalosporins, especially amoxicillin, a broad spectrum antibiotic [208]. In this example, the primary representational formalism for the design is a flow-sheet formalism, which visualizes the flow of intermediate product in a network of unit operations.

Abstract representations of a design support communication between people with different roles and different disciplinary backgrounds. Abstract representations function also as documentation that can be studied in later times and other locations.

In particular, abstract representations support highlighting different subsystems of the design. For instance, in the blueprint of a building we can separately highlight the model of the plumbing or the model of the electrical system. In the abstract representation of most information systems we can separately highlight a model of the database, and a model of the user interface. Abstract representations also support the structuring of discussions and can make argumentation transparent.

Box 4-1 The need for paradigm examples in Pedagogical Content Knowledge (PCK)

> A theory requires paradigm examples in order to clarify the meaning of variables and parameters and laws [1]. Paradigm examples in DOR clarify the meaning of laws or guidelines, which are provided by theory. The example in this box shows that such laws by themselves are not always sufficiently clear to be used as guideline.
> A law to be taken into account in the design of learning scenario's and learning material is: "Being active while learning is better than being inactive: activity is a good in itself" [9]. Relating activity with learning is a good example of a starting point for instructional design and for learning that is widely acknowledged [9-13]. However, the statement does not provide practical handles to select or define adequate activities for a student who has to learn the essence of a specific scientific concept. More specific directions for selecting or defining activities are for instance given [9] within a framework of constructive alignment, in [10] with a focus on communication and collaboration and in [14] derived from whole task analysis. Even then, paradigm examples are needed to give meaning to the term 'activity' in the context of specific subject matter and specific learning goals and objectives. As such, most of the FBT learning materials clarify in considerable detail what activities are deemed essential by subject matter experts in the corresponding disciplines. Design-oriented research can and should evaluate and report about the applicability and usefulness of design guidelines that are delivered by basic research in learning and instruction and cognitive psychology. The application of such design guidelines is seldom a trivial operation. Therefore, paradigm examples of how to apply such design guidelines in educational practice are of great value [15].

In particular, an abstractly represented design can already be regarded as a paradigm example for the way in which theory should be applied and elaborated (see Box 4-1). **Note, that this does not require an empirical evaluation of the operation of the realized design.**

Finally, abstract representations of a design are a basis for analysis of the costs and benefits of implementing and operating the realized design, for assessing related risks and for planning the realization. In particular, in cases where realization, implementation and setting into operation of the design are expensive and/or imply high risks, review of the abstract representation is essential. Examples are the design of a nuclear waste facility, an information system for making flight reservations or a new law that defines a nation wide education policy. Even in cases where risks are less high, it often makes sense to review the abstract representation of the design before we invest in realization, implementation, setting into operation and empirical evaluation.

Thus we argue that, with respect to the design and realization of digital learning materials, it is important to initiate a shift toward representing, reviewing and publishing abstract representations of the design. This requires representing 'blueprints' of learning material that are suitable to be reviewed by peers of the SME acting as reviewer of the most relevant journal and, if they satisfy the standards set by the journal, published accordingly.

4.3.3 Output-class 2a: the architecture of a design

An important result of a process of designing a large system such as a building, an information system, a body of learning material or a national educational system, is a top-level structure and a set of related overall characteristics. In information systems design, such a top-level structure is called 'an architecture'. As such, the result-type or output-class of 'architectures' comprises top-level models that, in principle, can be instantiated. While a large system is ideally composed of many subsystems, the functionality of the system as a whole is in general more than the sum of the functionalities of the subsystems.

The architecture of a large system is a component of the design. In most design disciplines, the architecture is considered as an output-class by itself. Examples of architecture publications in the field of information systems for learning support can be found in [49, 209-214].

An architecture might support clear separation of different aspects of a system. For instance, the architecture of the Proteus system [49] consists of a database of closed questions, a model of the knowledge of the student in relation to the learning objectives, a subsystem that updates the model of the students' knowledge and a subsystem that selects a new question from the question database.

For learning materials, learning environments and more generally for highly interactive information systems we propose to distinguish an 'external' architecture, i.e. the main structure of the user interface, and an 'internal' or system architecture. This system architecture should both enable the intended overall user experience as well as efficient maintenance of the system (see e.g. [215]).

ICT allows us to separate architectural aspects that are directly related to the experience of the user from architectural aspects that are related to realizing the functionality of the system. For instance, a set of LO's might internally be stored in a relational database. For different target populations the system might present different 'external architectures'. For instance, based on the same 'technical architecture' the system might present on the screen different hierarchies or different preferred learning path's to members of different target populations.

The 'external' architecture of digital learning material will usually be related to one or a few overall types of learning goals. The 'external' architecture of most designs resulting from WU projects was the result of combining *cases* with a hierarchical view of a library of LO's and sometimes a glossary and an index to formulas and symbols.

A *case* is a combination of a situation, an assignment and a role for the student. Examples of such a role for the student were ‘assistant of the lab instructor’ and ‘consultant’. In such a role, the student had to complete an assignment. This generally required the student to make series of decisions and to answer questions. The student could anytime revert to a relevant *presentational Learning Object (LO)* in the library by following a link from within the case or by entering the library directly. Presentational LO’s are LO’s that *present* information in terms of texts, diagrams, screen-recordings and animations. The student can anytime revert to a definition or formula by calling a pop-up for instance by moving the mouse over a symbol. Most learning materials in the FBT program fit the following technical architecture description presented in Box 4-2.

We will refer to the combination of external and technical architecture presented here as 'the FBT architecture'. In general, an architecture is intended to match a vision on society, business processes, knowledge management, learning and so on and so forth. The external FBT architecture is intended primarily to support the vision that students should learn specific tasks such as the task to design a purification line [146] or a

Box 4-2 The 'internal FBT architecture'

- A database of reusable presentational LO's and interactions.
- One or more definitions of an ordering of these LO's. Different definitions enable the presentation of one or more different user views of relationships between LO's. For instance, different hierarchical orderings may be applied to present different *library structures.* One library structure might be intended to highlight the structure of the disciplinary knowledge. Another library structure might be intended to suggest a preferred learning path for novices.
- *Cases* that are constructed of interactions. Interactions allow the learner to *interact* with the LO or module. Through an interaction, the student selects or constructs a response. The interactions used in the FBT architecture are conceptually the same as the interactions defined in the QTI 2.0 specification [3]. Depending on the learners' response in a certain interaction a new instance of one of the available interactions is presented to the learner.
- An *author-defined data storage* that enables storage and processing of student's responses [4-6].
- A *database of definitions and formula's* that supports on-demand presentation of information for instance in the form of pop-ups.
- A simple *design and/or experiment environment* that enables the student to design models and to run virtual experiments.

specific type of experiment [2, 99, 100] or the analysis of data from human intervention trials or observational studies in human nutrition research [37]. The external architecture supports the vision that students can learn such tasks by carrying out analogous tasks in a virtual environment. In addition, it was supposed that learning the actual task can be supported by providing *seamless access* to necessary information and automated feedback on the students' actions. We define seamless access as access that does not by itself impose any cognitive load. At the top-level the underlying vision of the external FBT architecture is in agreement with visions on learning and instruction presented in [12, 14, 216-218].

The design of a top-level structure or architecture can be a consequence of decisions with respect to clustering of components into larger subsystems or be founded on a theory about learning and instruction. In both cases, we regard this as a case of strategic decision-making (see Chapter 2). Exposing such a top-level structure and the underlying decisions and their arguments in publications is important. As we indicated at the start of this subsection, architectures are already a widely accepted output in various disciplines in particular in relation to digital learning support.

4.3.4 Output-class 2b: interface definitions

For systems that can be replaced by other systems without the need to impose new requirements on the outer environment the term *module* is very common. Usually, designers aim at an output of the design process that is modular. A modular system is a synthesis of modules. The basic concept that is used to realize modularity is the concept of *interface*. The *interface* of a system describes a set of assumptions about the environment of the system and a definition of the function(s) of the system (see section 2.8). This often refers to the system's inputs and outputs.

In instructional design a modular approach has been described in [31, 149]. In this chapter, we use the term 'module' in its widest sense. A module can be a course or a part of a course and may include or not include instructors. A specific class of modules is the class of modular learning materials. A textbook can be regarded as a module. A video can also be regarded as a module. A special class of learning materials is the class of digital learning materials. Modules of digital learning material have been called ‘units of learning material’, ‘learning objects’ or ‘reusable content objects’ [85, 108, 219-222]. We use the term Learning Object (LO) for a module of digital learning material with one or only a few functions and few assumptions with a small scope. For instance a LO may have as its sole function to present to the student what the Polymerase Chain Reaction (PCR) is. Another LO may have as its sole function to present to the student for what purpose the PCR can be used. Yet another LO may have as its sole function to let a student experience something that can easily go wrong in using PCR. A LO is essentially a self contained information system.

Note that modularity is typically a property of the design proper. Thus, evaluation of conformance to requirements for modularity primarily involves inspection of the design proper by reviewers.

Modules and Learning Objects are defined by their interfaces. In innovative design, each interface definition of a module is an output of design-oriented research. We propose that interface definitions should be acknowledged as such and that interface definitions should be submitted for review and publication. In information systems design, a coherent set of interface definitions is already considered sufficiently valuable output of a design process. Such a set of coherent interface definitions should be reviewed, i.e. requires inspection by peer reviewers. It must be noted however, that the peer review process of interface definitions in engineering sciences, for instance in the field of information systems, usually takes place in working groups within professional organizations [223-225]. Academics are involved and reviews and results are public but usually in reports or specifications and not in journals. In line with this, a coherent set of interface definitions of a module and the learning objects within the module should be regarded as potentially valuable output of a design process. Such an interface should for instance include a description of the prerequisite knowledge. To a large extent a specification like the IMS learning resource meta data specification [119] already indicates what should be part of the interface description of a learning resource.

Note that an interface can best be judged based on its description and that realization of the corresponding design bears little relevance for the review of the interface.

4.4 Output-class 3: learning scenarios

A *learning scenario* defines roles for all actors and for learning materials in a teaching-learning process or just a learning process. Furthermore, the scenario may define which actions are planned and in what order. A well-known learning scenario is ‘problem-based learning (PBL)’ [226-228]. Point of departure in a PBL process is a short description of an intriguing problem. Part of the activities in a PBL process are discussions in groups of about 10 students. Students analyze the problems, formulate learning goals, consult external sources of information (partly in self study setting) and reconvene to report their results. A tutor coaches the group discussions. Another well-known learning scenario is the lecture. In a real lecture, the lecturer lectures and the students listen to the lecture and try to process what the lecturer presents and to make notes. Variations on this type of scenario include the lecturer asking questions to the audience and/or the students asking questions to the lecturer. The lecture can be based on or supported by learning material that is purely presentational. This is the most common lecture scenario. Alternatively, a variation on the lecture scenario can be based on a virtual experiment environment. In such a scenario the lecturer can ask questions or stimulate the students to ask questions with respect to configuring an experiment set up, setting parameters and expected experimental outcomes. In such a scenario, the lecturer and the whole group work together towards a satisfactory design of the experiment. The lecturer leads the discussion and carries out the necessary actions within the design environment, virtual experiment environment or modeling toolset, and shows intermediate results by means of a beamer on a large screen in a lecture room. In another learning scenario, the same tool might be used by students who each work individually with a computer. Again the goal would be that the student learns to design and carry out experiments given a certain need for information.

One learning scenario can often be realized with several different learning materials and one module of learning material might be used in different learning scenarios. Again, we can compare this with a textbook. Textbooks are used in a wide range of different learning scenarios.

Systems that try to control the learning process of the student at a very detailed level would also allow only one learning scenario: the scenario in which the system is an individual tutor of the individual student. On the other hand, in higher education, we want the student to learn to control his/her own learning process. In most WU projects, there was a certain tension between this latter idea and a desire to provide guidance. Resolving such a tension is not possible without taking into account many other constraints such as budgetary and timing constraints and constraints that express a desired balance between the amount of student time that should be spent on presentational material and the amount of student time that should be spent on active learning.

Moreover, in the FBT program, one of the design requirements in most of the projects was that the overall learning scenario should not be embedded in the learning material. In other words, in most of the projects we did not want that the learning material restricts the teachers to only one learning scenario. The reason for aiming at learning material that is useful in a range of learning scenarios was that we did not want to restrict the use of the material to those learning scenarios that are used at Wageningen University. However, some learning materials do constrain the set of possible learning scenarios more than others. In particular, the primary benefit of the Proteus system is based on its behavior that is adapted to the individual student. Using the Proteus system in group settings or collaborative settings implies forsaking its primary benefit.

Like any scenario, a learning scenario is the result of a design process. In particular the class of innovative scenarios that are articulated to a level at which they can be directly applied in education such as for instance the PBL scenario, is a class of potentially valuable outputs of design-oriented research in higher education. In the FBT program, no innovative scenarios were designed.

4.5 Output-class 4: realized designs (instantiated artifacts)

An obvious class of candidate outputs of a design process is the class of realized designs. Often a realized design is the realization of an abstract representation of a design. For instance the actual construction of a building based on its blueprint. Another instance of a realized design is a learning object based on a storyboard and a set of interface definitions.

Alternatively, a realized design can be a direct realization. This means that there is no abstract representation of the design proper. There is a house but no blueprint. Currently, most digital learning materials, including those resulting from the FBT program, are direct realizations.

A realization of a design is seldom fit for publication in a journal. A typical exception would be the design of a law. However, a real building does not fit in a journal. Neither does a piece of digital learning material. Nevertheless, the realized design of learning materials should be considered as a candidate output of design-oriented research and a potential contribution to the growth of PCK. It must be possible to submit these outputs to the scientific community for review. Currently a design-oriented publication in a scientific journal usually only contains a pointer to a realized design instead of the realized design itself. Examples of such pointers are descriptors of the physical location of a building or a web address of a learning object. In practice, the characteristics of the realized design often seem to bear little weight in the review process of the corresponding article **about** the realized design.

One problem is that few Subject Matter Experts (SME's) make themselves available for review of realized designs. There have been some attempts to set up review processes for LO's in several projects [229-231]. However, there is not yet an established practice of publishing realized designs of digital learning materials that includes a scientific refereeing process. Furthermore, publications **about** digital learning materials are not yet linked to reviewed and published digital learning materials. We propose that reviews of publications about realized designs are accompanied by review and publication of the realized designs themselves. This should hold in particular for digital learning materials, because for these products access can be realized at low costs. The possibility to access and use innovative digital learning materials by any lecturer at any university would also enable attempts to reproduce empirical evaluation results.

A realized design might be presented as a ‘prototype’. If so, the meaning of the term ‘prototype’ should be intended to indicate that the realized design is a good starting point for further research. The Memorandum on design-oriented information systems research [16] regards "concrete solutions implemented as prototypes or production systems" as one of the result-types of design-oriented information systems research. Note that this is different from the 'mock up' type of intermediate prototypes that are intended to support a learning process of users and other stakeholders. The latter meaning of 'prototype' is typically derived from *rapid prototyping* approaches.

4.6 Output-class 5: case study results

4.6.1 Output-class 5a: a record of an implementation

Fitting the realized design into its outer environment is called ‘implementation’. The design of a system is based on assumptions with respect to the outer environment. These assumptions are at the same time requirements to the outer environment of the instantiated artifact. It often occurs that these assumptions turn out not to hold any more, once we try to fit the realized design in its environment. In other words, that the requirements as to the outer environment are not met. In such a case, additional changes in the outer environment may be needed. Such changes may require considerable efforts. In order to reduce the risk of expensive implementations we usually design a product in which a number of parameters can be set at implementation time. In such a case, implementation involves setting the operating parameters of the realized design. For an electrical appliance this might be: setting a switch from 230 volts to 110 volts. An example of an operating parameter for a learning object might be a parameter that determines the language. Thus, implementing the learning object might involve setting its language from English to German or Arabic. Implementing a learning object using a learning scenario involves among other things realizing access to this learning object. If the learning scenario includes a role for an instructor, then implementation also involves preparing the instructor for this role.

A record of the implementation process of an innovative instantiated artifact can be a valuable output of design-oriented research in education. In particular, a list of those assumptions in the interface that did not hold and a list of the succeeding adjustments that were made, are likely to contain important information. An example of an implicit assumption that turned out to be wrong in the FBT Program is the following.

One of the design requirements for digital learning material in the FBT Program was that the material can be implemented and used in a wide range of different learning scenarios in different universities. Many digital FBT learning materials were indeed implemented in different universities, some in slightly different learning scenarios. These implementations at least revealed one set of implicit assumptions that repeatedly turned out to be incorrect. Implicitly it had been assumed that material after one successful trial at a test site would automatically be used more times at that test site. In practice, continuation of use after initial successful use, turned out to be an unexpected problem. Moreover, it has not been possible to attribute discontinuations to one common factor. Sometimes, the next year a new staff member was assigned to the course in which the material was used. Sometimes, local changes in the organization of the ICT infrastructure led to a failure at an inconvenient moment. Sometimes a lecturer asked too late for allocation of ICT facilities and so on and so forth.

Usually, publications of case studies do not clearly distinguish the implementation process from the other parts of a case study. We propose that a record of an implementation process is acknowledged output of design-oriented research. An article that presents a case study of the use of digital learning material should contain at least a separate section dedicated to a description of the implementation process. This section should contain a list of the assumptions that did not hold and a description of the adjustments that were needed. Furthermore, such a section should list implicit assumptions that were revealed at implementation time.

4.6.2 Case study result 5b: a record of a realized learning scenario

In the FBT Program, tools, design environments, modeling environments, virtual experiment environments and cases have been designed and developed such that they could be used in different learning scenarios. Except for the Proteus system, all these tools, environments and cases can be used in a regular classroom in a group discussion scenario as described above. A record of the execution of a learning scenario is fit for publication. The latter will usually be described as part of a case study description. Both the plan for the scenario and the underlying arguments as well as the record of its execution should be standard sections in publications of case studies about the use of digital learning materials.

4.6.3 Case study result 5c: proof of feasibility

When a design is realized and implemented in a certain context (outer environment) and has been put in operation, the operation can be studied in a case study. A case study can answer the question if it is possible to satisfy all constraints (including all requirements) in practice. If we have demonstrated in one case study that the realized design in operation has satisfied all constraints, this means that we have been able to achieve the design goal at least once. Note that the goal is almost never just one point in design space, but usually a large set of points in design space. This set of points is defined by all constraints, including design requirements, laws of nature and human behavior as well as the budgetary limits. The statement ‘the goal has been reached’ thus means that one of the many points that define the goal in design space has been reached. In other words, the case study has provided a proof of feasibility. If we also adequately documented the design process, then arrival at the goal also implies that we found at least one 'path' to the goal in one relevant design space.

A proof of feasibility is essentially the falsification of the hypothesis “that it is not possible to design, develop, implement and operate a design that satisfies all constraints”. Well-known proofs of feasibility are the first aircraft flights with human beings and the first arrival on the moon. Other examples in various disciplines can easily be found (see e.g. [232-234]). In line with this, we consider many contributions to PCK as proofs of feasibility. The primary intended output of a case study should be a proof of feasibility.

4.6.4 Case study result 5d: proof of concept

We define a *proof of concept* as consisting of (1) one or a few explicitly described concepts, (2) a proof of feasibility, (3) an explanation that links the proof of feasibility to the concept(s) in a satisfactory way and (4) the absence of alternative plausible explanations. In particular, (3) and (4) are subject to expert judgment. Literature in a variety of different knowledge domains shows many publications that use the term 'proof of concept'. A good example that reflects our definition is [235]. However, for the sake of our argument we will use some examples that require less disciplinary knowledge.

The flights of the Wright brothers [236] were proofs of the concept of the combination of artificial wings, propulsion by a propeller and power delivered by an internal combustion engine. Note, that, if the goal is simply 'to fly hundred yards' a completely different concept such as the ‘Zeppelin concept’ can also be proved to ‘work’. This does not reduce the value of the proof of concept delivered by the Wright brothers. A proof of concept does not require that there is only one design with which to reach the goal.

In the case of a zeppelin, the possibility to float in the air is attributed to the concept of a shape that is lighter than the air it replaces. In other words, the zeppelin is an application of Archimedes' law. When we want to discuss the zeppelin concept as a concept for air traffic, other variables and constraints besides variables and laws of nature, become relevant. For an air traffic concept, the Zeppelin concept must satisfactorily match with available infrastructure, operational costs, customer satisfaction, safety and dependence on weather conditions. Thus, a proof of concept should provide evidence, based on which we can attribute the achievement of a design goal to one or a few basic concepts. The design goal to be able to fly is different from the design goal to set up a sustainable air traffic system. Evidence that a concept contributed to the achievement of one goal is not per se evidence that this concept contributes to the achievement of a related goal.

In the design of digital learning materials, it is also sometimes possible to describe one or a few basic concepts that are the core of the learning material. For instance, van der Schaaf et. al. [146] describe how students achieve learning objectives by designing downstream processes (purification processes) using an educational environment for learning to design linear purification processes. The concept of this can be described as in Box 4-3.

Box 4-3 Support for learning to design a linear purification process

- A student can design a complete purification line by successive selection of simplified unit operations and setting control variables for each of the unit operations.
- The design environment computes after any selection and setting immediately the values of all relevant output variables (e.g. cost, yield, …).
- The design environment and the cases that are presented to the student are stripped from all details that are not relevant for the learning goal.
- At any point in the design process, the student can press a context sensitive help button.
- The cases presented to the student are such that a bioprocess engineering student can be expected to recognize the design problem as relevant in bioprocess engineering.
- Students can inspect each other's designs.
- The overall learning goals are:
  - given a fermentation broth and a set of requirements, the student is able to design a purification process that satisfies the requirements;
  - the student is able to make effective and efficient use of the characteristics of the unit operations.

In practice, a design can imply too many concepts for the term 'proof of concept' to be informative. Moreover, distinguishing a core concept from a concept that just contributes to achieving the design goal is not always straightforward. A core concept should be a concept that cannot be replaced by another concept without changing many other characteristics of the design. The example given above already highlights the problem of identifying the core concept(s) on which the learning material is based. For instance, we can ask ourselves: is it a core concept that the student can at any point in the process press a context sensitive help button?

With some exceptions (see e.g. [237]), publications on digital learning materials in higher education seldom use the term 'proof of concept' and the term 'proof of feasibility'. Nevertheless, more implicitly, publications sometimes provide a suggestion of a 'proof' of 'a concept'. We propose to be more explicit about claims and to distinguish between a proof of feasibility and a proof of concept. The latter should be understood to imply that the feasibility is attributed to the application of a specific concept or small set of concepts. The arguments underpinning this attribution will be primarily theoretical.

4.6.5 Case study result 5e: paradigm example

For realized digital learning materials, we define a *paradigm example* as an example of a realized design for which a proof of concept has been provided. In general, paradigm examples are examples that are worth to be followed, because they clarify theory and because of their success in one or more case studies. Paradigm examples of this type demonstrate a way to elaborate concepts from a general theoretical framework for a specific content. Paradigm examples of this type also imply that the design can be realized and that using the design actually can result in achieving the design goals. Designing and delivering such paradigm examples implies contributing to pedagogical content knowledge.

4.6.6 Concluding remarks about case study results

A case study should not be regarded as a sample from a population in the statistical sense [238]. A proof of feasibility or a proof of concept is valuable in itself and valuable as a starting point for incremental improvement. Aiming at a proof of feasibility or proof of concept is not the same as aiming for general applicability of a design. The more accessible the details of a design and a proof of feasibility or proof of concept, the better it will be possible to reuse the knowledge that is implied by this design and the case study. Implementing the realized design in a new context (outer environment) is likely to require new parameter settings and changes in the new context as well. The less the design is dependent on assumptions about the context the more likely the design will also ‘work’ in this new outer environment.

Follow-up research aimed to exclude alternative explanations or to provide more empirical evidence for proof of concept claims or aimed to investigate if the design works in a range of related contexts does not fit a DOR research mode but might be classified as a task that fits in one of the other design-related research methodologies.

4.7 Design patterns: a borderline output-class

We provided a definition of the term 'design patterns' in section 2.4.9. A concept of 'design patterns' was introduced by Alexander and Ishikawa [88] in architecture. Later, design patterns were adopted in other fields such as software engineering [89] and more recently in the field of learning and instruction [106, 128-131, 189, 239, 240].

For different subject matter disciplines, it makes sense to define classes of design patterns that are an abstraction of design patterns for the individual subject matter disciplines. An example of a class of design patterns that can be recognized in FBT learning material is the combination of a problem and a set of interactions [2, 99] in Box 4-4.

A design pattern itself is often related to specific subject matter. The example of the pattern that is presented here has recurred several times in the learning material in the FBT Program, but we cannot yet claim that it is generic enough to be applied in every science course.

Design patterns often bridge the gap between design guidelines and paradigm examples. They are not as abstract as guidelines but not as concrete as examples. They are not as generic as guidelines but more generic than examples [66, 240].

Insofar design patterns are made explicit in order to structure documentation of the design by highlighting recurrent patterns and presenting them in the same format, they can be regarded as outputs of design-oriented research. Moreover, design patterns should certainly be 'units of analysis' in a case study as defined in [241].

However, publications of design patterns mostly suggest or even claim that these patterns are reusable in new designs. Design patterns presented in this way belong, along with design guidelines, to developmental research as defined in [242] or to design-based research as defined in [18].

Box 4-4 A design pattern for supporting students in acquiring design skills

Given the challenge to teach the student to design an experimental procedure:

- Present a set of types of research objects. For instance, if the research question is to find the effects of free fatty acids on gene expression in the liver and which transcription factors are involved [2], then the types of research objects might be ‘cells’, ‘mice’ or ‘human’.
- For each of these types of research objects make a further selection what to study from a set of subtypes or foci or parts. For instance, when the student decides to study humans (given the same research question) he might be presented the set 'liver cells', 'red blood cells', 'blood plasma', or 'urine' to select from.
- For each of these, present a set of treatments or processing steps and let the student select a treatment or processing step. Depending on the actual object which we want to prepare for measurements and depending on the research question, different treatments or processing steps make sense. For instance, heating, drying, mixing, crossing, fasting, giving a special diet, et cetera.
- Present possible variables to measure and let the student select a variable.
- Present a set of possible measurement techniques for this variable and let the student select a measurement technique.
- Classify combinations of possible selections. Each class would imply specific feedback. This feedback would be based on what the combinations in the class have in common. For instance, it makes sense to define for every underlying concept a class for all combinations of selections that are related to understanding of that specific concept.
- Provide feedback on each of the resulting classes of combinations.

4.8 Concluding Remarks

This chapter lists, defines and articulates classes of potentially valuable or ‘candidate’ outputs of DOR in higher/science education. The classes are linked to core concepts in design literature. Examples of such outcomes in higher education have been given as well as arguments why outcomes of these types should be considered valuable. The list is intended to help SME's in their role as researcher and teacher and other researchers who intend to contribute to PCK in a specific subject matter field to focus DOR activities and to organize their publications.

Many design-oriented publications that have been accepted by journals describe output that fits some of the listed output-classes. However, it has been argued that several other output-classes do not receive enough attention in publications in the categories of journals that were described in 1.1.2. For instance, interfaces or operational design requirements are currently seldom regarded as candidate outputs. Furthermore, ‘blueprints’ of digital learning materials are seldom submitted for review and published in scientific journals. To some extent, this can be explained by the lack of an adequate representational formalism for digital learning material. At the same time, realized digital learning materials are in practice insufficiently subject to peer review.

Finally, an output-class not discussed in this chapter as a separate output-class is the class of evaluation results. Evaluating DOR involves evaluation of both outputs of DOR as well as evaluation of DOR processes. With respect to DOR in higher education, this chapter specifies **what categories of outputs** should be evaluated. The question **how** products in these categories should be evaluated and the issue of evaluation of the design process itself will be discussed in Chapter 7

Chapter 5 Defining the goal and the focus of innovation

Abstract[7]

Reflection on the definition of goals in the FBT program and successive WU projects reveals that choosing the focus of innovation in design-oriented research for digital learning materials in higher education is a strategic decision. In this chapter, we discuss the two major dimensions along which to distinguish DOR projects aiming to deliver digital learning materials for higher education. Along the first dimension we distinguish DOR projects in which the focus of innovation is a new learning goal and DOR projects in which the learning goals are already well-established and operationally defined. The second dimension is based on the type of learning goal. For this we make a threefold distinction. One focus in higher education is often on understanding of concepts, laws and methods. A second focus is on acquiring and understanding the main structure of an approach to certain academic tasks within the discipline. A third focus may be on 'whole task competency'. Whole task competency implies that the student is competent in carrying out a real-life task as defined by one of the professions for which the study program prepares the student. In this chapter we explain the relationship between the definition of the goal, the focus of innovation and various knowledge domains. Early awareness of this relationship is important in

[7] A version of this chapter has been published as
Hartog, R., A. Beulens, and J. Tramper, Faculty-Based Design-Oriented Research on Digital Learning Materials: Defining Project Goals in Proceedings of Global Learn Asia Pacific 2010, Z.W. Abas, I. Jung, and J. Luca, Editors. 2010, AACE: Penang, Malaysia. p. 1418-1427.

setting up a DOR project on design, realization, implementation, use and evaluation of innovative digital learning materials.

5.1 Introduction

In this chapter, we discuss major distinctions that are important for DOR projects aiming at digital learning materials in higher education. The first distinction is a twofold distinction and based on the **primary focus of innovation** of the project. The second distinction is a threefold distinction and based on the **type of learning goal or type of target competency**. At the end of this chapter, Table 5-1 summarizes the characteristics of the resulting typology, and Table 5-2 summarizes the most relevant knowledge domains.

5.2 The first major distinction: the primary focus of innovation

In the FBT program the focus of innovation was often a 'new' learning goal such as the ability to apply a specific approach to modeling or a specific design competency. However, in a number of projects we primarily aimed at another focus of innovation. Examples are level of motivation or support for learning in heterogeneous target populations. In this sub section we will discuss the consequences of the choice between aiming at a new learning goal (see left hand columns in Table 5-1 and Table 5-2) or aiming at other innovations (see right hand columns in Table 5-1 and Table 5-2).

5.2.1 Aiming at new learning goals

One of the challenges in university education will be to define new learning goals. These new learning goals might be generated by recent research findings (see for instance [35]), the genesis of a new discipline (such as nutrigenomics see [2]) or increased awareness of 'holes' in educational programs at the university. For instance, the educational programs that were supported by the FBT program provided very little opportunity to acquire design competencies. A range of FBT projects aimed to 'repair' this situation. Five FBT projects [2, 92, 99, 100, 139] focused on the design of experiments. Two FBT projects [37, 137] referred directly to the design of an epidemiological study and the design of a questionnaire respectively. In [146] the focus was on the design of an industrial purification line. In [35, 51, 95, 96, 134, 190] the focus was on the design of qualitative and quantitative models.

Apart from the type of learning goals in these examples, there is often a range of learning goals that have been mentioned by the staff for many years as desired, but were actually never part of a course and an exam, and consequently not achieved by students. We regard these also as 'new', precisely because these goals have never been articulated and operationally defined. For instance, for students who have completed a bioprocess engineering BSc program one would want that they are able to demonstrate adequate knowledge and understanding of most well-known issues of scaling up a bioreactor design.

A major challenge in DOR projects in higher education aiming at 'new' learning goals is the definition of assessments that operationally define these goals [66, 244]. Even when there is an initial formulation of the 'new' learning goals such as the formulation on scaling up a bioreactor design such a formulation is only a starting point for further articulation, definition of corresponding assessments and design of corresponding learning materials.

5.2.2 Aiming at established learning goals: possible directions for innovation

A project aiming at new learning goals implies the necessity to articulate these new learning goals and to invest considerable efforts in formulating operational definitions of the requirements that define these learning goals.

A project aiming at established learning goals does not require efforts for the task of defining and articulating learning goals and designing assessments. This is an important difference between aiming at 'new' learning goals and aiming at other innovative design goals, some of which we will discuss below. Rather, one of the challenges when we aim at established learning goals is to relate the **innovative other design goals** to **existing** assessments.

One good focus for innovation in a project starting with existing assessments could be to cast existing assessments into the form of digital assessments [66]. Such an assessment might be a set of digital closed questions, but it might also be more advanced. In fact, a whole class of innovative design goals is defined by the field of computer-based assessment [66].

Literature on ICT in education provides many more innovative design goals that, at least in principle, do not imply new learning goals. Examples are goals that are formulated in terms of an integrated learning experience, a life long learning attitude, levels of motivation, learning efficiency and support of distance learning. In particular *distance learning* becomes more relevant every year. Here *distance learning* means that the students are not physically present in any of the buildings of the university. Further articulation of this design goal might imply that students should achieve the same learning goals in the same study time as in a regular course setting. This naturally suggests research into virtual environments as one of the research goals. For instance, for this purpose virtual microscopy experiments have been realized [90] and are currently being evaluated in a series of case studies. This implies that students acquire the skills of using microscopes and interpreting microscopic images by carrying out tasks with a **virtual** microscope.
Another innovative design goal related to 'old' learning goals might be to provide students with a completely *integrated learning experience* [3]. In Chapter 2, we defined ‘integrated learning experience’ as a learning experience that does not involve any form of cognitive load due to switching between different tasks or between the use of different media. Such a goal tends to point to what we will call *‘eBook Plus’ functionality* (see Box 5-1).

Box 5-1 eBooks

> The term eBook is widely used for a range of readable and viewable digital resources.
> In this thesis, eBook functionality means that the student can view and read from the screen and that the possibilities for navigation are at an abstract level the same as those that are available in a printed textbook. However these navigation possibilities are supported in such a way that the required effort, both mental as well as physical is much less than in a printed textbook. The main possibilities are, page turning, opening a specific chapter or section by clicking on its heading, opening a specific paragraph by clicking on a keyword in an index, opening a specific paragraph by clicking on a cross reference. In addition eBook functionality includes free text search. Finally, eBook functionality usually includes the possibility to run video clips and animations on part of the screen. Many publishers now support eBook formats. Some eBook formats require dedicated devices. Other eBook formats can be read within a browser.
> We define 'eBook Plus' functionality as extended eBook functionality that includes activating learning objects based on interaction types or based on animations, which in turn may be driven by simulations (see also [245]). In a university context we would also require that lecturers should be able to update the learning material with little effort [47, 246].

5.2.3 Aiming at established learning goals: shifting 'centre of mass'

A DOR project aiming at digital learning materials in higher education is likely to involve research that is relatively generic once it can start with established and well defined learning goals. The relationship with support for learning certain tasks that are very much related to specific subject matter, is still there, but it tends to be an indirect relationship. In this subsection, we will illustrate that the 'centre of mass' of the efforts in the project tends to shift into the direction of knowledge domains such as Information and Knowledge Management (IKM) or Information and Communication Technology (ICT) or even into other directions.

For instance, in process engineering education we might want to avoid any cognitive load involved in switching from reading a text to setting up a set of linear differential equations (see for instance [96]). Alternatively, in the context of learning to apply a psychological theory about nutrition behavior, we might want to avoid any cognitive load involved in switching from reading a text to defining a questionnaire (see for instance [137]). In both examples there is some relationship between the design requirement (i.e. no extraneous cognitive load due to switching between activities) and the subject matter. However this relationship is very indirect in comparison to the relationship between a learning goal and the subject matter. Moreover, not only is the relationship with the subject matter differs both cases, but also the technical challenge. This illustrates that the design goal of 'providing an integrated learning experience' will not be related to the subject matter in the same way as a learning goal is related to subject matter.

The same applies for 'supporting low-cost updating of learning material'. If the learning material is purely text based, such support will be very different from supporting low-cost updating of learning material that relies on three-dimensional visualization. Within many social science disciplines, purely text based learning material can often be satisfactory. Within many engineering disciplines and natural science disciplines such as chemistry, there is much need for three-dimensional visualization (see for instance [247]). Thus, the relationship between low-cost update functionality and knowledge on the subject matter or knowledge on how students are likely to learn the relevant subject matter related tasks is still there, but it is an **indirect** relationship.

DOR aimed at the combined design goal to provide an integrated learning experience to students and low-cost update functionality will require relatively little knowledge of the subject matter domain and relatively much expertise in the IKM domain. In most contexts in which the digital learning materials will operate, it will also be necessary to involve expertise on eLearning standards such as SCORM 2004 [202] and QTI 2.0 [122].

More generally, for a project aiming at an innovative 'other' goal, the DOR project, i.e. a project in the right hand columns of Table 5-1 and Table 5-2, is likely to require considerable input from knowledge on IKM. This may imply that the project team will invest considerably in the realization of generic functionality. Obvious examples of investments in generic functionality are investments in interaction types and investments in *eBook plus* functionality.

However, also a tool or environment that has a tight relation with the subject matter and of which the development is requested by the Subject Matter Expert (SME) can be much more generic than one might expect. For instance, the need for a virtual microscope and digital cases that are based on this microscope and 'integrated' with this microscope came primarily from cell biology courses. On the other hand, a virtual microscope can actually be incorporated in many other courses as well. Design and realization of a virtual microscope and corresponding digital cases will be partially generic, partially dedicated to specific categories of courses.

Within a certain course, the quality and the functional scope of the virtual microscope should be determined by the operational design requirements that articulate the learning goals. Apart from design requirements that articulate learning goals there might be other design requirements. These will come from students and lecturers involved in those courses. For instance, there might be requirements regarding a specific type of interaction that appeals to those students and lecturers.

Across different DOR projects on digital learning materials for different courses, other design requirements will carry more weight. Examples are requirements that define possibilities for reuse and requirements that aim to enable low cost implementation in different courses and faculties. Realization of these requirements requires knowledge of IKM and knowledge of organizational contexts as well.

For two reason it is important to indicate explicitly what a specific project goal will imply for the 'centre of mass' of the efforts in that project. Firstly, it is important for making connections with the right knowledge domain and being aware where the necessary knowledge is most likely to be found. Secondly, it is important for timely selection of the right journals and a timely indication of the required background of peer reviewers. Alternatively, when the research project is set up with a specific balance between certain disciplines this is likely to have consequences for the resulting focus of innovation in the design goal.

5.3 The second major distinction: type of learning goal or target competency

The second dimension along which we should distinguish design goals for digital learning materials in DOR projects is defined by the type of goals or type of target competencies. In this subsection we will distinguish three types corresponding to three rows in Table 5-1 and Table 5-2.

In several WU projects we have initially referred to the Four Component Instructional Design Model (4C/ID [12]). More recently a representation of the four components and corresponding guidance on how to use these [14] was published that requires much less knowledge of cognitive science than the original book. This version should be comprehensible to Subject Matter Experts (SME's). In the later version the four basic components of the 4C/ID model are 'Learning Tasks', 'Supportive Information', 'Procedural Information' and 'Part-task Practice'. The authors define 'learning tasks' as "authentic whole-task experiences based on real-life tasks that aim at the integration of skills, knowledge and attitudes." [14]. We will call the corresponding type of goal 'whole-task competency'.

'Supportive information' is relevant for the non-recurrent aspects of the learning task. Balance equations and their meaning typically constitute 'supportive' information in many sciences. 'Supportive information' should enable elaboration and help to acquire understanding. 'Procedural information' is relevant for the routine aspects of the learning task. The information on how to insert a specific unit operation in a flowsheet in a design environment such as SuperPro Designer [248] is an example of procedural information. 'Part-task Practice' aims at the routine aspects of the task and at achieving a very high level of automation. Complete incorporation of the four components in a university context is often not possible because the aim of university teaching at the level of a specific course is not always 'whole-task competency'. In particular, achieving a very high level of automation was often not an explicit element of our goals.

Against the background of the 'norm' of 'whole-task competency' as a learning goal we distinguish (see examples in the corresponding rows of Table 5-1),

(1) learning materials that are primarily aimed to support acquisition and **understanding** of concepts, laws, methods and other forms of **knowledge**,

(2) learning materials that are primarily aimed to support acquiring an **overview** and understanding of certain authentic **tasks** and

(3) learning materials that aim to support students in becoming competent to carry out a **complete** authentic **task**.

5.3.1 Aiming to support learning and understanding of concept, laws and methods

Activating digital learning material that is intended to support learning and understanding of a range of concepts, laws and methods, can consist of a set of digital exercises and questions with feedback. Sometimes it is possible to incorporate these exercises and questions in one or more cases. Recall that we have defined a case as a combination of a situation, an assignment and a role for the student. Such a case can be regarded as a network of interactions (see Figure 2-1). Such a case illustrates the relevance and some of the possible uses of the concepts, laws and methods that the students need to learn and understand.

Often it is not possible to design a case or set of cases that provides adequate opportunity to support all learning objectives of a course or a topic. In those cases that primarily provide a context for a cluster of exercises and questions, the task represented in the case is not the primary learning goal. The primary learning goal is an aggregate of the learning objectives that are defined by the exercises and questions. The exercises and questions enforce active learning and should stimulate induction (in particular, generalization and discrimination) and *elaboration* [14]. *Elaboration* refers to a "category of learning processes by which learners connect information elements to each other and to knowledge they already have available in memory" [12]. Moreover, the exercises and questions provide opportunities for the student to display behavior. We need this behavior in order to acquire information about the extent of their understanding [136].

For generic concepts, the range of exercises and contexts that are intended to define understanding can be very wide. Being able to apply a concept in a task and/or context that is very different from previously executed tasks and/or contexts, is called ‘Far Transfer’ [249]. Perkins & Salomon [250] argue that such transfer requires “intimate intermingling of generality and context specificity”. For instance, we might define ‘understanding of the Reynolds number’ as being able to calculate the Reynolds number for a set of combinations of geometrical form, fluid parameters and velocities, and to characterize the corresponding flow situations and as being able to identify flow situations in which the Reynolds number is actually not defined. Merriënboer et.al. [14] advise much practice and many different exercises in random order. However, the planned study effort of most courses in a university is usually too low to provide time for many different exercises.

When the learning objectives are that the student has knowledge of and understands concepts, laws and methods, the functions of the case are to promote a perception of relevancy, to reduce the amount of context presentation (because the same context is used in more than one exercise or question) and provide a context that inspires the designer of the learning material to formulate exercises and questions. However, the skills that the student should demonstrate in the assessment must preferably be analyzed and formulated in terms of the concepts, laws and methods and not per se in terms of the context provided by the case. Thus, the exam questions, which are operational definitions of the learning goal, will often be defined apart from the case environment that is used during the course.

When the learning objectives are more directly task related (see below) the cases and case environments will also be more directly related to the task, and the operational definitions of the learning goal will also be more directly related to the task.

5.3.2 Supporting acquisition and understanding of overviews of authentic tasks

In many courses in higher education the learning goal is to acquire insight in the main structure and characteristics of certain authentic tasks in research and in the role of specific concepts, laws and methods in these tasks. For such learning goals a number of cases were developed in the FBT program. A few examples are: a 'Mixing and Oxygen Transfer' case, a 'Membranes' case and a 'Heat Transfer' case [3], a 'Personalized Diets' case, an 'Obesity' case and a 'Leptin Pills' case [2], a 'Downstream Processing' case [146], a 'Brain' case and a 'Light Induction' case [99]. Such cases are highly streamlined and stripped of almost all details. Understanding of concepts, laws and methods is still relevant but much more relevant is their role in the authentic tasks. Operational definitions of learning objectives that articulate this 'task overview' goal, should therefore be much more tightly related to the case than operational definitions of learning objectives that articulate the 'understanding' goal. Actually, the case defines the learning goal.

5.3.3 Aiming at whole-task competency

We define *whole task competency* as the competency to carry out a specific task completely, i.e. including all its details. Usually 'whole task competency' in higher education requires more than understanding. A whole task might be to formulate a research question, design an experiment in order to answer this research question, plan and carry out the experiment, analyze and interpret the results and write an article presenting this work. The main distinction between whole task competency and having knowledge and understanding of an overview of an approach to an authentic task is that the former requires also that the student has adequate routine on recurrent demands of the task. Digital learning materials enable students to carry out tasks with a high degree of authenticity. In this context a high degree of authenticity, means that the task is largely the same as a task in real life. Such tasks are in practice always composed of many sub tasks. Also the environment in which the tasks will be carried out is an environment that is normally used in the corresponding professional context. Examples of such environments are SuperPro Designer [248] in a process engineering context or SPSS [251] in an applied data analysis context. Examples of learning materials that come somewhat in the neighborhood of supporting the acquisition of such whole-task competency are described by [37, 92, 137]. Even these examples provide relatively little exercise aimed at routine. All in all, most FBT learning materials aimed to support the acquisition of concepts and understanding in the context of a whole task and to provide insight in whole tasks, but did not aim at building routine on recurrent cognitive skills in the context of an authentic task.

5.3.4 The role of the discipline of Information and Knowledge Management

Table 5-2 shows in the right hand column a considerable list of innovative contributions of IKM. Combining innovations along these lines with defining new learning goals is in principle possible. In practice, designers and SME's may even be tempted to do so because they underestimate the task of defining and articulating new learning goals. However, at the start of the project it is important to be aware that combining the innovations in the right hand column with defining and articulating new learning goals will imply a research project that requires input from more knowledge domains, and will be far larger and more complex than 'staying within one column'.

5.3.5 Consequences of learning goals that are strongly whole- task related

Focusing on learning goals that are strongly related to a whole authentic task such as in section 5.3.2 and section 5.3.3 has important consequences for the operational definition of the learning objectives and thus for an exam.

Many learning processes as well as assessment processes in higher education will require the student to execute a set of tasks. In an often used paradigm, the student is said to be supported by 'scaffolds' during a sequence of tasks and successive tasks have to be carried out with less and less scaffolding [14, 216]. Means for scaffolding are for instance guiding questions, analogies, hints and feedback. The learning goal is often that the student can carry out the task with a well-defined maximum set of scaffolds in specified conditions. Alternatively, the learning goal is often that the student is able to carry out the task without any scaffolding. This line of reasoning implies a definition of an assessment and therefore a definition of a design requirement. In order to make these requirements operational, we distinguish two main approaches.

One approach is to specify requirements with respect to the students' performance or with respect to the results of the students' performance. For some tasks, this can be formalized in advance. For instance, we can ask the student to design a purification process in a given virtual environment and define a set of specified requirements for the design that the student has to deliver. We can also set requirements on the students' path in the design space. For instance, we might require that certain types of mistakes are never made or not made more than once. Or we might require that certain sub tasks never take more than a specified amount of time. Thus, for a task like this we can define the learning objectives operationally before the student takes the exam. For a task like this, we should also be able to realize computer-based assessments.

An alternative approach is to require the student to carry out the task in front of an expert panel and/or to submit the results to an expert panel. Then, the expert panel has to decide if the performance and or the results are satisfactory.

Note that both approaches require expert panels. In the first approach, the expert panel has to decide that the task, the requirements on student performance and the requirements on student output are satisfactory. In the second approach, the expert panel has to decide if the task, the student performance and the students' results are satisfactory.

Thus, there are three moments to define the learning objectives operationally: (1) before designing the case (2) just before the exam (3) after the exam.

For those FBT projects in which the students learn by carrying out tasks in a class of digital cases it would be natural to test for the ultimate achievement of the learning goals on the basis of an 'unscaffolded' task in the same 'family' of cases. Thus it would make sense to automate assessment of achievement of the learning goal based on an assignment that requires the student to carry out an unscaffolded task in the task environment. Indeed, in the last years of the FBT program and in the DiMoBio B-Basic project [252], several SME's have suggested this. In none of the FBT projects, we have done this yet.

In summary, an obvious challenge for design-oriented research as follow-up to FBT projects is the challenge of designing and realizing corresponding digital assessments. There are three reasons to allocate resources to this challenge. The first and main reason is that this implies a sharp definition of the learning objectives. It forces us to express what we really want to achieve. A secondary reason is related to the need for large-scale use of DOR products as already mentioned in section 1.7.2. This need will also induce a need for automated assessment of large numbers of students. The third reason is that a number of teachers already asked for this.

Note that realizing computer-based assessment in this way is much more generic than operationalizing a learning objective into a specific exam question. Further generic [8]research will have to address the following main problems: (1) how to define student behavior in terms of variables that can be monitored effectively and efficiently, (2) how to define scoring rules for student behavior based on these variables, (3) how to define scoring rules for student results, and (4) how to define the boundary between satisfactory score and unsatisfactory score. The first question requires much knowledge from the domains of computer science and information systems. The other three questions require much knowledge from the domains of assessment in education and of psychometrics. As to defining scoring rules for student results, wide availability of innovative closed-question types might suggest that at least for these question types the discussion about scoring rules has been concluded. However, recent research shows that even for multiple-response question types the theoretical justification of scoring rules is incomplete and far from trivial [253]. Computer-based assessment will also involve the question on how to organize secure web-based exams [254].

5.4 Concluding Remarks

In this chapter we have defined six types of projects aiming at innovative digital learning materials in higher education. We have explained that in design-oriented research for such learning materials it is of strategic importance to choose a clear focus for innovation.

Along one dimension we can distinguish between projects aiming at new learning goals or target competencies and projects that aim at established learning goals. A major challenge in projects defined by the former is the articulation of the learning goals into learning objectives and corresponding operational design requirements in

---

[8] Here we use the term 'generic' to indicate research aiming at solutions that can be used across a wide range of different subject matter knowledge domains.

the form of assessments. In this respect, innovative projects that do not imply new learning goals or target competencies are very different. Such projects can make use of existing definitions of learning objectives and existing assessments. Such projects might focus for instance on another goal such as high motivation, an integrated learning experience as focus for innovation, a high student/staff ratio, et cetera.

The second dimension is based on the type of learning goal or target competency. Along this dimension we have distinguished projects that aim primarily at understanding of concepts, laws or methods, projects that aim primarily on overviews of specific approaches to specific authentic but streamlined tasks and projects that aim primarily at the competency to carry out a specific task completely including all procedural details.

These dimensions define six major types of DOR projects for digital learning materials. The classification is intended to raise in an early stage of the project awareness of the most important relevant knowledge domains, of the different primary audiences for whom the results will be sufficiently interesting, and to provide direction to the definition of design requirements and therefore to evaluation as well. In particular when the project is a PhD project, it makes sense to aim for a good match between the competence and prior knowledge of the PhD candidate and the focus of innovation. A conscious allocation of the focus of innovation to one of the six classes helps to prevent the project side slipping into a class for which the right resources are not sufficiently available.

Goals will often change during the project. Nevertheless initial formulations of goals will have to be evaluated before budgets are assigned to the project. Several authors (e.g.[255, 256]) discuss the question which design research project proposals should deserve funding. For design-oriented research proposals this is also a relevant question. The arguments provided in this chapter suggest that the six fold classification of DOR projects for digital learning materials is a tool that can support decision making with respect to funding for such DOR projects. At least, proposals should be realistic and consistent with respect to which cell in the matrix would best characterize the output of the project and how much capacity to which knowledge domain should be allocated.

Tables 5.1 and 5.2 highlight that each of the types of DOR projects defined in this chapter tends to require a different configuration of disciplinary input. While interpreting Table 5.2 one should keep in mind that a DOR project in on digital learning material in higher education does not aim to provide innovation in general fields of learning, instruction or research in education. Rather, a DOR project may rely on established literature from such fields. Furthermore, the contribution of ICT and IKM is not restricted to software engineering and programming. In particular, IKM aims to extend the design space and to provide new possibilities to express educational goals.

Table 5-1 DOR projects aimed at digital learning materials: a typology

	new learning goals	*established learning goals → other goals*
'deep understanding'	1. Learning objectives still have to be articulated. 2. Learning objectives have to be operationally defined; when we aim at digital learning material the most concrete approach is to define the learning objectives in terms of digital interaction types. 3. This implies design and realization of a set of interactions. 4. These interactions can be given some coherence by aggregation into a case [3].	1. Develop new interaction types for instance interaction types that include manipulation of three dimensional objects (e.g. [138]). 2. Map available operational definitions of learning objectives onto digital closed questions based on interactions [131]. 3. Design materials & systems for heterogeneous target populations. 4. Aim at integrated learning experience [47, 257]. 5.Aim at motivation level [52].
'task overview'	1. Analyze overall task class. 2. Design/develop task environment and several cases including learning paths in this environment. These allow student to carry out instances of the task class; initially much support, gradually less. E.g. in [146], student should learn overall approach on 'how to design a process for purifying a 'fermentation broth' or in [134] 'how to build qualitative models of early development processes in living organism'. 3. Invest little effort in procedural information.	Provide guidance for structured approaches to problem (SAP) solving [191]. For instance design/develop an environment for students to design/develop/run models [95, 96].
'whole task'	As in cell above but in addition: 1. Just In Time (JIT) presentation of procedural information. 2. More part task practice. 3. Follow "Ten steps to complex learning." [14].	The difference with the innovation in the cell above is that projects in this cell will make use of professional task environments. Examples are Mathematica, Matlab, MathCad, SPSS, SuperPro Designer. A real wet laboratory is also a professional task environment. 1. Provide seamless connection from virtualized version of task to real version of task. 2. Provide performance support. 3. Enable low-cost part task practice. 5. New forms of on-line communication. 6. Computer-based assessment of the task see [66] for specific challenges.

This six fold classification aims to raise awareness of the different relevant knowledge domains and suggest which research platform is most likely to be interested in the project output. The right column gives a small selection of many possible innovations.

Table 5-2 DOR projects aimed at digital learning materials: focus of innovation ←→ knowledge domain

	new learning goals	*other goals*
'deep understanding'	Articulating new learning goals and corresponding assessment will require much effort from SME's [55]. This articulation may partly rely on new interaction types that extend the design space for digital learning material and provide new ways to express forms of 'deep understanding' that could not be expressed before.	A project in this cell will most likely have its centre of mass in overlapping knowledge domains : ICT, IKM, HCI, learning technology. Some examples are: 1. Design of new interaction types. e.g. interaction types based on requests to manipulate 'three-dimensional' images (e.g. based on VRML). 2. Provide a technical implementation for QTI 2.1 3. Develop a design and realization environment for web-based digital learning materials*. 4. Design adaptive systems as empty shells that can be filled with learning materials (e.g. a system that adapts selection of closed questions to student (e.g. [49]). 5. Reduce the need to make assumptions about prior knowledge of students in the target population. [151, 258] 6. Enable integrated learning experiences (see [47, 257]).*
'task overview'	Designing and developing a virtual task environment for new learning goals/target competencies implies a process of means-ends analysis. The new means are mostly in the domain of IKM. The SME's will try to express their ends in terms of these newly suggested means. .	While the learning goals/target competencies are well established, focus of innovation can be on innovative interfaces, improved synergy with LMS's and on reusability. The more we focus on reusability, the more generic the value of the efforts in this cell will be. The FBT projects demonstrate that the class of web-based virtual task environments contains very diverse tasks. Instances are doing lab experiments, generating microscopic images and operating chemical reactors. This induces the need for an environment for designing and building virtual task environments. This design and realization environment will have to include a dedicated library of object classes.*
'whole task competency'	A project in this cell should mainly be defined by the SME's. SME's should take care to make use of only those interfaces of other knowledge domains that they can understand without much additional effort.	A project in this cell will most likely have its centre of mass in the IKM knowledge. Some challenges are: 1. To provide integrated learning experiences involving the transition from the from virtualized version of task to the real version of the task. 2. To provide efficient and effective ways of synchronous communication as part of the integrated learning experience. 3. To make effective use of progress in IKM technology [see 258] in supporting achievement of whole task competency. 4. To enable computer-based assessment of the task (see [66]).*

For goals marked with * a start has been made for research beyond the FBT program
ICT = Information and Communication Technology; IKM = Information and Knowledge Management; LT = Learning Technology; HCI = Human Computer Interaction; LMS = Learning Management System

Chapter 6 Requirements engineering for large-scale use

Abstract[9]

Sustainable quality of design and realization of digital learning materials will only be possible when these learning materials are used by many teachers and students. Articulation of the design goal of digital learning material should imply the formulation of design requirements that must be satisfied in order to realize large-scale use. This chapter lists and discusses eight potential scenarios for large-scale use of digital learning materials in higher education and defines and discusses design requirements that are important for large-scale use.

[9] a version of this chapter has been accepted for publication in the Journal of Universal Computer Science as:
Hartog, R., A. Beulens, and J. Tramper,
"Digital Learning Resources in Higher Education: Designing for Large-scale Use"

6.1 Introduction

In Chapter 1, we have already argued that sustainable quality of learning materials requires large-scale use. This chapter is based on the hypothesis that increasing numbers of users enable increase of the quality of digital learning materials and that quality as enabled by increase of number of users can increase considerable before saturation will set in.

In this chapter, we discuss potential large-scale use scenarios in higher education. Next, we articulate the large-scale use goal in terms of design requirements for digital learning materials in higher education and discuss the consequences.

6.2 Eight scenarios for large-scale use

In this subsection, we give a short description of a number of well-known scenarios for large-scale use of digital learning materials. In practice, many hybrid forms of these scenarios are being tried out as well.

6.2.1 The cooperating SME's scenario

In this first scenario, Subject Matter Experts (SME's) of the same discipline share learning materials across different institutes of higher education. Each SME develops some learning material but the learning material is used by many other SME's. This scenario is already very old and has been successful in the past (see for instance [259, 260]). However, for sustainability this scenario relies on stable personal contacts and not primarily on links between positions in the university. When a person moves to another position within the university or to another job, there is a considerable chance that his contribution to sustainable large-scale use of the material that he was using is discontinued. Use of FBT learning materials at universities outside Wageningen

University (WU) was mostly based on personal contacts. Indeed this use often stopped when the teacher who used the material moved on to another position.

6.2.2 The funded cooperation scenario

In this scenario, cooperation within a disciplinary field is supported by a program that is funded by governmental or non-governmental organizations. Examples are NuGO [261], B-Basic [252] and ALTB [66, 262]. Often in such a disciplinary cooperation, one of the goals is to ‘spread’ or ‘disseminate’ knowledge. One way to do this is by the design, realization, implementation and use of learning materials in dedicated work-packages or sub projects [2, 66, 263].

6.2.3 The Learning Object Repository scenario

A Learning Object Repository (LOR) is a searchable store of digital learning objects that can be accessed over the internet. With the term LOR scenario we will refer to a cluster of scenarios ranging from completely open to the world to relatively restricted access. Over the years there have been many such repositories or referatories (i.e. 'portals' or 'brokers' that connect to digital learning materials in other repositories). Some that are currently still accessible are ARIADNE [264], the BEN portal [265], DLESE [266], MERLOT[267], the Open Educational Resource initiative [268], SCORE [269], SMETE [270] , Wikimedia-Commons [271], Wikiversity [272] and WISC-ONLINE [273]. Most of these are or were partly supported by funding organizations such as the National Science Foundation or the European Commission or cooperatives of schools and/or universities.

For university teachers a successful LOR scenario implies that the effort of finding the learning material they need within the LOR should be 'small'. This effort includes the effort of formulating search questions, filtering the results, downloading and installing the learning objects[10] and evaluating the learning objects. Thus, the LOR scenario requires at least support for standard search functions and corresponding metadata. Here, metadata are data that provide information about the learning object. Ideally, these metadata should include descriptions of the interface of the LO's and ratings by SME's.

Large-scale use of learning objects in a LOR scenario is only likely when the following conditions are met. Firstly, a 'large' number of university teachers and students must know which concepts, topics and methods for which subjects are supposed to be covered by the LOR. What the university teacher needs is a short and clear description of what he may expect to find and what not. This in turn requires coherence. The less coherent the contents of the LOR, the more extensive the description of these contents will have to be. Secondly, the probability that a 'quick' search based on these expectations fails, must be 'low'.

On the one hand, we cannot determine if a certain LOR and its contents satisfy these two constraints as long as 'short', 'clear', 'coherent', 'large', 'quick' and 'low' are not operationally defined. On the other hand, for the concepts, topics and methods that were relevant in the FBT program, we performed many times a search of about an hour. None of these searches produced descriptions of available learning objects or resources that led to a decision to adopt these objects/resources in one of our projects.

[10] Over the last years more and more interactive learning objects and resources become available that can be 'run online'. However, there are still many interactive learning objects and resources that need to be downloaded and installed on the computer of the user.

The general LOR scenario is not aimed at one discipline or one category of learning goals. Also a direct connection with a disciplinary community of lecturers and therefore a direct connection with those who can express the needs for learning material and with potential peer reviewers is less direct.

6.2.4 The Open Courseware scenario

In an open courseware (or open educational resources) large-scale use scenario, an organization[11] provides free web-based access to many or all resources that have been developed within the organization. The examples that are probably best known are MITOPENCOURSEWARE [274], OpenLearningInitiative [275] and LearningSpace [276]. Recently, institutes that provide open courseware are offering more and more dedicated links to access activating or interactive digital learning resources such as applets. However, most learning resources currently made available as open courseware are mainly presentational. In the institutional open courseware scenario, benefits for the university seem to be in the first place corporate image benefits. Depending on her primary source of funding, a university might also regard it as a moral obligation to provide free access to the learning resources of the university.

Open courseware is likely to be very valuable in many countries or regions in the world were university-level teaching capacity for certain knowledge domains is not locally available. In many of those countries or regions, lecturers are scarce and the workload for lecturers is often too high to be able to achieve satisfactory quality levels. Many institutions and foundations are funding capacity building projects for developing regions. One approach is often to send experts to the region in order to temporarily support and train local SME's. Another approach is often to provide grants

[11] or individual

for students from those regions to study in developed countries. The latter approach tends to lead to brain drain. Beyond a break-even number of locations, providing digital learning materials to local university teachers will clearly be cheaper than both of these approaches. Thus, the option to provide web-based activating learning materials in order to support local university teachers in regions with scarce SME capacity, deserves to be considered as a valuable complement to traditional capacity building approaches and a reason to support open courseware initiatives.

Over the last few years, the movement towards open access and open educational resources seems to gain impetus. More and more, universities invest in setting up attractive web-based repositories for open courseware. Faculty desiring to draw attention to their activating digital learning resources and to provide easy access can make use of the university's open courseware repository. This saves faculty the work of setting up such a repository themselves or submitting their applets on some of the repositories listed in the previous section. When a university invests in an open courseware system, it is likely that the efforts of faculty for submitting their digital learning resources to that open courseware system will be relatively low. It is yet too early to establish the benefits of open courseware initiatives for faculty or the extent to which open courseware initiatives contribute to sustainable large-scale use of learning resources.

6.2.5 A scenario limited to a single type of learning objects and relying on sponsors

In the PhET approach [277], the interactive learning objects are computer simulations that provide primarily opportunities for inquiry-based learning. The learning objects are JAVA or FLASH applets. Much effort has been invested to lower any barrier that might impede teachers or students from anywhere to run the simulations or to download the simulations. Sustainability relies on what the authors call the 'Mother Teresa Model', i.e. on charitable support from public and private foundations [278]. In the case of PhET this has been very successful. In 2008 already millions of downloads were reported (see supplementary material to [279]). While the PhET project is one of the most inspiring projects for digital learning resources in higher education, bootstrapping an approach like PhET is likely to be beyond the possibilities of faculty in most universities[12].

6.2.6 On campus large-enrollment course scenarios

In many universities a number of the courses have enrolments of hundreds of students. For such a course it is attractive to design, develop, implement, use and evaluate digital learning resources. The first reason is that sufficiently high enrolment numbers enable acceptable costs-per-student. Of course, one should still quantify 'sufficiently high' and 'acceptable', but this is out of the scope of this chapter because actual costs and benefits will vary a lot across different institutes in different countries. The second reason is that in large-enrolment courses there is a strong need to leverage capacity of the teacher. Digital learning resources are tools that can provide such leverage. In on-campus large-enrolment courses, the main issue is not how to lower barriers for other faculty to adopt the learning resources. Moreover, in this scenario, more than in any other scenario, the benefits of investments in digital learning resources will be experienced by the primary problem owners, i.e. by those faculty who are involved in design, realization, implementation, use and evaluation of the resources for their own course. In this scenario it will still be a challenge to shift the investments to an earlier point in time: instead of incurring costs during course activities in a number of years,

---

[12] The Khan Academy
280. Academy, K. *Khan Academy*. [last accessed May 15 2012]. http://www.khanacademy.org/. falls outside the scope of this thesis. Initially, most of the material in the Khan Academy was purely presentational and also it was not directed to higher education. However, the history of the Khan Academy is instructive for individual SME's in higher education. It shows that barriers for developing short movies based on screen recordings, and for making these available to students are very low. Moreover, it is likely that demand for such movies is high not only in secondary education but also in higher education.

the design and realization of digital learning resources incurs costs that have to be made in advance.

6.2.7 Distance learning large-enrollment program scenarios

Currently more and more web-based distance learning programs are initiated by universities [281, 282]. The challenge for a university is to identify the worldwide demand for the knowledge that matches her core competence and to match this demand with suitable distance learning programs. The reasons to take up this challenge may vary from a public sense of responsibility to the world to an expected return on investment. In general we believe that sustainability of a distance learning program will rely on possibilities to reduce the necessity of communication between teachers and students and secondly, to reduce costs of communication between teachers and students, both without loss of quality. The first calls for high quality digital activating learning materials. The second typically calls for application of knowledge management concepts such as 'active documents'. For instance, a student may ask a question about a phrase in such an 'active document', this question is answered by an SME, the question and the answer are stored with the document, later questions are automatically compared with questions that are already answered by SME's and if there is a match automatically answered [151].

6.2.8 A publisher's scenario

A publisher will define and implement a business model. A business model should describe the products, services, business processes, resources, supply chains, customers, value propositions, and a revenue model. For instance, a number of publishers currently develop digital learning material in connection to textbooks. When a university teacher decides to prescribe the textbook in his course he can import a corresponding course cartridge in the Learning Management System (LMS) of his university [283]. Then, the LMS makes the digital learning material in the course cartridge available to the students who are enrolled in the course. More and more publishers currently offer cartridges for major LMS. Alternatively, a publisher can host a publisher based LMS and provide students the opportunity to buy access rights to digital versions of textbooks or chapters of textbooks or to *eBooks*. The costs of access to these learning resources are considerably lower than the costs of buying the corresponding hardcopy. Furthermore, the digital learning resources and the hardcopy of the textbook essentially can provide different value to the student. For instance, digital learning resources can include sound, video and interaction, while on the other hand a traditional hardcopy does require no technology in order to be read.

Some publishers offer teachers the possibility to configure their own common cartridge or textbook or eBook, using resources made available by the publisher. In addition they offer support of an editor to the teacher. Alternatively or in addition, they provide web-based access to digital learning materials on a publisher managed system. Thus, a student can, instead of or in addition to buying a textbook, buy access to the publisher's learning management system [284].

Currently, several publishers are experimenting with different business models along these lines. Thus the abstract description of this scenario actually represents a whole class of different 'publisher-based' scenarios.

The most important aspect of these publisher-based scenarios is that these scenarios are most likely to provide adequate resources for marketing and communication.

In the remainder of this chapter we will discuss a number of design requirements for digital learning materials that match specific large-scale use scenarios. The first set of design requirements is based on the fact that the university teachers are the ones who will decide which materials will be used in university education.

6.3 Basic technical design requirements

When we aim at worldwide large-scale use, the definition of the target population is not per se part of the design requirements. For this use, the target population of a module of learning material is partly defined by the assumptions in the interface. For instance, when the interface assumes a bandwidth of 54 Megabits/second this will de facto exclude a number of regions in the world. This is in particular relevant for large-scale use scenarios such as a scenario aimed at leveraging SME capacity in developing regions. It may also be relevant for a large-enrollment distance learning scenario. Furthermore, even though publishers aim primarily at teachers, they will not want to restrict their target population too much by making too strict assumptions about the technical facilities that are available to these students.
When technical requirements are important, a requirement that limits the necessary bandwidth is likely to be the most important[13]. Requirements that limit screen resolution or processing power seem less important. Rather than formulating low screen resolution and processing power constraints, one should make available some up to date hardware and aim for a learning scenario that assumes a limited number of computers (laptops or tablets)[14].

6.4 Design requirements related to the evaluation effort of teachers

For a lecturer who has seldom or never reviewed digital learning material for use in his own course, such a review may often take more effort than reviewing a corresponding

---

[13] When a teacher in a region with low bandwidth wants to implement the learning resources in his/her course, it will be impossible for him/her or for the developers to make higher bandwidth in that region available.
[14] e.g. when a teacher wants to implement the learning resources in his/her course, and does not have hardware with the same screen resolution and processing power as that of faculty in the development team, one of the option is to make available one computer and classroom screen.

part of a textbook. Firstly, an experienced reader can scan pages of text relatively fast. Scanning multimedia learning material requires much more time because of its sequential nature. Secondly, we tend to believe that generally, a teacher reviewing a textbook seldom takes the time to answer all the questions and make all assignments at the conclusion of each chapter. It is our impression that teachers primarily evaluate a textbook on the basis of the presentational parts of the textbooks. On the other hand, many digital learning materials designed in WU projects were digital cases. A digital case is a combination of a situation, an assignment and a role for the student. In order to evaluate such a case or any other type of interactive learning material, teachers want to 'walk' through these materials and cases.

Thus we have to design digital learning materials that require low review effort of a teacher who wants to implement the material in a course. Of course 'low review effort' will still have to be defined. Currently we might define 'low effort' as 'comparable with the review efforts required by a text that covers the same subject matter'. Comments of different experts who reviewed parts of the material in the FBT program suggest that this new design requirement at least implies that an evaluator must be able to:

- continue at any moment from any computer,
- if desired, reset at any moment his state within the learning material (restart at the beginning),
- walk back and forth through any of the possible learning paths with 'previous' and 'next' buttons, without having to answer each question and without having to complete each assignment,
- immediately see what learning objectives will be achieved with the learning material,
- scan all content-related chains of inference in the material without having to carry out each inference step,
- separately inspect each model that is incorporated in the material.

In fact the last two requirements are classical requirements for knowledge based systems and intelligent tutoring systems [48, 285]. In these systems subject matter knowledge is represented in a specific format that is readable by SME's. A set of inference rules that are relatively generic constitutes a separate inference engine.

Another way to enable university teachers to quickly scan content related chains of inference and models would be to document the materials by means of a visual design language analogous to blueprint formalisms used in construction. Research in this area is only quite recently receiving more attention (see Box 2-2 Obvious candidates for representing the design of digital learning materials).

Different scenario's induce different requirements. The requirements in this subsection are primarily relevant in scenarios in which teachers need to decide on the use of materials developed elsewhere. The requirements in this subsection may be less urgent in the Open Courseware scenario or in large-enrollment scenarios in which the teachers who use the material are very much involved in the design and realization of the material.

6.5 Design requirements aimed at low implementation efforts of the teachers

Experience in the FBT program has made clear what type of efforts implementation of FBT materials in courses at other institutions does require.

Within 'cooperating SME's ' scenarios, we have some experience with implementation efforts for FBT materials at the Technical University of Łódź, Ecole Polytechnique Fédérale de Lausanne, Cornell University in the US, Asian Institute of Technology, Graz University of Technology, The Royal Veterinary and Agricultural University in Copenhagen.

Within 'funded cooperation scenarios', we have some experience with implementation efforts within the NuGO organization [261], (see also [2]) and a number of efforts in the B-Basic DiMoBio project [263].

These experiences suggest three categories of implementation efforts and corresponding design requirements aimed to limit these implementation efforts:

(1) efforts related to 'making space' for new learning goals within a course,

(2) efforts related to realizing an adequate learning scenario,

(3) efforts related to providing authorized access to the learning material.

6.5.1 'Making space' and design requirements

Learning materials focused on new learning goals imply that other learning goals in the course or curriculum must be achieved in less time or must be removed. We refer to this as to 'making space'. In the FBT program, the effort associated to 'making space' for new learning goals has already been invested within the context where the learning material was designed and developed. Implementing that learning material in an existing program or course of another university tends to induce 'making space' discussions. As the other university was not involved in the initial design this should not be surprising.

The effort of 'making space' will be related to the study effort that is imposed by the learning material. A larger study effort goes hand in hand with a need for making more space.

Efforts to 'make space' for new learning goals, point to the role of DSE journals or to leading disciplinary journals in the field (see Sub Section 1.1.2.). We believe that these journals are the place for the discussion about introducing new learning goals and dropping old learning goals. Note that a discussion about dropping old learning goals becomes more relevant as soon as the possibility to achieve new learning goals becomes more tangible due to proofs of feasibility and proofs of concept.

One requirement for digital learning materials focusing on new learning goals is that university teachers quickly can grasp what students will learn from this material and how much study effort this will cost. Ideally, this would imply that a package of digital learning material includes the operational definitions of the learning goals and objectives, i.e. a set of corresponding exam questions or assignments. The latter, together with a normative indication of the student's study effort should clearly convey the weight of the new learning goal per unit of study effort. This approach could also take away part of the load of the teacher to develop the exams. In the WU projects, we have no experience with this approach.

6.5.2 Learning scenarios and design requirements

Even though the intention of the FBT projects was to deliver learning materials that can be used in a range of different learning scenarios, most case studies in the FBT projects were based on the following learning scenario.

- Initially, students work more or less synchronously on the same cases in one computer room at the university.
- A staff or faculty member is present for technical problems, error corrections and students who want to question and discuss issues that actually reach beyond the learning objectives.
- Students are stimulated to work together, for instance in pairs.
- Part of the time a number of students will work in other settings, for instance at home, in order to finish their assignments.

At WU, this learning scenario was also believed to provide an appreciated alternation with other learning scenarios such as lecture-based learning scenarios and problem-based learning activities. In addition, at WU as well as at other universities, ‘live’ interaction between students and teachers and students among each other is often highly valued. This was another reason to promote working with the learning materials in one computer room and working together. Learning scenarios that imply a high level of such direct ‘live’ interaction are favored at many universities where such a scenario can be realized.

In many universities, the mostly used learning scenario in the WU projects requires the lecturer who is responsible for the course to invest time in organizational efforts. The lecturer has to make a reservation for a computer room and make sure that the right desktop - or laptop configuration is active at the right moment. The currently prevailing technical and organizational structure in many universities makes this a time consuming and error prone activity. We estimate that this usually will require somewhere between four to sixteen hours of a teacher's or assistant's time. At least in the universities that used WU materials this activity often required too much attention from the lecturer. Besides, this learning scenario requires the availability of computer rooms. For certain timeslots or even during the whole year, computer room capacity can be scarce. If this is the case, this tends to require early reservations and additional organizational efforts. Of course, such issues will not be relevant in universities where every student has a laptop and wireless access to the local area network of the university.

Alternatively, the learning scenario wherein the lecturer uses the material as the core of his lecture (see section 6.2.5) tends to require less organizational effort. Moreover, this learning scenario can be carried out in universities that do not have sufficient computer room capacity. Another learning scenario that requires less organizational efforts from the lecturer, is the scenario wherein the students use the material in their own time and in a place of their choice and where the students are responsible for their own laptop or desktop configuration. Experience with this scenario would also be important for use of learning materials in distance education.

We conclude that an important requirement for realizing the large-scale use goal is that the digital learning materials, just like textbooks, can support a range of learning scenarios and not just one learning scenario. This range of learning scenarios should in particular include lecture-based scenarios and learning scenarios in which students do not necessarily have to work in one room.

6.5.3 Design Requirements Related to Authentication, Authorization and Integrated Learning Experience

In most universities, an LMS is now a standard component of the facilities that support teaching and learning. The prevailing paradigm for handling digital learning resources is that the teacher configures a collection of resources for his course in the LMS. In particular, this may involve learning resources that have been uploaded by the teacher into the LMS. Alternatively, a learning resource might be a web-based application that 'lives outside the LMS'. We would make such a learning resource available to the student from within pages generated by the LMS. This is because we want a configuration that provides an integrated learning experience within the LMS environment. A learning experience is called integrated if it does not involve any form of cognitive load that needs to be attributed to switching between different tasks or to switching between the use of different media and if any other effort needed for such switches is negligibly small.

Authentication and provision of authorized access is straight forward for resources that have been uploaded and are stored within the LMS. Providing authorized access to resources that 'live outside the LMS’ requires a protocol for communication with the system that manages those learning resources. Establishing the implementation of such a protocol across universities in practice still tends to require interaction with administrators of LMS's of different universities [257] and in practice sometimes also involvement of the teachers of courses.

In order to provide the student an integrated learning experience and in order that the learning resource can delegate certain tasks to the LMS, it is a requirement that both conform to a common interface.

In relation to the upload paradigm, the most well-known specifications for such interfaces have been SCORM2004 [202], IMS Content Packaging [286], and Common Cartridge [283]. These interfaces aim to be standards that will realize interoperability. In the context of this chapter, interoperability means that any learning object or package or course cartridge can be functional in any LMS as long as both conform to the same specification. Non-conformance to SCORM2004, IMS content packaging and IMS Common Cartridge specification can impose additional implementation efforts on teachers who use an LMS, thus contributing to a barrier for large-scale use.

A publisher aiming to sell copies of learning resources to universities that want to incorporate these learning resources within their own LMS's will probably define conformance requirements as to SCORM2004, IMS Content Packaging and IMS Common Cartridge. Conformance to these specifications might also be required by funding agencies in a funded cooperation scenario.

In relation to paradigms that provide web-based access to learning resources, applications or tools that live 'outside' the LMS, the Learning Tools Interoperability (LTI) specification has been developed [287].

However, it is important to be aware that interoperability is not a strict requirement in several other scenarios. Scenarios that are not based on providing an integrated learning experience within a certain LMS environment could ignore the specifications mentioned above. For instance, the Open Courseware scenario is more aimed at students than at teachers. The intention is rather that students can directly access the learning resources and there is no role for an LMS. Thus, in an Open Courseware scenario it is less likely that conformance to learning technology standards is a design requirement. Moreover, the Open Courseware scenario is primarily a scenario in which the institution provides free access to resources that are developed for students already enrolled in a course of the university.

In large-scale use scenarios that are not primarily based on cooperation, but, for instance, at realizing an attractive business model there can be two reasons to forsake conformance to standards. Firstly, specifications that are adopted as standards such as the SCORM2004, IMS Content Packaging and IMS Common Cartridge are likely to be compromises. Consequently, these specifications are likely to be inadequate for certain specific large-scale use scenarios. Secondly, when a large-scale use scenario is based on realizing a competitive edge, this competitive edge might be realized by defining interfaces that provide more possibilities than the standards. Alternatively it is likely that a competitive edge might be realized with approaches in which the standard interfaces do not make sense. For instance, a publisher aiming to sell a learning service and attract students to her own server might well decide to set her own standards.

The same may hold for a university aiming at a distance learning program. In a competition between a few universities with overlapping core competences each aiming a distance learning program at the same target population of students, there is no reason to invest in interoperability of their own systems and resources with those of other universities. On the contrary, each university might primarily want to realize a competitive edge. However, when a consortium of universities decides to set up a common distance learning program, interoperability becomes very relevant.

The conclusion is that the importance of interoperability requirements depends on the large-scale use scenario at which one aims.

6.6 Design requirements related to prior knowledge of students

We already stressed that, when we aim at worldwide large-scale use, the assumptions in the interface of a module of learning material de facto define the target population. Alternatively, for learning material that will be used within a course in an existing program within a university, the target population is already defined. For this material, a design requirement is that the assumptions in the interface will hold for 'this' cohort of students being registered for 'this' curriculum in 'this' year. If not, implementation of the module or learning object may involve additional efforts in order to enhance the prior knowledge of some or all of the students in the target population (see Box 6-1). This is in particular relevant for large-scale use scenarios such as the cooperating SME scenario, the funded cooperation scenario and on campus large-enrollment courses. Within one university, it is usual to derive assumptions about the prior knowledge of the student from a description of courses that the student already is supposed to have

completed. Furthermore, we often assume that students in one group or one class have more or less the same background in terms of previously attended courses.
For a worldwide target population this approach of deriving assumptions about the target population does not work. A usual approach is to define prerequisite knowledge and clarify this definition by providing self tests [66].
An alternative approach is to minimize the number and scope of assumptions about the prior knowledge of members of the target group. This is also relevant for a publisher trying to define a target population. In order to handle this, we distinguish three approaches: designing self-contained learning material, designing adaptive learning material and adaptive systems, designing learning material that can be extended 'on the fly' when necessary.
A body of self contained learning material provides the student with all information and opportunities for learning that are needed to achieve a learning objective or a set of learning objectives. A body of self contained learning material is based on minimal prior knowledge assumptions. Learning objects are self contained but a textbook or a course can also be self contained.

Box 6-1 Implementation efforts due to wrong prior knowledge assumptions in the FBT program

In two subprojects in the FBT program, assumptions about prior knowledge turned out to be wrong.
The first case studies reported in [99] revealed implicit assumptions with respect to the prior knowledge of the students. This concerned lack of knowledge about a number of necessary concepts (requiring 'supportive information' see [14]). For this reason, the next time when the materials were used by a comparable target group a 'basics' case providing supportive information about the necessary concepts and opportunities to work with this information were provided in advance.
When the first cohort of students worked with the Downstream Process Designer it turned out that students who had never designed before, could not handle the very 'open' character of a design assignment [146]. This led to the embedding of the DSPD in a case to give students of the next cohort in the same target population more structure and a clear design goal.

One problem with self contained material is that it imposes unnecessary cognitive load on students who do already know much of what is necessary to achieve the learning objectives. These students have to process a lot of information that does not or only very little contribute to knowledge construction related to the learning objectives. Another problem is that designing and developing self contained material often implies a considerable investment.
The first problem of self contained materials can be solved by matching the needs of each individual student to the presentation of materials. For instance, the Proteus system [49], dynamically measures the performance of the student with respect to learning objectives and makes at any moment a selection of a specific question to offer the student. Systems like Proteus adapt what they present to the individual students. Adaptivity can go much further than Proteus. For instance, an adaptive system might 'know' for every user which definitions are mastered by this user and present a specific definition only to a user who does not yet master this definition. The type of adaptive systems that we described here are de facto still self contained. However, contrary to a self contained book, the system avoids imposing unnecessary cognitive load for the individual student. The system does this by filtering out interactions and presentations that refer to knowledge which the student sufficiently masters.

The second problem of aiming at self contained learning material is that assuming no prior knowledge at all is not realistic. On the one hand, there will always be people for whom a body of learning material is not self contained. On the other hand creating material that aims to be self contained often implies an investment of which a considerable part may never be used. Thus it may happen that an adapted system after a few years in use still has not presented some of the materials to any user. In fact, self contained materials and adaptive systems that are essentially based on self contained content, imply an investment in a large stock of material. In many industries and supply chains we would prefer just in time (JIT) production instead of production to stock.

In higher education, we are used to approach the second problem by answering questions of the student in the lecture room. This can be regarded as JIT production, but at the same time these answers evaporate and cannot easily be reused. An alternative is to design systems that enable interactions between student and teacher to be attached to a specific anchor in the learning material and thus enhance the learning material, for instance with 'active documents' as described in section 6.2.7 (see also [47, 151])[15].

6.7 Design requirements related to the size of what to deliver

A lecturer will not easily invest much time in implementing one small piece of learning material in his course. It will be different if one implementation effort is adequate for a large body of learning material, many students and several years. Thus, most lecturers will implicitly look at the **relative** implementation effort $\mathcal{E}_{\texttt{relative}}$.

$\mathcal{E}_{\texttt{relative}} = \mathcal{E}/(\boldsymbol{n} * \boldsymbol{e} * \boldsymbol{y})$

where

$\mathcal{E}$ is the absolute implementation effort

$\boldsymbol{n}$ is the average number of students that will use the material in a specific year

$\boldsymbol{e}$ is the average number of hours of study effort that the material generates and

$\boldsymbol{y}$ is the number of years for which the implementation effort is valid.

As long as $\mathcal{E}$ is relatively large, the barrier for the lecturer to actually make this effort will be lower for larger packages of learning material i.e. for a package that covers a larger study effort and a larger part of the course.

---

[15] compare this e.g. with the possibility to attach comments to a specific word or paragraph in a document.

Also, larger packages will make it easier to minimize the number and scope of prior knowledge assumptions because the necessary knowledge can be provided in the package.

In general, design and realization of large packages will imply incorporating many well-established learning goals and may shift the focus of innovation towards other goals. For instance, larger packages will tend to require more attention for the goal of the integrated learning experience. In Chapter 5 we have argued that this shift of focus of innovation towards other goals tends to a shift towards DOR projects that require less subject matter knowledge and more at ‘ICT and higher education’ (such as knowledge on learning tools interoperability) and/or Information and Knowledge Management (IKM). In addition, the necessity to combine the DOR tasks with routine design tasks will require special attention, both in carrying out the projects as well as keeping track of the focus for intended publications.

6.8 Conclusion

It is likely that sustainable design and realization of innovative high quality digital learning materials for higher education is only possible in design & realization teams and not by individual teachers (see also [184]). Furthermore, this sustainability requires effortless retrieval of, and seamless access to digital learning materials. We doubt if the first three large-scale use scenarios will ultimately be able to sustain both such teams as well as an organization that realizes large-scale retrieval and access. Rather, these first three scenarios can provide good points of departure for a distance learning scenario or for one of the publisher-based scenarios.

We believe also that no sustainable large-scale use scenario exists that relies only on funding for DOR projects. While innovation will be relevant in all large-scale use scenarios, the major part of the digital learning materials to be produced will have to be the result of routine design. An isolated focus on new learning goals and new target competencies resulting in 'unconnected islands of innovative digital learning material' is unlikely to result in sustainable availability of digital learning materials. This implies that future DOR projects for the design of digital learning materials should be embedded in a program for realization of a sufficiently large body of digital learning materials. For most subject matter fields this program will require much routine design and realization. Defining the interface of the DOR projects with this wider program is not within the scope of this thesis.

Different large-scale use scenarios can lead to different priorities and different design requirements. For different large-scale use scenarios we have articulated a number of design requirements and discussed the consequences.

In the last decades the concept of learning object has received much attention. An ideal has been sketched of bodies of learning material that can be configured on the basis of LO's. However, in some large scale use scenarios the primary demand is delivery of relatively large coherent modules. We now believe that in these scenarios, early focus on strict design requirements for digital learning material aimed at the LO-LMS paradigm is counter productive. In this respect it can be instructive to recognize that also in other areas besides the field of digital learning materials early focus on a high modularity has been contested [147, 148].

Even though there are worldwide hundreds of thousands of learning objects available, large-scale use at the level of **individual** interactive learning objects has not yet been as successful as expected. We believe that such reuse requires the existence of **well known clusters** of learning objects. These clusters should be sufficiently coherent in order to be able to describe them concisely. They should also provide almost complete coverage of what may be expected based on the description. The developments of the last two decades suggest that expecting the spontaneous emergence of such clusters within the next five years is not realistic. This suggests that it may be better to start with developing packages of digital learning material that are in terms of coherence and completeness very much like most textbooks.

There are several other arguments to aim design and realization at such packages. Firstly, the relative costs (i.e. per student, year and hour study effort) of implementing such a package within a university course would be relatively low. Secondly, with respect to prior knowledge, minimizing assumptions with respect to prior knowledge in the interface at the package level will in general be easier than minimizing assumptions in the interface of a learning object. The main reason is that any necessary information can be built into the package. Thirdly, informing lecturers of universities worldwide about a package requires considerable less effort than informing lecturers about individual learning objects.

When the primary aim is to design and deliver coherent packages, this requires much routine design and realization of digital learning materials as well. This again points to larger design and realization contexts such as a publisher's program or a distance learning program. It is likely that those questions within such a program that require a DOR approach are not directed at new learning goals but at innovations that are more generic as in the right hand columns of Table 5-1and Table 5-2. The main challenge is to design innovative responses to such questions and at the same time keep SME's in various subject matter domains involved.

Publishers have since long been able to realize successful business models for realizing large-scale use of textbooks. Since a few years, publishers are starting to offer several services that are based on digital learning materials. Without access to rather detailed financial statements it is difficult to be sure about success of certain business models. Currently, it seems that already several successful business models have evolved. We believe that one approach for combining DOR on digital learning materials with realizing large-scale use will have to fit within a publisher's business model. The most obvious other opportunities for realizing large-scale use of digital learning resources in higher education are based on large enrollment in university courses. Such large enrollment could be in on-campus courses but also in distance education.

Large-scale use considerations are important in the very first definition of DOR projects aiming to deliver innovative digital learning materials in higher education. In this very first stage, it is already important to foresee the consequences described in this chapter

Chapter 7 Evaluation in design-oriented research and in the Food and Biotechnology (FBT) program

Abstract

This chapter discusses evaluation in Design-Oriented Research (DOR). The discussion is based on the concepts that were defined in Chapter 2, 3 and 4. The main concepts are design goal, operational requirements, operational guidelines, and output-classes. Every output gives rise to product evaluation and possibly also to evaluation of the process that led to the product. For each of the output-classes defined in Chapter 4 opportunities for evaluation are articulated in this chapter. It is argued that evaluation efforts should be in balance with the claims that are made. Claims should be supported by corresponding evaluation results. However, if a claim with little weight requires expensive evaluation efforts, then it is better to drop the claim. Finally, the chapter includes a discussion on evaluation in Design-Oriented Research (DOR) projects in the Food and BioTechnology (FBT) program.

7.1 Introduction

In Chapter 2 we defined design-oriented research (DOR) as research that primarily aims to produce an innovative design and applies typical concepts of design methodology.
This implies the following research questions:

1. what are, in a specific real life context, goals that make sense and why,
2. how can these goals be articulated in terms of measurable quantities,
3. is it possible to achieve these goals,
4. if so how?

Evaluation aims to determine the value of outputs of a DOR project in relation to these questions.

With respect to the first two research questions, the discussion in sections 2.5 and 4.2 showed that a goal is seldom very clear without corresponding operational requirements. For this reason we have defined the combined formulation of the design goal and the operational design requirements that articulate this goal as **one** output-class.

The third question can ultimately only be answered once a realized design (i.e. an instantiated artifact) has been implemented in an intended context and has been used. A positive answer to the third question is the basis for a proof of feasibility. Such a proof requires outputs in the following output-classes: case study results, implementation records, a realized design, goals & requirements and possibly also a scenario. A scenario for using the instantiated artifact is not necessarily output, it can also be input for the DOR project. Thus, it might be that the scenario is part of the context definition (formulated in terms of assumptions in the interface). In case of learning material, the scenario is a learning scenario. For a proof of feasibility an abstract description of the architecture and most of the interfaces is not strictly necessary.

In order to determine if the goal has been achieved we must measure the values of variables in operational design requirements and determine if they satisfy these requirements.

Box 7-1 The distinction between summative evaluation and formative evaluation

A well-known distinction in social sciences is the distinction between summative and formative evaluation [83].
Product evaluation of the implemented instantiated artifact in operation is often called summative evaluation in order to distinguish it from formative evaluation. Summative evaluation is intended to answer the question if we have achieved the goal. Formative evaluation is intended to provide information that can help us to make improvements in order to come closer to the goal. Often formative information requires us to measure more variables than just the variables in the operational design requirements. For instance, when the goal is that 80% of the students in a cohort successfully completes an assignment and there is doubt if the goal will be achieved, it makes sense to define additional variables to be measured. These should provide information that gives direction to activities for improving the design or its realization. Alternatively, a formative evaluation might just be a set observations. For instance, a recording of all actions of a student working with the learning material might give information that gives directions to improvements.
In educational research and instructional design the definitions of summative evaluation and formative evaluation are often slightly different. For instance, Romiszowski [31] writes:
"It has become fashionable to speak of
1. summative evaluation, which sums up the results or outputs of a course. This normally takes place after the course is over and has no effect on the structure or processes of the course.
2. formative evaluation, which measures the outputs specifically in order to change the form of the course to modify either its structure or the processes which go on. Formative evaluation takes place during the course as well as at the end of the course."

Summative evaluation aims to determine if the goal has been achieved and to relate achievement of the goal to the operation of the realized and implemented instantiated artifact. For a proof of feasibility a *summative evaluation* is needed. Most literature in social sciences and in educational research distinguishes *summative evaluation* and *formative evaluation*. *Formative evaluation* aims to determine what changes in the design or the realized design are needed to achieve the goal. Box 7-1 gives some background information on this distinction.

Finally, the answer to the fourth question is implied by the architecture and interfaces, one or more (learning) scenarios, the realized design, the implementation records and a record of a realized learning scenario. These can be considered 'milestones' or 'landmarks' along a 'path through design space' (see section 2.4.8). A more complete trail of a path that has actually been traversed would require the use of a design rationale system (see e.g. [288]).

In this chapter, we discuss for each of the output-classes opportunities for evaluation activities. Table 7-1 displays an overview of these output-classes and the potential value for an output in each of the output-classes. Evaluation should produce arguments and evidence for the actual value of an output. In addition, we discuss evaluation in the projects of the FBT program and provide suggestions for 'getting more' out of future PhD sized DOR projects.

7.2 Balancing evaluation efforts with the value of the evaluation outcomes

Evaluation implies measurements, interpretation of measurement results, and expert reviews of theoretical outputs and of the given interpretation of the measurement results. These efforts are necessary to validate claims with respect to outputs in the output-classes described in Chapter 4 and listed in Table 7-1.

The more interesting these claims, the more important it is to validate these claims. It would be wrong if so many efforts have been invested in delivering a potentially valuable output, that no resources are left for its evaluation. On the other hand, claims with little value deserve little effort.

Table 7-1 Overview of DOR tasks, output-classes and potential value per output-class
The ordering in the table does not indicate a time ordering. Architecture & Interfaces can be described after the delivery of the instantiated artifact. Design requirements can be formulated parallel to realization. RQ = Research Question within the DOR project.

TASK ***sub tasks***	***Output - class***	***Potential value of output is based on***
DEFINE GOAL ▪ Carry out focus group interviews with representatives of stakeholders. ▪ Carry out surveys aimed at communities of experts. ▪ Search for design patterns for operational design requirements.	Goals & Operational Design Requirements	▪ importance of goal ▪ clearness of goal
DESIGN BLUEPRINT ▪ Distinguish levels of decision making when making decisions about architectures. ▪ Study available architectures. ▪ Search and use design patterns. ▪ Use an adequate available representational formalism.	Abstract Representation Architecture, Interface Definitions, Interactions	contribution to answer on RQ-4 "… if so how", and in particular in ▪ the way in which the goal has been/might be achieved ▪ which assumptions are based on basic research and which assumptions are based on other considerations
DESIGN SCENARIO NOTE: details are outside the scope of this thesis.	Learning Scenarios	contribution to answer on RQ-4 "… if so how"
REALIZE ▪ Make explicit choice for technology. ▪ Use an adequate integrated design and realization environment.	Instantiated artifact (Realized Design)	contribution to answer on RQ-4 "… if so how", and in particular in ▪ the way in which the design can be realized with the technical means and the resources that were available at the time of the project.
USE ▪ Make a plan for implementation. ▪ Carry out implementation according to plan. ▪ Evaluate for each assumption in the interface if the assumption holds. ▪ Identify implicit assumptions that don't hold. ▪ Prepare a protocol for the case study. ▪ Take care that the protocol aims at the intended claim. ▪ Carry out case study according to protocol.	Case Study Results a. implementation record b. measurement results c. proof of feasibility d. proof of concept	contribution to answer on RQ-3 "… possible to achieve the goal?" and RQ-4 "if so how…" ▪ proof of feasibility ▪ proof of concept

In line with most literature on evaluation (see also section 2.7) we define *product evaluation in DOR* as: determining if an output in an output-class satisfies a set of requirements.

In section 7.3, we will scan outputs in each of the output-classes for evaluation opportunities. We define *process evaluation in DOR* as: determining if a process or task that produces an output satisfies a set of guidelines or procedures. In section 7.4 we discuss opportunities for process evaluation. In addition, we compare the identified opportunities for evaluation with the actual evaluation experience in the FBT program and suggest improvements that satisfy the 'little value, little effort' criterion. Finally, this scan of output-classes and potential evaluation activities results in a number of questions for further research.

7.3 Product evaluation in DOR

7.3.1 Product evaluation of goal & operational design requirements

The main questions to be answered by product evaluation of the goal are:

- is the goal clearly defined?
- does the goal make sense?
- if so why?

In answer to the first question we aim to define the goal in terms of operational design requirements. In Chapter 5 we made a major distinction between new learning goals as focus for innovation and other innovative goals.

7.3.1.1 Evaluating new learning goals & operational design requirements

For learning goals, defining operational design requirements is defining assessments, for instance exams. Product evaluation of these assessments is mainly a matter of evaluation by reviewers. In particular, Subject Matter Experts (SME's) should confirm that the assessment articulates the new learning goal. In addition, SME's should give their opinion on the value of the innovative learning goal. As to the former, literature on assessment in education provides sets of requirements for assessment questions and for assessments [289-291]. Examples of such requirements (see for instance [290]) are 'the item should not imply a trick question', 'should not be opinion-based', 'should be grammatically correct', 'as brief as possible' etcetera. These requirements are in fact meta-requirements. They allow an SME to evaluate an assessment not only with respect to the content validity, but also with respect to generic aspects of assessment in education.

Currently it is uncommon to approach evaluation of assessments in higher education in a scientific way. We have not been able to find publications or evaluations of complete assessments for subjects in the field of Food and Biotechnology in higher education. Currently, it is also not usual to provide much room for publication and argumentation of the operational definitions of the requirements, i.e. the assessment, in DSE journals and other journals mentioned in section 1.1.2.

In a number of FBT publications, we did include a typical exam question. While these exam questions did provide some insight as to the operational meaning of the learning goals, these exam questions were not accompanied by strict directions for scoring nor were they closed questions. Moreover, for none of the exams operational requirements as to precision or reliability were specified. This is in contrast to engineering design were we would often see specifications of required precision as well. Thus, exams and questions were not fully operationalized. Literature suggests that the challenge to realize more and better defined assessments will require considerable research effort [27, 66, 120, 244].

One approach to improve evaluation of design goals and corresponding operational design requirements requires that the SME's who publish in DSE journals and other journals as mentioned in section 1.1.2, acknowledge the importance of operational definitions of innovative learning goals and the necessity to review and share such definitions (for inspiration on how this could be done see e.g. [197-199]). In other words, ultimately, assessments should get the scientific status of other measuring instruments, such as for instance intelligence tests.

Improving evaluation - 1 One of the first improvements in future DOR projects should be to organize reviews of learning goals and corresponding assessments by independent SME's. When assessment primarily has to measure ‘understanding’ and not has to take into account integration of skills such as laboratory skills or presentational skills, it should be possible to formulate a large part of the assessment in a format that just consists of text, formulas, diagrams and other pictures [66]. The costs of one SME evaluating such an assessment does not have to be high. For an exam to be completed by students within three hours, the time needed by an SME to evaluate that assessment in relation to the learning goals and target competencies is likely to be of the same order of magnitude.

Improving evaluation - 2 Product evaluation of assessments has not been carried out systematically in the FBT program. For future DOR projects aiming at digital learning materials in higher education we propose at least one formative evaluation of each assessment before the assessment is taken by the students. It is very unlikely that an assessment cannot be improved after its first conception [55]. Furthermore we propose that SME's carry out a summative evaluation of the assessment in a later stage. In case of an assessment instrument that has been used for some time, statistics of the answers or scores for parts of the assessment can be used to guide such evaluation.

Finally such assessments should be published in DSE journals or in the scientific journals of the subject matter discipline.

The costs of producing output in this first output-class in terms of assessments are considerable [55]. The costs of evaluating these assessments in the way suggested above are negligible in comparison to other costs in DOR for digital learning material. The additional benefits are twofold. First, if a goal is achieved it becomes very clear what goal is achieved. Second, they focus attention on the value or importance of the goal.

During the FBT program it became more and more clear that it should often be possible to found an assessment on a version of a digital case. This is in particular a good option when we aim at design competencies or modeling competencies such as in most of the FBT projects. If assessment is based on a version of a digital case, evaluation of the assessment by Subject Matter Experts (SME's) would at least require a 'walk through' of this digital case. In these situations, expert evaluation of the assessment actually largely overlaps expert evaluation of the realized design of the digital learning material.

Further research - 1. Further research should be aimed at assessments based on the performance of the student in a digital case.

7.3.1.2 Other components of the goal, other operational design requirements

A number of FBT projects were not primarily or not exclusively aimed at support for achieving new learning goals. Additional requirements were operationalized in terms of questionnaire questions. For instance, students were asked to select an option from the set {'strongly disagree', 'agree', 'neither agree nor disagree', 'disagree' , 'strongly agree'} in response to a statement such as "I enjoyed studying the case". With respect to certain aspects of questionnaires, experts on questionnaire design were consulted. Transcripts or relevant sections of these questionnaires were published in refereed scientific journals (see for instance [36, 49, 135]). Thus, for these additional requirements the publications allow readers to understand their meaning.

In analogy with evaluation based on psychometric analysis of assessment results, it can make sense to evaluate the questionnaires by means of analysis of completed questionnaires. For instance, when more than one question is used to define one underlying variable (often called 'construct') it makes sense to measure the internal consistency between these questions. The need to do so will depend on the actual set of operational definitions. In the FBT program no such analyses have been carried out.

Recall that in DOR, the formulation of the operational design requirements often will lead to a goal formulation that differs from the initial goal formulation, which was the basis for allocation of project resources. In other words, the goal formulation that is input is most likely not the formulation of the goal that is the output.

Evaluation of the resulting goal and its articulation in terms of requirements will be based on meta-requirements. Table 7-2 displays some examples of such requirements on requirements. These meta-requirements are mostly not sufficiently operationalized. In the FBT program we have not used any formal approach to evaluation of the articulation of the goals, nor have we used a formal approach to determine the value of the resulting goal.

Further research - 2 How to evaluate in practice different DOR goals requires further research. Only for some pairs of goals, the question how to decide which goal is more desirable is easy to answer. In particular, if two goals differ only with respect to one variable with an ordered domain, it will often be clear which goal is more desirable. For example, suppose that two goals only differ with respect to the percentage of students in a cohort who achieve a certain level along the dimension of that variable. In such a case, the goal with the highest percentage of successful students is also the goal that is more desirable than the other.

7.3.2 Product evaluation of abstract representation of the design

7.3.2.1 Product evaluation of architectures

Product evaluation of architectures will mainly be a matter of evaluation by the scientific community based on the system of refereed scientific journals. Recall that the external architecture is the set of architectural aspects that are directly related to the experience of the user. The main requirement for the external architecture is that it is consistent with available knowledge from fields related to cognition, learning and instruction. For instance, the external architecture should satisfy constraints that are represented by the contiguity principle "place printed words near corresponding graphics" and "synchronize spoken words with corresponding graphics"[292]. Moreover, it should be possible to realize a technical (or internal) architecture that enables the external architecture.

Table 7-2 Examples of requirements on outputs in output-class 1: goal & operational design requirements

Generic meta-requirements: o The combination of goal and its articulation should be clear to all stakeholders. o The costs of the measurement procedure attached to each requirement should be in balance with the value (i.e. perceived importance) of the information that results from applying the measurement procedure. o The operational design requirements must be consistent with the goal. o The goal must be ethically acceptable. o The stakeholders together must have promised to support the goal. o If the goal and requirements are claimed to be derived from a demand, there must be empirical evidence for such a demand. Specific meta-requirements on requirements defining an innovative goal for digital learning material: o The formulation of goal and operational requirements reflects a clear focus for innovation (Chapter 5). o The formulation of goal and operational requirements fits at least one large-scale use scenario (Chapter 6). o A subset of the operational design requirements should define the assessment of the learning goal. o SME's confirm that the assessment de facto defines an important learning goal. o Variables such as 'increase of motivation', 'study efficiency', 'the students' perceived understanding of …' , et cetera are operationally defined in terms of questionnaire questions. o Requirements with respect to observed behavior of students who use the material include an observation protocol.

The first important requirement for the technical architecture is that it can be realized with available technology and that it enables the desired external architecture.

Other requirements on the technical architecture will be indirectly derived from a large-scale use scenario. In Chapter 6, we saw that different large-scale use scenarios may lead to different preferences as to architectures. For instance a scenario of cooperating universities may assume a service oriented architecture [44, 293]. Product evaluation of the technical architecture will have to take into account the relevance of such technical developments related to technical architectures. Product evaluation of representations of technical architectures often takes place within working groups in professional organizations such as IEEE [223], ADL [224] and IMS Global Learning Consortium [225]. In addition, it will be partly evaluation by referees of scientific journals in information systems design or computer science. In the FBT program, contributions of this type were restricted to the concept and realization of author-defined data storage [5, 6] and to the concept of adaptive selection of digital closed questions to students [49]. Reflection on published outputs of the FBT program displays a rather fragmented picture of the FBT architecture.

Improving evaluation - 3 We have argued that a good and explicit representation of an architecture and its underlying argumentation is a potentially valuable output of a DOR project. Such an architecture should be evaluated against operational design requirements. Such an architecture should also enable a specific articulation of theory of learning and instruction for specific subject matter. In particular, this refers to the type of activities and feedback that are enabled by the architecture. Evaluation will also have to take into account the consequences of specific choices with respect to large-scale use scenarios.

Further research - 3 In DOR projects in the follow up of the FBT program, new design requirements are being formulated. These requirements are still formulated in terms of shortcomings of the fit between standards for LMS's and educational server applications such as the downstream process designer or the Proteus system. Some attempts have been made to realize synergy between such educational server applications and LMS's [257]. The most relevant recent development in this respect has been labeled Learning Tools Interoperability [287].

7.3.2.2 Product evaluation of technical interfaces

Explicit descriptions of the interfaces of learning objects or larger modules of learning material with their intended context should be acknowledged as proper handles for evaluation. Recall that each interface essentially consists of a description of a set of assumptions about its context and one or more functions.

The interface of a learning object or package or any other module will contain a set of technical assumptions, such as 'there will be a browser that can handle the output' or 'there will be an LMS that is compliant to the SCORM 2004 fourth edition specification'. Evaluation of technical assumptions is often straightforward.

7.3.2.3 Product evaluation of 'generic' assumptions about students

The interface of a learning object or a package of learning material will also imply 'generic' assumptions about students, teachers and reviewers. The most relevant 'generic' assumption about teachers and reviewers in higher education is that their attention is a scarce resource. This also holds for students, but for students also assumptions related to learning are important.
With 'generic' assumptions about students in the target population we mean that these assumptions hold for all these students. Evaluation of these 'generic' assumptions should be based on the requirement that scientific literature in cognitive science and on learning and instruction reports sufficient evidence to support such assumptions.
In reflection on the FBT outputs we can formulate only few underlying 'generic' assumptions:

- attention and learning are positively related [294, 295],
- attention and motivation are positively related [296, 297],
- perceived relevance, confidence and satisfaction are positively related with motivation, once attention has been captured [298],

- working memory capacity is limited [14, 299].

On the one hand, most FBT publications provide references that can be read as support for these implicit assumptions. On the other hand, a problem with such assumptions is that they are not quantitative. Most FBT products also imply quantitative assumptions but for these we could not find evidence in literature. For instance, many FBT materials reflect the implicit assumption that a good frequency to prompt the student for input is of the order of magnitude of ten times per hour.

Further research - 4 should provide a more complete set of 'generic' assumptions, sufficient evidence for any 'generic' assumption and provide more quantitative indication for the values of the variables in these assumptions.

In addition, the designs were based on many assumptions that were not formulated in a form like the examples above. Much of this knowledge was part of theoretical models of learning and instructional design such as cognitive apprenticeship [216], the 4C/ID model [12] or the belief that a certain specific structured presentation of information is better than another presentation or than an unstructured presentation.

Once more, we emphasize that providing evidence for generic assumptions is not a task that fits a DOR approach. In addition, we recall that it is necessary to make assumptions, even though not all assumptions can be supported by adequate empirical evidence.

We do not see any possibility for improving evaluation with respect to these 'generic' assumptions **within** a DOR project. Currently, it is difficult to find explicit references to the interface paradigm in design in cognitive science literature or literature on learning and instruction. While this literature is usually rather explicit in terms of guidelines (that are mostly suggestions for requirements (see section 2.7.1), we have not been able to translate the presented information in terms of generic interfaces.

7.3.2.4 Product evaluation of prior knowledge assumptions

The main DOR challenge in designing a body of learning material is to formulate interfaces that assume as little specific prior knowledge as possible. For instance, the function of the learning objects in the down stream process design environment [146] are intended to support students in learning what a specific unit operation for purification of bioreactor output can do and how to apply this unit operation in a purification process. None of these learning objects does make the specific prior knowledge assumption that the student knows what a protein is, even though the primary aim of the whole process is to deliver one specific more or less purified protein.

Evaluating interfaces by scanning them for unnecessary prior knowledge assumptions requires at least subject matter expertise. An explicit formulation of these assumptions such that this scan can be carried out quickly would support SME's.

Improving evaluation - 4 In the FBT projects no interfaces where explicitly defined. Thus, in case of expert evaluation experts had to carry out mental reverse engineering on the realized digital learning materials in order to evaluate the interfaces. Descriptions of specific prior knowledge assumptions in interfaces of learning objects and cases would allow a faster and more precise form of expert evaluation. Such expert evaluation should be part of a publication culture in which it is normal to publish interfaces of learning objects in refereed scientific journals.

7.3.3 Product evaluation of a scenario

In the FBT program, we aimed to design digital learning materials that should match not just one new, but a range of existing learning scenarios. For this reason we do not discuss the option of improving the evaluation of a learning scenario or its corresponding design process.

7.3.4 Product evaluation of realized design (instantiated artifact)

7.3.4.1 Product evaluation of the realized design by experts

If there is an abstract representation of a design, product evaluation of the instantiated artifact implies first of all, answering the question if the realized system conforms the abstract representation. This is essentially verification. This implies code inspection by software engineers.
In the FBT program, no sufficiently complete abstract representations of the design have been produced. In some cases 'storyboards' have been developed, but these were by far not complete enough to pass on to a technical developer without additional communication. The main reason for this is that satisfactory representational formalisms and corresponding software development tools that support design and realization based on these formalisms were lacking. When these become available code inspection for verification will become relevant. Apart from this, reverse engineering and documentation of the code can be an approach to deliver the missing abstract description of the design. This will allow for a higher level evaluation of the design.

7.3.4.2 Product evaluation of the realized design by means of systematic testing

For the cases of the FBT program, empirical evaluation by an SME of all possible combinations of user input and screen view is feasible. We call this a complete 'walk through'. More generally, it will depend on the number of possible different states how realistic it is to have really all possible combinations checked by this approach. In most information systems, this option is not realistic but for a limited set of interactions all answers of the user can be tested systematically.

All FBT materials have been evaluated by a few SME's. This involved checking for factual errors, inference errors or calculation errors. A number of materials were also corrected for incorrect use of the English language. Moreover, the SME's gave an overall impression of the materials, including an opinion on the value of the learning goal and the implicit design requirements. For all materials this overall impression was positive.

SME's are scarce and SME's have seldom time to spare. In the FBT program we asked SME's from outside the project, to invest considerable time in hands-on experience with digital learning materials and careful registration of errors and comments. This is much more than what a journal normally asks from a reviewer of a submitted paper. Besides, reviewing a submitted paper is more in line of what an SME is used to do.

In section 6.4 we already described the same problem with respect to the barriers to evaluate the material by an SME in order to decide if the material is useful in his own course. There, we formulated a few requirements to digital learning material that are necessary to keep the effort of evaluating the material within reasonable bounds. However, in order to induce an SME who is not in the position to use the material, these measures are by far not enough.

In addition to the other reasons for providing descriptions of assessments, architectures and interfaces, the difficulty of finding SME's who will review the instantiated artifact itself, is an additional reason for aiming at these descriptions. When well described assessments, architectures and interfaces would be submitted in the form of papers to journals and be considered suitable for review, the consequence would be that they can be adopted in the standard circuit for scientific review.

In general, experience in the FBT projects suggests that we can only require an expert to check a list of requirements and accordingly to deliver a complete and detailed evaluation report if we can offer financial compensation.

7.3.5 Product evaluation of case study results

7.3.5.1 The implementation record

An implementation record makes clear if design requirements have been adjusted or not and if the context in which the design will be used has been adjusted or not.

Adjusting the context may have been necessary because one or more assumptions in the interface of the design turned out not to hold. When adjusting the context is not feasible, it may be necessary to adjust the design requirements and consequently the design and its realization.

One function of the implementation record is to support interpretation of the case study results. If adjusting the context required considerable effort, it might be questioned to which degree goal achievement can be attributed to the realized design in operation. Alternatively, if the design itself has been adjusted, records of these adjustments are needed to update the abstract representation of the design.

7.3.5.2 Product evaluation of measurement results

Evaluation of measurement results implies primarily answering the question if all the necessary data have been produced. In particular, in order to determine if the goal has been achieved, all variables that take part in a requirement must be measured. Additional data may be required to exclude the possibility that the goal was already achieved at the start of the case study[16].
Two sets of measurements may also be needed to exclude alternative explanations for the achievement of the goal (see section 2.10.6.6).
Two sets of measurement results of case studies are not needed when we want to compare two different goals with respect to a certain quality (for instance when we want to discuss if one goal is more desirable than another). **Two different goals** should be compared on the basis of the goal definition (see section 2.10.6). Comparing

[16] Note that not all variables have to be measured at the start of the case study. For instance we can ask in a 5 point Likert item with options running from 'strongly disagree' to 'strongly agree': "My motivation to understand the concept of confounding was raised by this module". And we can require that 80% of the students in the case study answer with 'agree' or 'strongly agree'. Such a requirement would not induce the need for two sets of measurements.)

two goals **as to their feasibility** does require two case studies and corresponding data sets.

With respect to measurement results it is particularly important to evaluate the **measurement process**. This will be discussed section 7.4 below.

7.3.5.3 Proof of feasibility

If all measurement results are in the set of values that satisfy all constraints, which are defined by the context and the operational design requirements, then the goal is achieved. Evaluating this is straightforward.

In addition, a proof of feasibility requires that the goal achievement can be attributed to the use of the instantiated artifact. In particular, we want to provide arguments that exclude other plausible explanations. For digital learning material, this may suggest a randomized controlled trial. This would imply two case studies. Students of a cohort would be randomly allocated to one or the other case study. In one case study, students would use the learning material, in the other, students would not have access to the material. Next, in each of the two case studies, the same set of variables would be measured and the results would be compared. The set of variables to be measured would include the variables that are part of the goal definition. In addition, it may be deemed necessary to measure other variables of which it cannot be guaranteed that they will have the same value in both case studies.

The arguments that are provided for the claim that the use of the design is the only plausible explanation for the goal achievement should be evaluated by experts and in particular by reviewers of refereed scientific journals.

In the relevant FBT projects, we could not imagine alternative explanations that needed to be excluded (see also 2.10.6.6). Apart from the practical difficulties, it would have been very unlikely that the additional cost of carrying out a randomized controlled trial would be balanced by resulting information [255, 300].

7.3.5.4 Proof of concept

When the design actually embodies only one concept, and when the proof of feasibility has been provided we can also claim a proof of concept. When the design is highly modular and a certain module embodies one concept, the argumentation for the proof of feasibility is the same as the argumentation for the proof of concept. However, when the design is based on more concepts that are **integrated** into the design, a proof of concept requires a theoretical link between the proof of feasibility and these concepts. This theoretical link requires transparency of the design. Such transparency must at least be realized by means of clear representations of the architecture and the major interfaces. Often, the strategic design decisions and thus the architecture of the design will be based on a set of core concepts. If so, then it will be impossible, or very expensive, to make a new design, leaving out one concept, and realizing it, implementing it and evaluate it empirically. A good example that illustrates this impossibility is the concept of 'linear case' that is deeply integrated in many of the learning resources that we realized in WU projects.

In section 4.6.4 it is already explained that evaluation of the proof of concept **within DOR** is essentially expert evaluation of line of reasoning that links the concept to the proof of feasibility. In practice, we expect that such a proof of concept will seldom be delivered for digital learning materials that are based on **integrated** application of a set of concepts. Thus, additional evaluation activities aiming at delivering a proof of concept are likely to be a waste of effort. In case of integrated application of a set concepts, DOR is seldom an adequate approach to a proof of concept.

7.3.5.5 Final remarks on case study results

In the FBT program, we aimed to design and develop learning material fit for several learning scenarios and educational settings. The underlying idea is the same as for most learning materials produced by publishers: materials that are only fit for one learning scenario in one university will not be used at a large-scale. A range of case studies would be needed to support claims with respect to the question if the design goals can also be reached in other contexts that satisfy the originally defined assumptions about the context. The experience in the FBT program showed that implementing and carrying out case studies in a range of universities is a task that requires too much effort to be combined in one or a few PhD sized projects that are primarily aimed to produce innovative learning materials.

Finally, it is often practical to combine a summative and a formative evaluation in one case study. For instance, the questionnaires for students in the FBT program partly defined a subset of the operational design requirements, but also contained one or more open questions aimed to generate hypotheses for improvement.

7.4 Process evaluation in DOR

Within DOR process evaluation has primarily two functions.

(1) Improve the process during the process.

(2) Give direction or contribute to product evaluation. In particular, results of measurements that were not according to the originally defined procedure are considered suspect.

Besides, process evaluation may help to improve future design processes. Using process evaluation for the latter would be outside the scope of DOR, but does fit within the scope of several other design-related research approaches.

7.4.1 Evaluation of the measurement process

Operational design requirements describe how variables in these requirements have to be measured. Evaluation of the measurement process implies a check if the actual measurement process was conform the operational design requirement. If not, it may be necessary to correct the data, or if correction is not possible, to drop the data.

7.4.2 Evaluate consumption of resources over time: time registrations

Except for [60, 134, 141], the publications resulting from the FBT program provide little insight in the design and realization processes of the FBT program. Indeed, there has been little formal registration of what happened during the FBT projects. Consequently, process evaluation based on records of execution of each process relies mainly on the memory of the team members.

An important type of process variables is the attention of the designer team dedicated to any specific output class or any characteristic of any output. We cannot directly measure this attention but the average time per day or per week is likely to give an indication. It is possible to formulate requirements on the set of values of such process variables, i.e. to formulate process guidelines. For instance, a guideline could require that the designer at least dedicates an average of one hour per day to attempts in realizing high **ARCS** values. This would mean that the designer at least for one hour a day spends time on conscious searching for ideas and inspiration on how to capture the **Attention** of the student, how to guarantee that the student perceives the **Relevance** of the learning tasks and goals and examples, how to support the student's **Confidence** and how to generate student **Satisfaction**. Another guideline could be that the designer does not spend time on sub tasks beyond his competence. Such specific process guidelines will have to be formulated in relation to the goal of the DOR project. Because the goal will develop during the project, the process guidelines should not be too specific in order to limit the number of adjustments of process guidelines during the project.

Improving evaluation - 5 Process evaluation based on a comparison of actual allocation of attention and capacities with specific guidelines. These specific guidelines have to be formulated within the project and should define the intended allocation of attention and capacities to sub tasks and aspects. Such process evaluation during the project is valuable to improve the process within the DOR project. After a few weeks, this form of process evaluation will require a few minutes a day.

Within the DOR project such process evaluation will make the team aware of project costs that do not contribute to the DOR goal. Also, time registrations can point product evaluators to outputs that deserve extra evaluation attention because the invested amount of time is not conform the intended amount of time.
It should be noted that time registrations as proposed here, provide also valuable information for reconstructive studies aimed to provide guidelines for future projects.

7.4.3 Scheduling issues with respect to evaluation in the FBT projects

Experience in the FBT projects highlighted two timing issues that require much attention in DOR projects on design, realization, implementation, use and evaluation of (digital) learning materials in higher education. First, scheduling expert evaluations within the project turned out to require much attention and communication effort with experts. Second, the university schedule strongly constrains the evaluation schedule. In the next three subsections we will now discuss these issues.

7.4.3.1 The point in time when the assessment becomes available

In Chapter 5 we distinguish innovative designs for which learning goals, learning objectives and assessments still have to be defined, from innovative designs based on well-established learning objectives and existing assessments.
For the latter DOR projects, assessments are already available in an early stage of the project. Even when the innovation would imply the design of equivalent assessment, for instance in the form of a computer based assessment, the learning goal and its requirements are essentially well known early in the project.
For the DOR projects in which the innovation is primarily a new learning goal, we make a further distinction. For these designs we distinguish those that primarily aim at 'understanding' on the one hand from those that primarily aim at larger tasks involving complex cognitive skills on the other hand (see Table 5-1and Table 5-2 lower two cells in left column). Examples of these larger tasks are described in [12, 37, 95, 100, 146].

If the primary aim is: to support the achievement of 'understanding' concepts, then assessment can and should be defined in **an early stage of the project**. In these DOR projects the assessment formulation helps to give direction to the design of the digital learning material.

If the project mainly involves larger composite tasks, **the assessment should rather be a derivative of a case**. For instance, we can use the task of designing a purification process given a set of design requirements [146] or the task of designing an experiment given certain research questions [2, 99, 100] as an assignment. Next we can assess the students' performance on an analogical version of the task in the same digital design environment.

7.4.3.2 Scheduling case studies including summative evaluations

In many universities, a specific subject will be taught only once a year. Furthermore, the schedule for teaching will often be fixed for many years. Within the time frame of a PhD project, there will be only a few possibilities for evaluation of the realized design, implemented and in operation in a regular educational setting. If a DOR project on digital learning materials is restricted to one university, the university schedule severely constrains the time schedule for the sub tasks. At first sight, it would seem that the solution is to look for more opportunities in other universities and implement realized digital learning materials in these other universities. In practice, the organization of such implementations and evaluations turns out to require considerable attention of the DOR team.

7.4.3.3 Fitting expert evaluation in the time frame of the projects

If expert evaluation is intended to be formative, it makes sense to realize the expert evaluation before implementation in the intended context of the material. If expert evaluation is intended to be summative, it can also be carried out after one or more case studies. Thus, for formative expert evaluation, time constraints will in general restrict the number of experts that can carry out the evaluation.

Improving Evaluation - 6 We now promote that evaluation activities are scheduled as soon as it has been decided in which exam session the achievement of learning goal and other objectives will be evaluated. This means that – starting with the date of this exam we should plan backwards in time to the beginning of the DOR project and explicitly plan all other evaluation sub tasks. This planning should take into account the issues discussed in this section. The plan should document explicitly who has taken responsibility for which evaluation action and who has taken responsibility for backup in case of incidents illness et cetera.

Note that the difference with the actual FBT projects is just the additional activity of early and explicit planning of evaluation sub tasks. The actual sub tasks themselves are not additional. Planning and scheduling evaluation activities will not disturb the balance between evaluation efforts and efforts in actual design and realization of the learning materials. Rather, the extra effort of planning and scheduling is likely to enable efficiency gains and prevent missed opportunities later on. We strongly believe that early planning of evaluation sub tasks will reduce the probability of missing data. In some of the FBT projects missing data did not allow us to draw clear conclusions or make certain claims.

At least the following evaluation activities need to be planned.

- Formative expert evaluation of realized designs.
- Formative empirical evaluation in case studies.
- Summative empirical evaluation in case studies:
  - by means of assessments aimed to investigate achievement of learning goals;
  - by means of questionnaires aimed to investigate the feasibility of achieving other design goals, e.g. student satisfaction and teacher satisfaction.

7.4.4 Using the right tools and methods

Ideally, realization should be a translation of an abstract representation of the design into software. This would imply conformance to accepted software engineering guidelines and use of adequate tools. An example of such a tool is a version control system. Such a system would also allow the evaluators to inspect the history of the resulting materials and identify decision points.

Currently however, we have not yet found an adequate formalism for abstract representations of the design of digital learning materials that will support expert evaluation. Thus, improving evaluation or adding evaluation activities along this line is not yet an option.

Further research - 3 Further research is needed on the design of a design language that enables experts of different disciplines to communicate efficiently with each other and to evaluate abstract representations of the design of digital learning material (see Box 2-2).

7.4.5 Final remarks on process evaluation in DOR

Within DOR, process evaluation mainly pays off in terms of quality or quantity of the outputs. Information delivered by process evaluation **for future design** projects would not primarily be design-oriented research. Such information might be valuable in various other design-related types of research. Such information should help to improve 'generic' process guidelines or to formulate new process guidelines. Improving process guidelines or adopting new process guidelines **within** a DOR project is not the intention of DOR and might be counter productive. One of the main difficulties of DOR is to stay focused on the design goal and not to side slip unconsciously into other research modes. Allocation of attention to improvements of 'generic' design guidelines during the DOR project should be limited to those improvements for which it can be guaranteed that they will pay of within the DOR project.

7.5 Concluding remarks

For each of the output-classes it makes sense to find out how we can arrive at a more interesting claim for output in that output-class. Of course, any such claim should also be supported by positive evaluation results.

The best opportunities for making interesting claims are in the goal definition. In Chapter 6 we already have formulated the requirement that the goal definition should enable one or more large-scale use scenarios. Note that the set of possible large-scale use scenarios is not necessarily complete at this moment. Rather it is likely that additional scenarios will be invented.

In Chapter 5 projects are distinguished with respect to their focus for innovation. On the one hand there are new learning goals as focus for innovation. Defining a really new learning goal and showing that this learning goal can be achieved in a large case study can be impressive. The primary requirement is that this learning goal is acknowledged as important for students in the specific program.

On the other hand, for well-established learning goals there are possible other foci of innovation. The 'centre of mass' of efforts in those projects could for instance be near innovative ways of constructive alignment, or innovative knowledge and information management.

In this chapter, we scanned the output-classes defined in Chapter 4 for opportunities to improve evaluation in future PhD DOR projects aimed at digital learning materials in higher education. Six opportunities were identified.

(1) Evaluation of 'blueprints' of assessments by SME's outside the project.

(2) Formative and summative evaluation of assessments (i.e. meta evaluation!).

(3) Evaluation of an abstract representation of the architecture.

(4) Scanning of interfaces for unnecessary prior knowledge assumptions.

(5) Comparison of intended allocation of efforts with actual allocation of efforts.

(6) Backward planning and scheduling evaluation activities as soon as the decision with respect to the exam date has been made.

In addition we formulated questions for further research.

(1) How to define assessments based on performance of students in a case in a virtual environment.

(2) How to evaluate the importance of a DOR goal in higher education.

(3) How to design a representational formalism for digital learning materials that enables experts from different disciplines to evaluate an abstract representation of the design.

(4) How to realize synergy between educational server applications and learning management systems.

(5) How to arrive at a sufficiently complete set of 'generic' assumptions about university students in Food and BioTechnology from anywhere in the world.

Finally, there is a grey area between really innovative learning goals and learning goals that are well articulated and well established. Experience in the FBT program suggests that for many learning goals in many subject matter fields, operationalization of 'understanding' still requires further research. In particular, "Measuring depth of understanding can pose challenges for objectivity. Much work needs to be done to minimize the trade-off between assessing depth and assessing objectively" [27]. Moreover, in most large-scale use scenarios it makes sense to require that scoring and marking of exams impose a minimal load on scarce SME capacity. In addition, experience with different types of interaction in the FBT program and in follow-up projects, suggests that results from research in knowledge and information management can still offer more opportunities to capture and operationalize what

SME's regard as 'understanding'. Thus, for this grey area, a valuable DOR goal is the delivery of computer-based assessments. Capturing 'understanding' will be the focus of innovation in these projects.

Summary

The context of the research described in this thesis is formed by a number of research projects that were aimed at the design, development, implementation, use and evaluation of innovative digital learning materials. Most of these projects were carried out mainly within Wageningen University. In this thesis, these projects are collectively referred to as 'WU Projects'. During this research it became clear that available literature provided insufficient support with respect to a number of issues. Examples are 'How to phrase research questions?', 'What output to expect?', 'What type of evaluation is relevant?' and 'What methods should be used?'. In fact, in parallel with the WU projects, the body of literature on methodology for design related research approaches in several disciplines was growing considerably. This thesis aims to contribute to this methodological discussion. In addition, this thesis presents a view on the characteristics and possibilities of digital learning materials in higher education.

In Chapters 2,3 and 4, a methodological framework for design, development, implementation, use and evaluation of innovative digital learning materials in higher education is defined and elaborated. Research that fits this framework is called design-oriented research (DOR). The framework is the result of a systems-oriented theoretical discussion of literature from a range of knowledge domains such as learning and instruction, knowledge and information systems research and engineering design. The concepts and terminology are illustrated with examples from publications that resulted from various WU projects. In addition, part of the framework is captured in a glossary of terms. The glossary aims to provide a coherent and ‘workable’ set of terms and corresponding definitions or descriptions. For many terms, this implies a compromise between natural language preferences of members of different disciplines.

In Chapter 4, a classification of outputs that are potentially valuable is presented. An important implication of the view presented in this chapter, is that the actual design

goal can be output, rather than input of a DOR project. This is in agreement with the view of instructional design and design of digital learning resources as processes of constraint exploration and constraint satisfaction. In Chapter 5 and Chapter 6 the most important strategic decisions in DOR projects that aim to deliver digital learning materials in higher education are discussed. In Chapter 5, a classification of design goals with their relation to various knowledge domains is presented. In Chapter 6, a classification of large-scale use scenarios with their relation to design requirements is given. Finally, Chapter 7 describes requirements and opportunities for evaluation in design-oriented research in education and reflects on evaluation in a number of WU projects. This leads to a number of suggestions for improvement with respect to evaluation in DOR.

Samenvatting

De context van het onderzoek dat beschreven is in dit proefschrift wordt gevormd door een aantal onderzoeksprojecten die gericht waren op het ontwerp, de realisatie, implementatie, gebruik en evaluatie van innovatieve digitale leermiddelen. De meeste van deze projecten zijn overwegend binnen Wageningen University uitgevoerd. In dit proefschrift wordt naar deze projecten verwezen middels de term 'WU Projects'. Gedurende deze onderzoeken werd het duidelijk dat de beschikbare literatuur onvoldoend ondersteuning bood voor dit soort onderzoek met betrekking tot vragen zoals: 'Wat is een geschikt format voor onderzoeksvragen?', 'Wat voor resultaten kan men verwachten?', 'Wat voor evaluatie is relevant?' en 'Welke methoden moeten gebruikt worden?'. In dezelfde periode waarin de WU projecten werden uitgevoerd groeide de hoeveelheid literatuur over methoden en technieken in ontwerpgerelateerde onderzoeksbenaderingen in diverse disciplines aanzienlijk. Dit proefschrift levert een bijdrage aan de methodologische discussies betreffende ontwerp gerelateerde onderzoeksbenaderingen. Bovendien presenteert dit proefschrift een beschouwing van de karakteristieken en mogelijkheden van digitale leermiddelen in universitair onderwijs.

In hoofdstuk 2,3 en 4, is een methodologisch kader voor ontwer, realisatie, implementatie, gebruik en evaluatie van innovatieve digitale leermiddelen gedefinieerd en nader uitgewerkt. Onderzoek dat past in dat kader wordt in dit proefschrift 'ontwerpgericht onderzoek' genoemd. Het kader is het resultaat van een systeem georienteerde theoretische discussie van literatuur uit een reeks van kennisdomeinen waaronder 'leren en instructie', 'kennis en informatiesystemen', en 'technologisch ontwerpen'. De begrippen en termen worden geïllustreerd met voorbeelden uit publicaties die voortkwamen uit diverse WU projecten. Bovendien wordt een deel van het kader vervat in een woordenlijst. Die woordenlijst is bedoeld

om een samenhangende en werkbare verzameling termen met bijpassende definities of beschrijvingen te bieden. Voor veel termen impliceert dat een compromis tussen natuurlijke taal en voorkeuren van onderzoekers in verschillende disciplines.
In Hoofdstuk 4 wordt beschreven tot wat voor soort potentieel waardevolle resultaten ontwerpgericht onderzoek kan leiden. Een belangrijke implicatie van het kader dat beschreven wordt is dat het eigenlijke ontwerpdoel eerder gezien moet worden als uitkomst dan als input van een ontwerpgericht onderzoeksproject. Dit past goed bij een visie waarin het ontwerpen van instructie en digitale leermiddelen gezien wordt als een process van 'constraint exploratie' en 'constraint satisfaction'. De Hoofdstukken 5en 6 behandelen de belangrijkste strategisch beslissingen welke genomen moeten worden in ontwerpgerichte onderzoeksprojecten die bedoeld zijn om digitale leermiddelen op te leveren. In Hoofdstuk 5, wordt een classificatie van ontwerpdoelen gepresenteerd en worden de verschillende soorten ontwerpdoelen gerelateerd aan diverse kennisdomeinen. Hoofdstuk 6 is gericht op ontwerpeisen die voortkomen uit een eventuele wens om te garanderen dat de betreffende leermiddelen op grote schaal gebruikt zullen worden. Deze ontwerpeisen zijn gekoppeld aan bekende scenarios voor grootschalig gebruik. Hoofdstuk 7beschrijft eisen en mogelijkheden voor evaluatie in ontwerpgericht onderzoek in universitair onderwijs en reflecteert op de evaluaties in een aantal WU projecten. Dit leidt tot een aantal suggesties voor verbetering met betrekking tot evaluatie in ontwerpgericht onderzoek.

References

1. Kuhn, T.S., *The Structure of Scientific Revolutions*. 2e edition 1973, Chicago: The University of Chicago.
2. Busstra, M.C., R. Hartog, S. Kersten, and M. Műller, *Design guidelines for the development of digital nutrigenomics learning material for heterogeneous target groups.* Advances in Physiology Education, 2007. **31**(March): p. 67 - 75.
3. Sessink, O.D.T., H. van der Schaaf, H.H. Beeftink, R.J.M. Hartog, and J. Tramper, *Web-based Education in Bioprocess Engineering.* Trends in Biotechnology, 2007. **25**(1): p. 16 - 23.
4. Sessink, O.D.T., R. Beeftink, J. Tramper, and R. Hartog. *Author-defined storage in the next generation learning management systems*. in *3rd IEEE International Conference on Advanced Learning Technologies (ICALT 2003)*. 2003. Athens Greece: IEEE.
5. Sessink, O.D.T., H.H. Beeftink, and R.J.M. Hartog, *Database Functionality for Learning Objects.* Technology, Instruction, Cognition and Learning, 2005. **2**(4).
6. Kassahun, A., A. Beulens, and R. Hartog, *Providing Author-Defined State Data Storage to Learning Objects.* Journal of Educational Technology and Society, 2006. **9**(2): p. 19 -32.
7. Hartog, R., H.v.d. Schaaf, and J. Verver, *eLearning.* Agro Informatica, 2003. **16**(4): p. 9-11.
8. Hartog, R., C.D.d. Gooijer, H.v.d. Schaaf, O. Sessink, and O.W. Vonder. *Comparing Web based Course Development With and Without a Learning Environment*. in *Proceedings of Webnet 2000: World conference on www and the internet*. 2000. San Antonio: Gordon Davies & Charles Owen.
9. Biggs, J. and C. Tang, *Teaching for quality learning at university*. 3e edition 2007, Maidenhead Berkshire England: Open university Press.
10. Collis, B. and J. Moonen, *Flexible Learning in a Digital World: Experiences and Expectations* 2001, London: Kogan Page Limited.
11. Anderson, J.R., *Learning and memory : An integrated approach* 1995, New York: Wiley.
12. Merriënboer, J.J.G.v., *Training Complex Cognitive Skills: A four-component instructional design model for technical training* 1997, Englewood Cliffs, NJ: Educational Technology Publications.
13. Nonaka, I. and H. Takeuchi, *The Knowledge-Creating Company* 1995, New York: Oxford University Press.
14. Merriënboer, J.J.G.v. and P.A. Kirschner, *Ten Steps to Complex Learning* 2007, Mahwah New Jersey: Lawrence Erlbaum Associates.

15. Larkin, J.H., C. Scheftic, and R.W. Chabay, *An Applied Science of Instructional Design* in *Towards a Scientific Practice of Science Education*, M. Gardner, et al., Editors. 1990, L. Erlbaum Associates: Hillsdale, NJ. p. 163 - 179.
16. Österle, H., J. Becker, U. Frank, T. Hess, D. Karagiannis, H. Krcmar, P. Loos, P. Mertens, A. Oberweis, and E.J. Sinz, *Memorandum on design-oriented information systems research.* European Journal of Information Systems, 2011. **20**(1): p. 7-10.
17. World Bank. *Capacity Development*. [last accessed april 18, 2012]. http://go.worldbank.org/O1RMGUIXX0.
18. Bell, P., C.M. Hoadley, and M.C. Linn, *Design-Based Research in Education* in *Internet Environments for Science Education*, M.C. Linn, E.A. Davis, and P. Bell, Editors. 2004, L. Erlbaum Associates: Mahwah NJ. p. 73-85.
19. Mulder, M., J. Gulikers, H. Biemans, and R. Wesselink, *The new competence concept in higher education: error or enrichment?* Journal of European Industrial Training, 2009. **33**(8/9): p. 755-770.
20. Dym, C.L. and P. Little, *Engineering Design: a project-based introduction* 2000: John Wiley & Sons.
21. Masterman, M., *The Nature of a Paradigm* in *Criticism and the Growth of Knowledge*, I. Lakatos and A. Musgrave, Editors. 1970, Cambridge University Press: Cambridge. p. 59 - 89.
22. Bronshtein and Semendyayev, *A Guide Book to Mathematics* 1973, Zürich: Harry Deutsch Verlag.
23. Smith, P.L. and T.J. Ragan, *Instructional Design* 1993, USA: Macmillan Publishing Company.
24. Hartog, R. and S. Osinga, *Keller: een database met screen-recordings* in *SPEC Symposium*, C.J.d. Groot and J.J.M. Smolenaars, Editors. 1996: Wageningen. p. 25-32.
25. Polanyi, M., *Personal Knowledge*. 3[e] edition 1969, London: Routledge & Kegan Paul.
26. W3C. *Web Services Description Language (WSDL) Version 2.0 Part 1: Core Language*. [last accessed 10 augustus 2012]. http://www.w3.org/TR/wsdl20/.
27. Bransford, J.D., A.L. Brown, and R.R. Cocking, *How People Learn: Brain, Mind, Experience, and School. Expanded Edition.*, ed. C.o.B.a.S.S.a.E. Committee on Developments in the Science of Learning and Committee on Learning Research and Educational Practice, National Research Council.2003, Washington DC: National Academy Press.
28. Briggs, L.J., ed. *Instructional Design: Principles and Applications*. 1977, Educational Technology Publications: Englewood Cliffs NJ. 532.
29. Dick, W. and L. Cary, *The Systematic Design of Instruction.* . 3[e] edition 1990, Glenview Illinois: Scott Foresman and Company.
30. Yelon, S.L., *Powerful Principles of Instruction* 1996, White Plains, NY: Longman Publishers USA.
31. Romiszowski, A.J., *Designing Instructional Systems* 1984, New York: Nichols Publishing Company.
32. West, C.K., J.A. Farmer, and P.M. Wolff, *Instructional Design: Implications from Cognitive Science* 1991, Needham Heights, MA: Allyn and Bacon.
33. Gagné, R.M., L.J. Briggs, and W.W. Wager, *Principles of Instructional Design*. 4[e] edition 1992, New York.
34. Reigeluth, C.M., *What is Instructional-Design Theory and how is it changing?* in *Instructional Design Theories and Models - A new paradigm of instructional theory*, C.M. Reigeluth, Editor 1999, Routeledge: Hillsdale, NJ. p. 5 - 30.

35. Aegerter-Wilmsen, T. and T. Bisseling, *Biology by Numbers-Introducing Quantitation into Life Science Education.* PLoS Biology 2005. **3**(1): p. e1.
36. Busstra, M.C., E.J.M. Fesken, R.J.M. Hartog, P. van 't Veer, and F.J. Kok, *Interactive Digital Learning Material on Collating Evidence from Human Nutrition Research.* e-SPEN the European e-Journal of Clinical Nutrition and Metabolism, 2008. **3**: p. e52-e61.
37. Busstra, M.C., A. Geelen, E.J. Feskens, R.J.M. Hartog, and P. van ’t Veer, *Design and Development of Digital Learning Material for Applied Data Analysis.* The American Statistician, 2008. **62**(4): p. 329 - 338.
38. Shulman, L., *Knowledge and teaching: Foundations of the new reform.* Harvard Educational Review, 1987. **57**: p. 1 - 22.
39. Driel, J.H.v., W.R. Veal, and F.J.J.M. Janssen, *Pedagogical content knowledge: an integrative component within the knowledge base for teaching.* Teaching and Teacher Education, 2001. **17**(8): p. 979-986.
40. Abell, S.K., *Research on Science Teacher Knowledge -Chapter 36* in *Handbook of Research on Science Education*, S.K. Abell and N.G. Lederman, Editors. 2007, Routledge. p. 1122 e.p.
41. Anderson, J.R., *Cognitive Psychology and Its Implications*. 5^{e} edition 2000, USA: Worth Publishers.
42. Blackboard. *Blackboard*. [last accessed december 17 2010]. http://www.blackboard.com/.
43. Moodle. *Moodle*. [last accessed 17 dec 2010]. http://moodle.com.
44. Sakai_Foundation. *Sakai*. [last accessed 7 augustus 2012]. http://www.sakaiproject.org/.
45. Maurer, H. and N. Scherbakov. *WBTmaster*. [last accessed march 04 2006]. http://coronet.iicm.edu/.
46. Hasebrook, J.P. and H.A. Maurer, *Learning support systems for organizational learning*. Series on innovative intelligence. Vol. 8. 2004, Singapore: World Scientific Publishing Company.
47. Kolk, K.v.d., G. Beldman, R.J.M. Hartog, and H. Gruppen, *Towards an Integration of Printed and Digital Learning Materials for Food Chemistry Education into a Full E-learning Course* in *Proceedings of World Conference on Educational Multimedia, Hypermedia and Telecommunications*, J. Luca and E.R. Weippl, Editors. 2008, AACE Chesapeake VA: Vienna. p. 6073 - 6078.
48. Hartog, R., *Computer assisted learning: from process control paradigm to information resource paradigm.* Journal of Microcomputer Applications, 1989. **12**(Spring): p. 15-31.
49. Sessink, O.D.T., H.H. Beeftink, J. Tramper, and R.J.M. Hartog, *Proteus: A Lecturer-Friendly Adaptive Tutoring System.* Journal of Interactive Learning Research, 2007. **18**(4): p. 533-554.
50. Hartog, R., *Qualitative Simulation and Knowledge Representation for Intelligent Tutoring* in *Computer Assisted Learning. Proceedings of the 2nd International Conference, ICCAL '89, Dallas, TX, USA, May 1989*, H. Maurer, Editor 1989, Springer Verlag: Berlin. p. 193-213.
51. Aegerter-Wilmsen, T., F. Janssen, D. Kettenis, O. Sessink, R.J.M. Hartog, and T. Bisseling, *Introducing molecular life science students to model building using computer simulations.* Journal of Computers in Mathematics and Science Teaching, 2006. **25**(2): p. 101-122.
52. Diederen, J., H. Gruppen, R. Hartog, and A.G.J. Voragen, *Design and Evaluation of digital learning material to support acquisition of quantitative*

problem solving skills within food chemistry education. Journal of Science Education and Technology, 2005. **14**(5-6): p. 495-507.
53. Gill, T.G., *Quick and Dirty Multimedia.* Decision Sciences Journal of Innovative Education, 2007. **5**(1): p. 197 - 206.
54. Hartog, R., S. Osinga, A.J.M. Beulens, and W. de Bruin, *Computerondersteund onderwijs op basis van audiovisuele screen-recordings.* Tinfon, 1997. **6**(2): p. 59-63.
55. Hartog, R., S. Draaijer, and L. Rietveld, *Practical Aspects of Task Allocation in Design and Development of Digital Closed Questions in Higher Education.* Practical Assessment, Research and Evaluation, 2008. **13**(2): p. 1 - 15.
56. Wager, W., *Instructional Technology and Higher Education* in *Instructional Design*, L.J. Briggs, Editor 1977, Educational Technology Publications: Englewood Cliffs New Jersey. p. 395 - 419.
57. Hartog, R.J.M., *The FBT programme Looking back and looking forward* 2009, WMMRC Wageningen University.
58. Gardner, M., J.G. Greeno, F. Reif, A.H. Schoenfeld, A. diSessa, and E. Stage, *Toward a scientific practice of science education* 1990, Hillsdale, N.J.: Lawrence Erlbaum Associates.
59. Anderson, T. *Utilizing Disruptive Technologies in the University: Confessions of an Agent Provocateur*. in *World Conference on Educational Multimedia, Hypermedia and Telecommunications 1999*. 1999. AACE.
60. Busstra, M.C. (2008). *Design and Evaluation of Digital Learning Material for Academic Education in Human Nutrition.* Wageningen University, Wageningen.
61. Sessink, O.D.T. (2006). *Bioprocess-Engineering Education with Web Technology.* Wageningen University, Wageningen.
62. Aegerter-Wilmsen, T. (2005). *Digital Learning Material for Experimental Design and Model Building in Molecular Biology.* Wageningen University, Wageningen.
63. Schaaf, H.v.d. (2007). *Design of Digital Learning Material for Bioprocess-Engineering Education.* Wageningen University, Wageningen.
64. Diederen, J. (2005). *Design and Evaluation of Digital Activating Learning Materials for Food Chemistry Education.* Wageningen University, Wageningen.
65. Kolk, K.v.d., G. Beldman, R. Hartog, and H. Gruppen, *Students Using a Novel Web-Based Laboratory Class Support System: A Case Study in Food Chemistry Education.* Journal of Chemical Education, 2011. **89**(1): p. 103-108.
66. Hartog, R.J.M., ed. *Design and development of digital closed questions: a methodology for midsized projects in higher education*. Vol. 13. 2008, SURF: Utrecht. 167.
67. Barab, S. and K. Squire, *Design-Based Research: Putting a Stake in the Ground.* Journal of the Learning Sciences, 2004. **13**(1): p. 1 - 14.
68. Brown, A.L., *Design Experiments: Theoretical and Methodological Challenges in Creating Complex Interventions in Classroom Settings.* Journal of the Learning Sciences, 1992. **2**(2): p. 141 - 178.
69. Collins, A., *Towards a design science of education* in *New directions in educational technology*, E. Scanlon and T. O'Shea, Editors. 1992, Springer: Berlin. p. 15-22.
70. Collins, A., D. Joseph, and K. Bielaczyc, *Design Research: Theoretical and Methodological Issues.* Journal of the Learning Sciences, 2004. **13**(1): p. 15 - 42.

71. Design-Based Research Collective, *Design-Based Research: An Emerging Paradigm for Educational Inquiry.* Educational Researcher, 2003. **32**(1): p. 5-8.
72. Gravemeijer, K. and P. Cobb, *Design research from a learning design perspective* in *Educational Design Research*, J.v.d. Akker, et al., Editors. 2006, Routledge: Abingdon. p. 17 - 51.
73. Hevner, A. and S. Chatterjee, *Design science research in information systems.* Design Research in Information Systems, 2010. p. 9-22.
74. Hevner, A., *The Three Cycle View of Design Science Research.* Scandinavian Journal of Information Systems, 2007. **19**(2): p. 87-92.
75. Hevner, A.R., S.T. March, J. Park, and S. Ram, *Design Science in Information Systems Research.* MIS Quarterly, 2004. **28**(1): p. 75 - 105.
76. Kelly, A.E., *Quality Criteria for design research* in *Educational Design Research* J.v.d. Akker, et al., Editors. 2006, Routledge: Abingdon. p. 107 - 117.
77. March, S.T. and V.C. Storey, *Design science in the information systems discipline: an introduction to the special issue on design science research.* MIS Quarterly, 2008. **32**(4): p. 725-730.
78. March, S.T. and G.F. Smith, *Design and natural science research on information technology.* Decision Support Systems, 1995. **15**(4): p. 251 - 266.
79. Reeves, T.C., J. Herrington, and R. Oliver, *Design Research: A Socially Responsible Approach to Instructional Technology Research in Higher Education.* Journal of Computing in Higher Education, 2005. **16**(Spring).
80. Richey, R.C. and J.D. Klein, *Design and development research: methods, strategies, and issues* 2007, Mahwah: Lawrence Erlbaum Assoc Inc.
81. van den Akker, J., K. Gravemeijer, S. McKenney, and N. Nieveen, eds. *Educational Design Research* 2006, Routledge & Kegan Paul: London/New York.
82. van den Akker, J., R.M. Branch, K. Gustafson, N. Nieveen, and T. Plomp, *Design Approaches and Tools in Education and Training* 1999, Dordrecht/Boston/London: Kluwer.
83. Verschuren, P. and R. Hartog, *Evaluation in Design Oriented Research.* Quality and Quantity, 2005. **39**(6): p. 733-762.
84. Winters, N. and Y. Mor, *IDR: A participatory methodology for interdisciplinary design in technology enhanced learning.* Computers & Education, 2008. **50**(2): p. 579-600.
85. Moonen, J., *The Design and Prototyping of Digital Learning Material: Some New Perspectives* in *Design Approaches and Tools in Education and Training*, J.v.d. Akker, et al., Editors. 1999, Kluwer: Dordrecht/Boston/London. p. 95-111.
86. Simon, H.A., *The sciences of the artificial*. 3e edition 1996, Cambridge, Mass.: MIT Press.
87. Asimow, M., *Introduction To Design*. Prentice-Hall Series in Engineering Design, ed. R.J. B.1962, Englewood Cliffs, NJ: Prentice-Hall.
88. Alexander, C., S. Ishikawa, M. Silverstein, M. Jacobson, I. Fiksdahl-King, and S. Angel, *A pattern language : Towns, buildings, construction* 1977, New York: Oxford University Press.
89. Gamma, E., R. Helm, R. Johnson, and J. Vlissides, *Design Patterns: Elements of Reusable Object-Oriented Software* 1994, Amsterdam: Addison and Wesley Professional Computing Series.

90. Hartog, R.J.M., H. van der Schaaf, A.J.M. Beulens, and J. Tramper, *Virtual Experiments in University Education* in *Looking Toward the Future of Technology Enhanced Education: Ubiquitous Learning and the Digital Native*, M. Ebner and M. Schiefner, Editors. 2009, IGI Global: Hershey New York. p. 373 - 393.
91. Kirschner, P.A. and W. Huisman, *Dry laboratories in science education.* International Journal of Science Education, 1998. **20**(6): p. 655-682.
92. Sessink, O.D.T., R. Beeftink, J. Tramper, and R. Hartog, *Virtual Parameter-Estimation Experiments in Bioprocess-Engineering Education.* Bioprocess and Biosystems Engineering, 2006. **28**: p. 379 - 386.
93. Cartier, J., *Chemical reactors and fermentors: Simulation and Modeling*, 2007, Ecole Polytechnique Federale de Lausanne: Lausanne. p. 117.
94. Jong, T.d. and W.R.v. Joolingen, *Scientific Discovery Learning With Computer Simulations of Conceptual Domains.* Review of Educational Research, 1998. **68**(2): p. 179-201.
95. Schaaf, H.v.d., M. Vermuë, J. Tramper, and R. Hartog, *Support of Modelling in Process-Engineering Education.* Computer Applications in Engineering Education, 2006. **14**(3): p. 161 - 168.
96. Schaaf, H.v.d., J. Tramper, R. Hartog, and M. Vermuë, *A digital tool set for systematic model design in Process-Engineering Education.* European Journal of Engineering Education, 2006. **31**(5): p. 619 -629.
97. Yelon, S.L., *Active Appropriate Practice* in *Powerful Principles of Instruction*1996, Longman Publishers USA: White Plains, NY. p. 187-210.
98. Biggs, J.B., *Teaching for quality learning at university* 1999, Buckingham, UK: Open University Press.
99. Aegerter-Wilmsen, T., T. Bisseling, and R. Hartog, *Web based Learning Support for Experimental Design in Molecular Biology: A Top-Down Approach.* Journal of Interactive Learning Research, 2003. **14**(3): p. 301-314.
100. Diederen, J., H. Gruppen, R. Hartog, and A.G.J. Voragen, *Design and Evaluation of Digital Assignments on Research Experiments within Food Chemistry.* Journal of Science Education and Technology, 2006. **15**(3,4): p. 227 - 246.
101. Goodyear, P., *Emerging Methodological Challenges* in *Methodological Choice and Design*, L. Markauskaite, et al., Editors. 2011, Springer Netherlands. p. 253-266.
102. Brown, D.C. and B. Chandrasekaran, *Design Problem Solving: Knowledge Structures and Control Strategies* 1989, London: Pitman Publishing.
103. Dasgupta, S., *Design Theory and Computer Science* 1991, Cambridge UK: Cambridge University Press.
104. Alexander, C., *Notes on the Synthesis of Form* 1964, Cambridge: Harvard University Press.
105. Vincenti, W.G., *What Engineers Know and How They Know It* 1990, London: the Johns Hopkins Press.
106. Baggetun, R., E. Rusman, and C. Poggi, *Design Patterns For Collaborative Learning: From Practice To Theory And Back* in *Proceedings of World Conference on Educational Multimedia, Hypermedia and Telecommunications*, L. Cantoni and C. McLoughlin, Editors. 2004, AACE Chesapeake VA: Lugano. p. 2493 - 2498.
107. Valcke, M. and R. Martens, *The problem arena of researching computer supported collaborative learning: Introduction to the special section.* Computers & Education, 2006. **46**(1): p. 1-5.

108. Duval, E., W. Hodgins, D. Rehak, and R. Robson, *Learning objects symposium Special Issue Editorial* Journal of Educational Multimedia and Hypermedia 2004. **13**(4): p. 1-195.
109. Kirschner, P.A., *Cognitive Load theory: implications of cognitive load theory on the design of learning.* Learning and Instruction, 2002. **12**: p. 1-10.
110. Bannert, M., *Managing cognitive load--recent trends in cognitive load theory.* Learning and Instruction, 2002. **12**(1): p. 139-146.
111. Gaines, B.R. and M.L.G. Shaw, *Concept maps as hypermedia components.* International Journal of Human-Computer Studies, 1995. **43**: p. 323-361.
112. Jonassen, D.H., *Computers as cognitive tools: Learning with technology, not from technology.* Journal of Computing in Higher Education, 1995. **6**(2): p. 40-73.
113. Derntl, M. and R. Motschnig-Pitrik, *coUML: A Visual Language for Modeling Cooperative Environments* in *Handbook of Visual Languages for Instructional Design*, L. Botturi and S.T. Stubs, Editors. 2008, Information Science Reference: Hershey USA. p. 154 - 184.
114. Botturi, L. and S.T. Stubs, *Handbook of Visual Languages for Instructional Design* 2008, Hershey USA: Information Science Reference.
115. Niegemann, H.M., *A Multimedia Storyboard Tool* in *Proceedings Ed-Media Conference 2002* Montreal *(C*anada*)*, M.P. Driscoll and T. Reeves, Editors. 2002, AACE: Charlottesville. p. 2692-2693.
116. Wolverton, J.A. and G. Waddoups, *Developing a Storyboarding Process for Online Content: From Microsoft PowerPoint to Macromedia Flash* in *Proceedings EdMedia Conference 2000 Montreal (Canada)*, J. Bourdeau and R. Heller, Editors. 2000, AACE: Charlottesville. p. 1188-1193.
117. OMG. *Unified Modeling Language*. [last accessed 7 aug 2012]. http://www.uml.org/.
118. WG12, I.L. *Learning Object Metadata Standard*. [last accessed jul 31 2007]. http://www.ieeeltsc.org/standards/1484-12-1-2002/
119. IMS Global Learning Consortium. *IMS Learning Resource Meta-data Specification Version 1.3 - Final Specification.* . [last accessed Dec 17 2010]. http://www.imsglobal.org/metadata/.
120. Anderson, L.W. and D.R. Krathwohl, *A Taxonomy for Learning, Teaching and Assessing: A Revision of Bloom's Taxonomy of Educational Objectives* 2001, New York: Longman.
121. Koper, R. and C. Tattersall, *Learning design: A handbook on modelling and delivering networked education and training* 2005, Berlin-Heidelberg: Springer
122. IMS Global Learning Consortium. *Question and Test Interoperability*. [last accessed Aug 7 2012]. http://www.imsglobal.org/question/.
123. Hauser, J.R. and D. Clausing, *The House of Quality.* Harvard Buisiness Review, 1988. **1988**(May-June): p. 63-73.
124. Simon, H.A., *Cooperation Between Educational Technology and Learning Theory to Advance Higher Education* in *Technology Enhanced Learning: opportunities for change*, P.S. Goodman, Editor 2001, Erlbaum. p. 61-74.
125. Sweller, J., J. van Merrienboer, and F. Paas, *Cognitive Architecture and Instructional Design.* Educational Psychology Review, 1998. **10**(3): p. 251-296.
126. Chandrasekaran, B., *Design Problem Solving: A Task Analysis.*, in *AI-Magazine*1990. p. 59-71.

127. Jonassen, D., *Instructional design as design problem solving: An iterative process.* Educational Technology, 2008. **XLVII**(3): p. 21-26.
128. Dimitrova, M. and A. Suthcliffe. *Designing Instructional Multimedia Applications: Key Practices and Design Patterns*. in *EDMEDIA '99*. 1999. Seatle: AACE.
129. Derntl, M., *The Person-Centered e-Learning Pattern Repository: Design for Reuse and Extensibility* in *Proceedings of Edmedia 2004*2004, AACE: Lugano. p. 3856 -3861.
130. Mislevy, R.J., L. Hamel, R. Fried, T. Gaffney, G. Haertel, A. Hafter, R. Murphy, E. Quellmatz, A. Rosenquist, P. Schank, K. Draney, C. Kennedy, K. Long, M. Wilson, N. Chudowski, A.L. Morrison, P. Pena, N. Butler Songer, and A. Wenk, *Design Patterns for Assessing Science Inquiry*, in *PADI Principled Assessment Designs for Inquiry*, A.M. Colker, et al., Editors. 2003, Center for Technology in Learning (SRI International): Menlo Park. p. 97.
131. Draaijer, S. and R.J.M. Hartog, *Design Patterns for Digital Item Types in Higher Education.* e-Journal of Instructional Science and Technology (e-JIST), 2007. **10**(October).
132. Hartog, R.J.M. and S. Osinga, *Sculptor: an environment for constraint network specification* in *Optimization-based computer aided modelling and design*, H.-J. Sebastian, Editor 1994, Lansa: Rijswijk. p. 197-211.
133. Linn, M.C., E.A. Davis, and P. Bell, *Internet Environments for Science Education* 2004: Lawrence Erlbaum Associates.
134. Aegerter-Wilmsen, T., F. Janssen, R. Hartog, and T. Bisseling, *Digital Learning Material for Model Building in Molecular Biology.* Journal of Science Education and Technology, 2005. **14**(1): p. 123-134.
135. Diederen, J., H. Gruppen, R. Hartog, and A.G.J. Voragen, *Evaluation of computer-based learning material for food chemistry education.* Chemistry Education Research and Practice, 2005. **6**(2): p. 64-82.
136. Simon, H.A., *Learning to Research about Learning* in *Cognition and Instruction: twenty-five years of progress*, S.M. Carver, Editor 2001, L. Erlbaum Associates: Mahwah, NJ. p. 205 - 226.
137. Busstra, M.C., C.d. Graaf, and R.J.M. Hartog, *Design of Digital Learning Material on Social – Psychological Theories for Nutrition Behavior Research.* Journal of Educational Multimedia and Hypermedia, 2007. **16**(2).
138. Busstra, M.C., R. Hartog, and P. van 't Veer, *The role of active manipulation of three-dimensional scatter plots in understanding the concept of confounding.* Epidemiologic Perspectives & Innovations, 2005. **2**(6).
139. Wilmsen, T., T. Bisseling, and R. Hartog, *Web based Learning Support for Experimental Design in Molecular Biology* in *Proceedings of ED-Media 2002*, P. Barker and S. Rebelsky, Editors. 2002: Denver. p. 2063-2068.
140. Ebert, C., R. Dunke, M. Bundschuh, and A. Schmietendorf, *Best practices in software measurement* 2005, Berlin: Springer.
141. Diederen, J., H. Gruppen, R. Hartog, G. Moerland, and A.G.J. Voragen, *Design of activating digital learning material for food chemistry education.* Chemistry Education: Research and Practice, 2003. **4**(3): p. 353-371.
142. Herbsleb, J., D. Zubrow, D. Goldenson, W. Hayes, and M. Paulk, *Software quality and the Capability Maturity Model.* Communications of the ACM, 1997. **40**(6): p. 30-40.
143. Boehm, B. and V.R. Basily, *Software defect reduction top 10 list.* IEEE Computer, 2001. (Januari).

144. Iding, M.K., *How analogies foster learning from science texts.* Instructional Science, 1997. **25**(4): p. 233 - 253.
145. Norman, D.A., *The Design of Everyday Things* 1989, New York: Doubleday.
146. Schaaf, H.v.d., M. Vermue, J. Tramper, and R. Hartog, *A design environment for downstream processes for Bioprocess-Engineering students.* European Journal of Engineering Education, 2003. **28**(4): p. 507-521.
147. Schilling, M.A., *Toward a General Modular Systems Theory and Its Application to Interfirm Product Modularity.* The Academy of Management Review, 2000. **25**(2): p. 312-334.
148. Hölttä-Otto, K. and O. De Weck, *Degree of Modularity in Engineering Systems and Products with Technical and Business Constraints.* Concurrent engineering - results and applications, 2007. **15**(2): p. 113 - 126.
149. Russell, J.D., *Modular Instruction: A Guide to the Design, Selection, Utilization, and Evaluation of Modular Materials* 1974, Minneapolis: Burgess Publishing Co.
150. Keren, G. and W.B. de Bruin, *On the Assessment of Decision Quality: Considerations Regarding Utility, Conflict and Accountability*. Thinking: Psychological Perspectives on Reasoning, Judgment and Decision Making2005: John Wiley & Sons, Ltd.
151. Heinrich, E. and H. Maurer, *Active Documents: Concept, Implementation and Applications.* Journal of Universal Computer Science, 2000. **6**(12): p. 1197-1202.
152. Dasgupta, S., *Technology and creativity* 1996, New York: Oxford University Press.
153. Dym, C.L., *Engineering Design: a synthesis of views* 1994, USA: Cambridge University Press.
154. Sommerville, I., *Software Engineering* 8th edition. International Computer Science Series2007: Pearson Education.
155. Stapleton, J., *DSDM, dynamic systems development method: the method in practice* 1997, Harlow, England ; Reading, Mass.: Addison-Wesley.
156. Crinnion, S.A., *Evolutionary Systems Development* 1994, London: Pitman Publishing.
157. Popper, K.R., *The Logic of Scientific Discovery* 1974, London: Hutchinson & Co LTD.
158. Campbell, D.T. and J.C. Stanley, *Experimental and Quasi-Experimental Designs for Research* 1966, Boston.
159. Hartog, R.J.M., A.J.M. Beulens, and J. Tramper, *Faculty-based projects on digital learning materials in higher education and design-related research methodologies*, 2009, WMMRC Wageningen University: Wageningen. p. 32.
160. Linn, M.C., E.A. Davis, and P. Bell, *Introduction* in *Internet environments for science education*, M.C. Linn, E.A. Davis, and P. Bell, Editors. 2004, L. Erlbaum Associates: Mahwah NJ. p. xv - no successiv e page numbers.
161. Davis, E.A., *Creating Critique Projects* in *Internet Environments for Science Education*, M.C. Linn, E.A. Davis, and P. Bell, Editors. 2004, L. Erlbaum Associates: Mahwah NJ. p. 89 - 113.
162. KIE-Research-Group. *Knowledge Integration Environment*. [last accessed 6 mei 2008]. http://kie.berkeley.edu/KIE.html.
163. Davis, E.A., *Prompting Middle School Science Students for Productive Reflection: Generic and Directed Prompts* Journal of The Learning Sciences, 2003. **12**(1): p. 91 - 142.

164. Glaser, R., *Progress Then and Now* in *Cognition and Instruction Twenty-Five Years of Progress*, S.M. Carver and D. Klahr, Editors. 2001, Lawrence Erlbaum Associates: Mahwah, NJ. p. 493 -507.
165. Hartog, R.J.M. and A.J.M. Beulens, *Digital Learning Resources, Design-Oriented Research and Basic Research* in *Proceedings of World Conference on Educational Multimedia, Hypermedia and Telecommunications 2012*, T. Amiel and B. Wilson, Editors. 2012, AACE: Chesapeake, VA. p. 2941-2950.
166. Bush, V., *Science: The Endless Frontier.* Transactions of the Kansas Academy of Science (1903-), 1945. **48**(3): p. 231-264.
167. Stokes, D.E., *Pasteur's Quadrant: basic science and technical innovation* 1997, Washington DC: Brookings Institution Press.
168. Gibbons, M., *Science's new social contract with society.* Nature, 1999. **402 (6761)**(December): p. C81-C84.
169. Ziman, J., *Real Science: what it is and what it means* 2002, Cambridge: Cambridge University Press.
170. Calvert, J., *What's Special about Basic Research?* Science, Technology & Human Values, 2006. **31**(2): p. 199.
171. Patton, M.Q., *Qualitative Research and Evaluation Methods*. 3e edition 2002, Thousand Oaks London: Sage Pub.
172. OECD, *Frascati Manual: Proposed standard practice for surveys on research and experimental development* 2002, OECD Publication Services: Paris.
173. Lidwell, W., K. Holden, and J. Butler, *Universal Principles of Design* 2003, Gloucester MA: Rockport Publishers.
174. Newell, A., *Unified Theories of Cognition* 1987, Cambridge, MA: Harvard University Press.
175. Posner, G.J. and A.N. Rudnitsky, *Course Design: a guide to curriculum development for teachers* 1997, White Plains, NY: Longman Publishers USA.
176. Halpin, T., *Conceptual Schema & Relational Database Design*. 2nd edition 1995, Sydney: Prentice-Hall.
177. Chen, P.P.-S., *The Entity-Relationship Model-Toward a Unified View of Data.* ACM TODS, 1976. **1**(1).
178. Fowler, M. and K. Scott, *UML Distilled: a brief guide to the standard object modeling language*. 2nd edition 2000, Berkeley, California: Addison and Wesley.
179. Schlesinger, S., R.E. Crosby, R.E. Gagne, G.S. Innis, C.S. Lalwani, J. Loch, R.J. Sylvester, R.D. Wright, N. Keir, and D. Bartos, *Terminology for Model Credibility.* Simulation, 1979. **32**(March): p. 103-104.
180. Fagan, M., *Advances in software inspections.* IEEE Transactions on Software Engineering, 1986. **12**(7): p. 741- 755.
181. Doyle, J.K. and D.N. Ford, *Mental models concepts for system dynamics research.* System dynamics review, 1998. **14**(1): p. 3-29.
182. Simon, H.A., *Models of bounded rationality* 1982, Cambridge, Mass.: MIT Press.
183. Einstein, A., *Relativity* 1954: Methuen & Co Ltd.
184. Burkhardt, H., *From Design Research to Large Scale Impact* in *Educational Design Research*, J.v.d. Akker, et al., Editors. 2006, Routledge: London/New York. p. 121-150.
185. Anderson, T.R., *Bridging the educational research-teaching practice gap.* Biochemistry and Molecular Biology Education, 2007. **35**(6): p. 465-477.

186. Kester, L., P.A. Kirschner, J.J.G.v. Merriënboer, and A. Baumer, *Just-in-time information presentation and the aquisition of complex cognitive skills.* Computers in Human Behavior, 2001. **17**(4).
187. Schreiber, A.T., J.M. Akkermans, A.A. Anjewierden, R.d. Hoog, W.v.d. Velde, and B.J. Wielinga, *Engineering of Knowledge: The CommonKADS Methodology version 0.3* 1997, Amsterdam: Department of Social Science Informatics.
188. Holzinger, A., *Usability engineering methods for software developers.* Commun. ACM, 2005. **48**(1): p. 71-74.
189. Cristea, A. and F. Garzotto, *Designing Patterns for Adaptive or Adaptable Educational Hypermedia: a Taxonomy* in *Proceedings of World Conference on Educational Multimedia, Hypermedia and Telecommunications*, L. Cantoni and C. McLoughlin, Editors. 2004, AACE Chesapeake VA: Lugano. p. 808-813.
190. Aegerter-Wilmsen, T., M. Coppens, F. Janssen, R. Hartog, and T. Bisseling, *Digital learning material for student-directed model building in molecular biology.* Biochemistry and Molecular Biology Education, 2005. **33**(5): p. 325-329.
191. Mettes, C.T.C.W., A. Pilot, H.J. Roossink, and H. Kramers-Pals, *Teaching and Learning Problem Solving in Science Part I: A general strategy.* Journal of Chemical Education, 1980. **57**(12): p. 882-885.
192. Glinz, M., *Rethinking the Notion of Non-Functional requirements* in *Proceedings of the Third World Congress for Software Quality*2005: Munich. p. 55 - 64.
193. Gilb, T., *Towards the engineering of requirements.* Requirements Engineering, 1997. **2**(3): p. 165-169.
194. Kibler, R.J. and R.E. Bassett, *Writing performance objectives* in *Instructional Design*, L.J. Briggs, Editor 1977, Educational Technology Publications: New Jersey. p. 49 - 95.
195. Steen, J., *Guideline for writing and using intended learning outcomes for degree programmes at Wageningen University.*, 2009, Wageningen University: Wageningen.
196. Laurillard, D., *Rethinking University Teaching A Conversational Framework for the Effective Use of Learning Technologies*. 2[e] edition 2002, London: Routledge Falmer
197. Hestenes, D., M. Wells, and G. Swackhamer, *Force concept inventory.* The physics teacher, 1992. **30**: p. 141 - 158.
198. Klymkowsky, M.W. and K. Garvin-Doxas, *Recognizing Student Misconceptions through Ed's Tools and the Biology Concept Inventory.* PLoS Biol, 2008. **6**(1): p. e3.
199. Garvin-Doxas, K. and M.W. Klymkowsky, *Understanding Randomness and its Impact on Student Learning: Lessons Learned from Building the Biology Concept Inventory (BCI).* CBE Life Sci Educ, 2008. **7**(2): p. 227-233.
200. Klymkowsky, M.W., K. Garvin-Doxas, and M. Zeilik, *Bioliteracy and Teaching Efficacy: What Biologists Can Learn from Physicists.* Cell Biol Educ, 2003. **2**(3): p. 155-161.
201. Wilson, C.D., C.W. Anderson, M. Heidemann, J.E. Merrill, B.W. Merritt, G. Richmond, D.F. Sibley, and J.M. Parker, *Assessing Students' Ability to Trace Matter in Dynamic Systems in Cell Biology.* CBE Life Sci Educ, 2006. **5**(4): p. 323-331.
202. ADL. *SCORM*. [last accessed May 15 2012]. http://www.adlnet.gov/capabilities/scorm.

203. Katz, H., *Designing a SCORM 2004 Compliant and ISD Grounded Reusable Learning Object Model (RLOM)* in *Proceedings of World Conference on Educational Multimedia, Hypermedia and Telecommunications*, L. Cantoni and C. McLoughlin, Editors. 2004, AACE Chesapeake VA. p. 287-292.
204. Ding, L., R. Chabay, B. Sherwood, and R. Beichner, *Evaluating an electricity and magnetism assessment tool: Brief electricity and magnetism assessment.* Physical review special topics- physics education research, 2006. **2**(1): p. 1-7.
205. Gronlund, N.E., *Measurement & Evaluation in Teaching*. 3rd edition 1965, New York: Macmillan Publishing Company.
206. Dennis, A., B.H. Wixom, and D. Tegarden, *Systems Analysis and Design with UML Version 2.0*. 2nd edition 2005: John Wiley and Sons.
207. Borgida, A., R.J. Brachman, M.J. Coombs, and J.F. Sowa, *Principles of semantic networks : explorations in the representation of knowledge* 1991, San Mateo, CA: Kaufman.
208. Tramper, J., *Immobilizing biocatalysts for use in syntheses.* Trends in Biotechnology, 1985. **3**(2): p. 45-50.
209. Wilson, S., *3 Architectures to Support Authoring and Content Management with Learning Design* in *Learning Design*, R. Koper and C. Tattersall, Editors. 2005, Springer: Berlin-Heidelberg. p. 408.
210. Brusilovsky, P., *KnowledgeTree: a distributed architecture for adaptive e-learning.* Proceedings of the 13th international World Wide Web conference on Alternate track papers & posters, 2004. p. 104-113.
211. Gütl, C., M. Pivec, C. Trummer, V.M. García-Barrios, F. Mödritscher, J. Pripfl, and M. Umgeher, *AdeLE (Adaptive e-Learning with Eye-Tracking): Theoretical Background, System Architecture and Application Scenarios.* EURODL, 2005. p. 2005-12.
212. Schmidt, A. and C. Winterhalter, *User Context Aware Delivery of E-Learning Material: Approach and Architecture.* Journal of Universal Computer Science (JUCS), 2004. **10**: p. 28–36.
213. Baghaei, N., A. Mitrovic, and W. Irwin, *Problem-Solving Support in a Constraintbased Tutor for UML Class Diagrams.* Technology, Instruction, Cognition and Learning, 2006. **4**(2): p. 4 - 28.
214. Jara, C.A., F.A. Candelas, F. Torres, S. Dormido, F. Esquembre, and O. Reinoso, *Real-time collaboration of virtual laboratories through the Internet.* Computers & Education, 2009. **52**(1): p. 126-140.
215. Clements, P., R. Kazman, and M. Klein, *Evaluating software architectures*. The SEI series in software engineering2002, Indianapolis IN: Pearson Education.
216. Collins, A., J.S. Brown, and S.E. Newman, *Cognitive Apprenticeship: Teaching the Craft of Reading, Writing, and Mathematics* in *Knowing, learning and instruction: essays in honor of Robert Glaser*, L.B. Resnick, Editor 1989, Lawrence Erlbaum: Hillsdale, NJ. p. 453-494.
217. White, B.Y. and J.R. Frederiksen, *Causal Model Progression as a Foundation for Intelligent Learning Environments.* Artificial Intelligence, 1990. **42**: p. 99-157.
218. Boyle, T., *Towards a Theoretical Base for educational multimedia design.* Special issue of JIME Journal of Interactive Media in Education - on Theory of Learning Technologies, 2002. **2002**(2).
219. Collis, B. and A. Strijker, *New Pedagogies and Re-Usable Learning Objects: Toward a Different Role for an LMS* in *Proceedings of World Conference on Educational Multimedia, Hypermedia and Telecommunications 2002*, P. Barker and S. Rebelsky, Editors. 2002, AACE: Chesapeake, VA. p. 334-339.

220. Duval, E., *A Learning Object Manifesto - Towards Share and Reuse on a Global Scale*, in *Presentation at E-learning Conference,* 2005, ProLearn: Brussels, Belgium,. p. 113 - 117.
221. ESCOT. *Educational Software Components for Tomorrow*. [last accessed 10 augustus 2012]. http://www.merlot.org/merlot/viewMaterial.htm?id=284173.
222. Roschelle, J., C. Digiano, M. Koutlis, A. Repenning, A. Jackiw, and D. Suthers, *Developing Educational Software Components.* IEEE Computer, 1999. **32**(9): p. 50-58.
223. IEEE. *IEEE home page*. [last accessed 2008]. http://www.ieee.org/.
224. ADL. *Advanced Distributed Learning home page*. [last accessed 2008]. http://www.adlnet.gov/.
225. IMS Global Learning Consortium. *About the IMS Global Learning Consortium*. [last accessed 10 Augustus 2012]. http://www.imsglobal.org/aboutims.html.
226. Gijbels, D., F. Dochy, P. Van den Bossche, and M. Segers, *Effects of Problem-Based Learning: A Meta-Analysis From the Angle of Assessment.* Review of Educational Research, 2005. **75**(1): p. 27-61.
227. Moust, J.H.C. and H.G. Schmidt, *Probleemgestuurd Onderwijs* 1998, Groningen: Wolters-Noordhof.
228. Schmidt, H.G., *Problem-based learning: Rationale and description.* Medical Education, 1983. **17**: p. 11-16.
229. Smith-Gratto, K., D. Wicks, and C. Berger. *MERLOT: Reaping the On-line Vineyard*. in *World Conference on Educational Multimedia, Hypermedia and Telecommunications 2002*. 2002. Denver, Colorado, USA: AACE.
230. Harrigan, K., T. Carey, and W. Ld Team. *Community Building in a Co-operative Learning Object Repository*. in *World Conference on Educational Multimedia, Hypermedia and Telecommunications 2005*. 2005. Montreal, Canada: AACE.
231. Cafolla, R., *Project MERLOT: Bringing Peer Review to Web-Based Educational Resources.* Journal of Technology and Teacher Education, 2006. **14**(2): p. 313-323.
232. Matzas, M., P.F. Stahler, N. Kefer, N. Siebelt, V. Boisguerin, J.T. Leonard, A. Keller, C.F. Stahler, P. Haberle, B. Gharizadeh, F. Babrzadeh, and G.M. Church, *High-fidelity gene synthesis by retrieval of sequence-verified DNA identified using high-throughput pyrosequencing.* Nature Biotechnology, 2010. **28**(12): p. 1291-1294.
233. Wagner, D. and D. Schmalstieg, *Making augmented reality practical on mobile phones, part 1.* Computer Graphics and Applications, IEEE, 2009. **29**(3): p. 12-15.
234. Bode, S., A. Fischer, W. Kunhauser, and M. Riebisch. *Software Architectural Design Meets Security Engineering*. in *Engineering of Computer Based Systems, 2009. ECBS 2009. 16th Annual IEEE International Conference and Workshop on the*. 2009.
235. Lee, J., M. Natarajan, V.C. Nashine, M. Socolich, T. Vo, W.P. Russ, S.J. Benkovic, and R. Ranganathan, *Surface Sites for Engineering Allosteric Control in Proteins.* Science, 2008. **322**(5900): p. 438-442.
236. Eckert, C. and J. Clarkson, *The reality of design* in *Design process improvement - a review of current practice*, J. Clarkson and C. Eckert, Editors. 2005, Springer: London. p. 1 - 29.
237. Baker, J., J. Thierstein, K. Fletcher, M. Kaur, and J. Emmons, *Open Textbook Proof-of-Concept via Connexions.* The International Review of Research in Open and Distance Learning, 2009. **10**(october): p. 1-13.

238. Yin, R.K., *Case study research: design and methods*. 2nd edition. Applied social research methods series; vol 51994, Thousand Oaks: Sage Publications.
239. Goodyear, P., *Educational design and networked learning: Patterns, pattern languages and design practice.* Australasian Journal of Educational Technology, 2005. **21**(1): p. 82-101.
240. Underwood, J.S., C. Hoadley, H.S. Lee, K. Hollebrands, C. DiGiano, and K.A. Renninger, *IDEA: Identifying design principles in educational applets.* Educational Technology Research and Development, 2005. **53**(2): p. 99 - 112.
241. Yin, R.K., *Case study research: design and methods*. 3nd edition 2003, Thousand Oaks (CA): Sage.
242. Richey, R.C., J.D. Klein, and W.A. Nelson, *Developmental Research: Studies of instructional design and development* in *Handbook of research on educational communications and technology*, D.H. Jonassen, Editor 2004, L. Erlbaum Associates: Mahwah NJ. p. 1099 - 1130.
243. Hartog, R., A. Beulens, and J. Tramper, *Faculty-Based Design-Oriented Research on Digital Learning Materials: Defining Project Goals* in *Proceedings of Global Learn Asia Pacific 2010*, Z.W. Abas, I. Jung, and J. Luca, Editors. 2010, AACE: Penang, Malaysia. p. 1418-1427.
244. Anderson, T.R. and K.J. Schönborn, *Bridging the educational research-teaching practice gap.* Biochemistry and Molecular Biology Education, 2008. **36**(4): p. 309-315.
245. Rowhani, S. and K. Sedig, *E-books plus: Role of interactive visuals in exploration of mathematical information and e-learning.* Journal of Computers in Mathematics and Science Teaching, 2005. **24**(3): p. 273-298.
246. Huber, T., W. Nagler, and M. Ebner, *The ABC-eBook System From Content Management Application to Mash-up Landscape* in *Proceedings of edmedia 2008*, J. Luca and E.R. Weippl, Editors. 2008, AACE: Vienna. p. 6015 - 6022.
247. Schönborn, K.J. and T.R. Anderson, *The importance of visual literacy in the education of biochemists.* Biochemistry and Molecular Biology Education, 2006. **34**(2): p. 94-102.
248. Intelligen. *SuperPro Designer*. [last accessed october 13 2007]. http://www.intelligen.com/superpro_overview.shtml.
249. Walker, D., *Fundamentals of Curriculum* 1990, New York.
250. Perkins, D. and G. Salomon, *Are cognitive skills context bound?* Educational Researcher, 1989. **18**(jan-feb): p. 16-25.
251. SPSS. *SPSS*. [last accessed 2008]. http://www.spss.com/.
252. B-Basic. *B-Basic*. [last accessed Oct 28 2007]. www.b-basic.nl.
253. Boxel, M.v. and H. Verstralen, *Multiple Response Questions in Computerized Testing* in *Design and development of digital closed questions: a methodology for midsized projects in higher education*, R.J.M. Hartog, Editor 2007, SURF Foundation: Utrecht. p. 35 - 51.
254. Sessink, O.D.T., R. Beeftink, J. Tramper, and R. Hartog, *Securing web-based exams.* Journal of Universal Computer Science, 2004. **10**(2): p. 145 - 157.
255. Phillips, D.C., *Assessing the quality of design research proposals* in *Educational Design Research*, J.v.d. Akker, et al., Editors. 2006, Routledge & Kegan Paul: London/New York. p. 93 - 99.
256. Edelson, D.C., *Balancing Innovation and Risk* in *Educational Design Research*, J.v.d. Akker, et al., Editors. 2006, Routledge & Kegan Paul: London/New York. p. 100 - 106.
257. Hartog, R.J.M., H. van der Schaaf, and A. Kassahun, *Towards Synergy between Learning Management Systems and Educational Server Applications* in

Proceedings of World Conference on Educational Multimedia, Hypermedia and Telecommunications, J. Luca and E.R. Weippl, Editors. 2008, AACE Chesapeake VA: Vienna. p. 2253 - 2258.

258. Maurer, H., *Important Aspect of Knowledge Management* in *Computer Science in Perspective*, R. Klein, H.W. Six, and L. Wegner., Editors. 2003, Springer: Berlin Heidelberg. p. 245 - 254.
259. Maurer, H., *Presentation type cal for classroom and lab use at university level* in *Proceedings LIFE Conference Calgary*1987: Calgary. p. 27-29.
260. Maurer, H. and F. Makedon, *COSTOC: Computer Supported Teaching Of Computer Science* in *Proceedings of the IFIP conference on Teleteaching 1986 - Budapest*, F. Lovis, Editor 1986, North Holland Publishing Company: Amsterdam. p. 107 - 119.
261. NuGO. *Nutrigenomics Organization*. [last accessed Dec 17 2010]. www.nugo.org.
262. SURF. *SURF Innovatie in het onderwijs : ALTB (Actief leren - Transparant beoordelen)*. [last accessed 2011]. http://www.surf.nl/nl/projecten/Pages/ALTB-Actief-Leren-Transparant-Beoordelen.aspx.
263. Hartog, R.J.M. and J. Tramper, *DiMoBio: 'Digital Modules for Heterogeneous Groups in university Biotechnology Programmes*, 2006, B-Basic organization.
264. Ariadne-Foundation. *Ariadne*. [last accessed April 9 2009]. http://www.ariadne-eu.org/.
265. NSDL. *The BEN Portal*. [last accessed May 15 2012]. http://www.biosciednet.org/portal/.
266. DLESE. *Digital Library for Earth System Education*. [last accessed Jun 23 2010]. http://dlese.org/library/index.jsp.
267. MERLOT. *Multimedia Educational Resource for Learning and Online Teaching*. [last accessed november 23 2009]. http://www.merlot.org/merlot/index.htm.
268. InternetArchive. *Open Educational Resources OER*. [last accessed Mei 15 2012]. http://www.archive.org/details/education.
269. SREB. *SCORE Sharable Content Object Repositories for Education*. [last accessed May 15 2012]. http://sreb-score.org/.
270. SMETE-Open-Federation. *SMETE Digital Library* [last accessed Jan 31 2011]. http://www.smete.org.
271. WikimediaFoundation. *Wikimedia-Commons*. [last accessed May 15 2012]. http://commons.wikimedia.org/.
272. WikimediaFoundation. *Wikiversity*. [last accessed May 15 2012]. http://en.wikiversity.org/.
273. Wisc-Online. *Wisconsin Online Resource Center*. [last accessed May 15 2012]. http://www.wisc-online.com/.
274. MIT. *MIT Open Courseware*. [last accessed May 15 2012]. http://ocw.mit.edu/OcwWeb/web/courses/courses/index.htm.
275. Carnegie Mellon. *OpenLearningInitiative*. [last accessed May 15 2012]. https://oli.web.cmu.edu/openlearning/.
276. The Open University. *LearningSpace*. [last accessed May 15 2012]. http://openlearn.open.ac.uk/.
277. University of Colorado. *PhET*. [last accessed May 15 2012]. http://phet.colorado.edu/.
278. Wieman, C.E. and K.K. Perkins, *A powerful tool for teaching science.* Nature physics, 2006. **2**(may).

279. Wieman, C.E., W.K. Adams, and K.K. Perkins, *PhET: Simulations That Enhance Learning.* Science, 2008. **322**: p. 682 - 683.
280. Academy, K. *Khan Academy*. [last accessed May 15 2012]. http://www.khanacademy.org/.
281. Arneberg, P., L. Guardia, D. Keegan, J. Lõsenko, I. Mázár, P. Fernández Michels, M. Flate Paulsen, T. Rekkedal, A. Sangrà, J. Atle Toska, and D. Zarka, *Analyses of European megaproviders of e-learning* 2007, Bekkestua, Norway: NKI Publishing House.
282. Mayadas, A.F., J. Bourne, and P. Bacsich, *Online Education Today.* Science, 2009. **323**(5910): p. 85-89.
283. Consortium, I.G.L. *IMS GLC Common Cartridge Working Group*. [last accessed May 15 2012]. http://www.imsglobal.org/commoncartridge.html.
284. Pearson. *Mastering Chemistry*. [last accessed May 15 2012]. http://www.masteringchemistry.com/.
285. Wenger, E., *Artificial Intelligence and Tutoring Systems: Computational and Cognitive approaches to the communication of knowledge.* 1987, San Francisco: Morgan Kaufmann Publishers.
286. IMS Global Learning Consortium. *IMS Content Packaging v1.1.4 final specification* [last accessed May 15 2012]. http://www.imsglobal.org/content/packaging/index.html.
287. IMS Global Learning Consortium. *Learning Tools Interoperability v 1.1 final specification*. [last accessed May 15 2012]. http://www.imsglobal.org/lti/.
288. Regli, W.C., X. Hu, M. Atwood, and W. Sun, *A Survey of Design Rationale Systems: Approaches, Representation, Capture and Retrieval.* Engineering with Computers, 2000. **16**(3): p. 209-235.
289. Haladyna, T.M., *Writing Test Items to Evaluate Higher Order Thinking* 1997: Needham Heights MA, Allyn and Bacon.
290. Haladyna, T.M., *Developing and Validating Multiple-Choice Test Items*. Third edition 2004, Mahwah, N.J.: Lawrence Erlbaum Associates.
291. Downing, S.M. and T.M. Haladyna, *Handbook of test development.* 2006, Mahwah, NJ: Lawrence Erlbaum Associates.
292. Clark, C. and R. Mayer, *E-learning and the science of instruction*. 2nd edition 2008, San Francisco: Pfeiffer.
293. Dagger, D., A. O'Connor, S. Lawless, E. Walsh, and V.P. Wade, *Service-oriented e-learning platforms: From monolithic systems to flexible services.* Internet Computing, IEEE, 2007. **11**(3): p. 28-35.
294. Simon, H.A., *The bottleneck of attention: connecting thought with Motivation.* in *Integrative views of motivation, cognition and emotion.*, W.D. Spaulding, Editor 1994, University of Nebraska Press: Lincoln. p. 1-21.
295. Kahneman, D., *Attention and Effort* 1973, Englewood Cliffs New Jersey: Prentice Hall.
296. Csikszentmihalyi, M. and J. Nakamura, *The Dynamics of Intrinsic Motivation: A Study of Adolescents* in *Motivation and Cognition : Goals and Cognitions*, C. Ames and R. Ames, Editors. 1989, Academic Press.
297. Ames, R. and C. Ames, *Introduction* in *Motivation in Education vol 3 Goals and Cognitions*1989, Academic Press. p. 1 - 10.
298. Keller, J.M., *Development and Use of the ARCS Model of Motivational Design* 1983, Enschede: Technische Hogeschool Twente.
299. Baddeley, A., *Working memory: looking back and looking forward.* Nature reviews. Neuroscience, 2003. **4**(10): p. 829.

300. Bhattacharjee, Y., *Can Randomized Trials Answer the Question of What Works?* Science, 2005. **307**(5717): p. 1861-1863.

Acknowledgement

In the first place I want to thank Kees Oomen for his work on the lay-out but also for being a valuable colleague and supporter of my work. The latter holds as well for Gerard Moerland and Jan Ockeloen. Kees, Gerard and Jan, thanks for your contributions to Wageningen MultiMedia Research Centre and for being reliable colleagues. Over the years Julia Diederen, Tinri Aegerter, Cora Busstra, Olivier Sessink, Hylke van der Schaaf and Koos van der Kolk have stimulated me to think about what we have called 'design-oriented research' and to elaborate a design-oriented research approach. Julia, Tinri, Cora, Olivier, Hylke and Koos, thanks for all your input to the main topic of this thesis. Olivier, Hylke and Koos, you have also contributed to most of the technological ideas and their technical implementation underlying this work. In this, you also have proven to be very reliable. Thanks again! To Ayalew Kassahun I want to express my thanks for his lessons to me on information system development. Ayalew, I am sure I learned more from you than you are aware of. To Sjoukje Osinga and Harry Lövenstein I want to express my thanks for their valuable assistance in the final stage of this work.

This thesis is itself the result of a design process. In terms of chapters two and three, this process has involved many requirements and constraints, both process constraints as well as design constraints. The contribution to those constraints by my supervisors, Adrie Beulens and Hans Tramper, is here gratefully acknowledged.

As this thesis is one of the many results of more than fifteen years of work on digital learning resources for higher education, many more people have contributed indirectly to this thesis. I thank them all and I sincerely hope I have not forgotten to give them due credit in other publications.

Curriculum Vitae

Robertus Johannes Maria Hartog was born on september 15, 1948 in Haarlem. He completed his secondary education (gymnasium β) in 1966 at the Triniteitslyceum in Haarlem. In 1974, he received his master degree in experimental physics from the University of Amsterdam. Next, he taught physics in several positions in secondary education in the Netherlands and in Surinam until 1979. From 1979 until 1986, he taught physics and didactics of physics at the teacher-training centre (NLO) in Nijmegen.

Since 1986, he is employed by Wageningen University. Besides lecturing information systems development, his primary task is to lead projects on the use of information and communication technology in higher education. In particular, he has managed a large educational innovation program, called the Food and BioTechology program. He was co-supervisor in five PhD projects that were completed within this program. As project leader he is also responsible for the Wageningen Multimedia Research Centre. This is a centre within the Social Sciences Group of Wageningen University focusing on the design, development and implementation of digital learning materials.