A MODEL FOR USING LEARNERS' ONLINE BEHAVIOUR TO INFORM DIFFERENTIATED INSTRUCTIONAL DESIGN IN MOODLE

R.G. LEPPAN

2019

A model for using learners' online behaviour to inform differentiated instructional design in Moodle

by

Ronald George Leppan

Thesis

submitted in fulfilment of the requirements for the degree

Philosophiae Doctor

in

Information Technology

in the

Faculty of Engineering, the Built Environment and Information Technology

of the

Nelson Mandela University

Promoter: Prof. JF van Niekerk
Co-promoter: Prof. RA Botha

April 2019

DECLARATION

I, Ronald George Leppan (s196475940), hereby declare that the thesis for Philosphiae Doctor in Information Technology is my own work and that it has not previously been submitted for assessment or completion of any postgraduate qualification to another University or for another qualification.

Ronald George Leppan

ACKNOWLEDGEMENTS

I would like to express my gratitude to:

My promoters, Professor Johan van Niekerk and Professor Reinhardt Botha for their probing questions that challenged me to clarify my thoughts.

The Nelson Mandela University Department of Research Capacity Development for their financial support and for the research development workshops, especially the SANTRUST Doctoral Proposal Development Programme.

My colleagues who helped lessen my workload so that I can focus on getting over that final hurdle.

My family and friends for their interest in my progress, for their patience while I became increasingly absent and for their support when I needed it the most.

My muse, Claudette, thank you for being a willing sounding board for my often-incoherent ideas, for the tough love during times when I needed a push and for the gentle touch when I needed a lift.

ABSTRACT

This thesis proposes a learning analytics-based process model, derived from a web analytics process, which aims to build a learner profile of attributes from Moodle log files that can be used for differentiated instructional design in Moodle. Commercial websites are rife with examples of personalisation based on web analytics, while the personalisation of online learning has not yet gained such widespread adoption.

Several Instructional Design Models recommend that, in addition to taking prior knowledge and learning outcomes into account, instruction should also be informed by learner attributes. Learning design choices should be made based on unique learner attributes that influence their learning processes. Learner attributes are generally derived from well-known learning styles and associated learning style questionnaires. However, there are some criticisms of learning style theories and the use of questionnaires to create a learner profile. Attributes that can be inferred from learners' online behaviour could provide a more dynamic learner profile.

Education institutions are increasingly using Learning Management Systems, such as Moodle, to deliver and manage online learning. Moodle is not designed to create a learner profile or provide differentiated instruction. However, the abundant data generated by learners accessing course material presented in Moodle provides an opportunity for educators to build such a dynamic learner profile. Individual learner profiles can be used by educators who desire to tailor instruction to the needs of their learners.

The proposed model was developed and evaluated using an iterative design focused approach that incorporates characteristics of a web analytics process, instructional design models, Learning Management Systems, educational data mining and adaptive education technologies. At each iteration, the model was evaluated using a technical risk and efficacy strategy. This strategy proposes a formative evaluation in an artificial setting. Evaluation criteria used include relevance, consistency, practicality and utility.

The contributions of this thesis address the lack of prescriptive guidance on how to analyse learner online behaviours in order to differentiate learning design in Moodle. The theoretical contribution is a model for a dynamic data-driven approach to profile building and a phased differentiated learning design in a Learning Management

System. The practical contribution is an evaluation of the expected practicality and utility of learner modelling from Moodle log files and the provision of tailored instruction using standard Moodle tools.

The proposed model recommends that educators should define goals, develop Key Performance Indicators (KPI) to measure goal attainment, collect and analyse suitable metrics towards KPIs, test optional alternative hypotheses and implement actionable insights.

To enable differentiated instruction, two phases are necessary: learner modelling and differentiated learning design. Both phases rely on the selection of suitable attributes which influence learning processes, and which can be dynamically inferred from online behaviours. In differentiated learning design, the selection/creation and sequencing of Learning Objects are influenced by the learner attributes. In learner modelling, the data sources and data analysis techniques should enable the discovery of the learner attributes that was catered for in the learning design.

Educators who follow the steps described in the proposed model will be capable of building a learner profile from Moodle log files that can be used for differentiated instruction based on any learning style theory.

TABLE OF CONTENTS

LIST OF APPENDICES

LIST OF FIGURES

LIST OF TABLES

LIST OF ABBREVIATIONS AND ACRONYMS

ABCD	:	Audience, Behaviour, Conditions, Degree
ADDIE	:	Analyse, Design, Develop, Implement, Evaluate
AES	:	Adaptive education Systems
AH	:	Adaptive Hypermedia
ARFF	:	Attribute Relations File Format
ASSIST	:	Approaches and Study Skills Inventory for Students
ASSURE	:	Analyse, State, Select, Utilise, Require, Evaluation
ATTLS	:	Attitudes Towards Thinking and Learning Survey
BKT	:	Bayesian Knowledge Estimation
COLLES	:	Constructivist Online Learning Environment Survey
DAG	:	Directed Acyclic Graphs
DBR	:	Design Based Research
DBSCAN	:	Density-Based Spatial Clustering of Applications with Noise
DeRTEL	:	Design Research for Technology Enhanced Learning
DSR	:	Design Science Research
EDM	:	Educational Data Mining
EM	:	Expectation Maximisation
FN	:	False Negatives
FP	:	False Positive
FPR	:	False Positive Rates
GMM	:	Gaussian Mixture Models
ICT	:	Information and Communication Technology
ID	:	Instructional Design
IDE	:	Integrated Development Environment
IEEE	:	Institute of Electrical and Electronics Engineers
IS	:	Information Systems

ISDT	:	Information Systems Design Theory
ITS	:	Intelligent Tutoring Systems
KPI	:	Key Performance Indicators
LA	:	Learning Analytics
LD	:	Learning Design
LKE	:	Latent Knowledge Estimation
LMS	:	Learning Management System
LO	:	Learning Object
LOM	:	Learning Object Metadata
LSAES	:	Learning Style Based Adaptive Educational Systems
LTI	:	Learning Tools Interoperability
MARC	:	Machine Readable Cataloguing
NIPALS	:	Non-Linear Iterative Partial Least Squares
OLM	:	Open Learner Model
PFA	:	Performance Factors Analysis
PHP	:	PHP: Hypertext Pre-processor
RDF	:	Resource Description Framework
RIO	:	Reusable Information Object
RLO	:	Reusable Learning Object
ROC	:	Receiver Operating Characteristic
SCORM	:	Sharable Content Object Reference Model
SGML	:	Standard Generalised Markup Language
SMART	:	Specific, Measurable, Attainable, Relevant, Timely
SNA	:	Social Network Analysis
SPADE	:	Sequential Pattern Discovery Using Equivalence Classes
TEL	:	Teaching Enhanced Learning
TN	:	True Negative

TP	:	True Positive
TPR	:	True Positive Rates
UI	:	User Interface
URL	:	Uniform Resource Locator
WEKA	:	Waikato Environment for Knowledge Analysis
WHERETO	:	• **W**here the Unit Is Going • **H**ooks to grab learners' attention • **E**xplore issues • **R**eflect • **E**valuate • **T**ailor • **O**rganise for engagement
xAPI	:	Experience Application Programming Interface
XML	:	Extensible Mark-Up Language

Chapter 1. Introduction

Chapter 1
Introduction

1.1) Research Context

1.2) Relevance of Research Problem

1.3) Rationale for and Significance of Proposed Solution

1.4) Research Aim, Question and Objectives

- 1.4.1) Primary Research Question
- 1.4.2) Research Objectives

1.5) Thesis Scope

1.6) Thesis Structure

- 1.6.1) Chapter 1: Introduction
- 1.6.2) Chapter 2: Research Methodology
- 1.6.3) Chapter 3: Instructional Technology and Learning Design
- 1.6.4) Chapter 4: Educational Data Mining
- 1.6.5) Chapter 5: Iterative Development and Evaluation of Proposed Solution
- 1.6.6) Chapter 6: Discussion and Conclusions

1.1 Research Context

The rapid growth of emerging technologies for teaching and learning is a double-edged sword. There is no question that educators are blessed with an abundance of instructional technologies that can be harnessed to provide engaging online learning experiences. However, the vast and continuously evolving choice and complexity of technologies makes the process of instructional design exponentially harder (Sinclair and Aho, 2017). Making appropriate learning design choices becomes so overwhelming that educators often use available technologies inadequately, haphazardly or not at all (Minielli and Ferris, 2005).

One example of the ineffectual use of technology is the phenomenon of teachers using the Moodle Learning Management System as a mere repository for lecture notes, despite its large collection of pedagogically sound features to choose from (Sinclair and Aho, 2016). Many South African institutions adopted a blended learning approach to complete the academic year when the #FeesMustFall protests halted classroom instruction in 2015 and 2016. This is resulting in a growing number of lecturers moving their courses online. The lecturers at institutions that adopted Moodle may need prescriptive guidance to move beyond simply uploading slideshows and documents.

Compounding the problem of instructional design (ID) is the diverse nature of our classrooms (Hall and Keynes, 2011). Each learner has unique characteristics, abilities and emotions that influence his or her approach to learning. South African higher education institutions must cope with rising learner numbers from a diverse background, against a backdrop of limited resources (Gravett and Geyser, 2009). Given the diverse background of our learners, we should try to create a learning environment tailored towards their unique characteristics, which according to Verdú and Regueras (2008) has the potential to improve the effectiveness of learning and the satisfaction of learners. However, with the high number of learners in a classroom, differentiating instruction towards this diversity is a daunting task. With the advances in Web 2.0 technologies, differentiated instruction through e-learning may offer some solutions (Lwoga, 2012). In particular, there is a growing interest in using Learning Analytics to extract meaningful information from large online datasets (Merceron, Blikstein and Siemens, 2015). Learning Analytics is defined by Siemens (2013) as the measurement, collection, analysis and reporting of data about learners and their contexts, in order to optimise the learning process and the learning environment.

1.2 Relevance of Research Problem

Contemporary ID theories offer recommendations on the most effective way of presenting instructional content (Mehlenbacher, 2012). Instructional design models are used by instructional designers to create technology enhanced learning experiences systematically.

There are two main types, procedural and conceptual instructional design models. Procedural models provide a simplified view of a complex task and prescribe sequential steps in the creation of technology enhanced instructional interventions. Conceptual instructional design models provide an abstract view of reality in the form of taxonomies, heuristics or conceptual frameworks. They do not impose any sequence but provide factors to be considered when creating online learning opportunities.

Both classes of instructional design models are useful for their ability to create meaningful instructional experiences. Both classes of instructional design models frequently highlight the importance of developing instruction based on known learner attributes. Following these instructional design models and pedagogical theories, instructional designers can develop an initial online learning opportunity while catering for various learner attributes. However, it is not known whether the learners working through the online material actually experience the learning design as intended by the instructional designer.

Developers of adaptive e-learning systems say that we should tailor online instruction based on a dynamic learner profile that is constructed from their online learning behaviours (Brusilovsky, 2012; Mulwa, Lawless, Sharp, Arnedillo-Sanchez and Wade, 2010; Seyal and Rahman, 2015; Van Seters, Ossevoort, Tramper and Goedhart, 2012). While ID models describe the need for catering towards different learner characteristics, they are lacking in prescriptive guidance on how to infer relevant learner characteristics from learners' dynamic online behaviours.

Instructional Designers can create a proprietary learning system or use existing Learning Management Systems (LMS) to facilitate instruction. An LMS provides a platform through which learning content can be delivered. Moodle is an example of an LMS that is widely adopted (Moodle, 2018). It provides Learning Resources and Learning Activities in its core functionality. Moodle is also open source, which means there is a community of developers producing plugins that Moodle users can incorporate into their learning material.

At its core, though, Moodle does not consider unique learner characteristics (Graf, 2009). Characteristics, such as cognitive style and motivation, influence the way learners engage with the instructional materials. So, while the initial learning design may have been developed to satisfy various learner characteristics, are they suitable for the actual learners currently engaging with the instructional materials? When improvements are necessary to the learning design, we need to dynamically, through learning analytics, determine the learner characteristics to motivate our learning design choices (Rienties, Nguyen, Holmes and Reedy, 2017). With the learning design frequently presented through a Learning Management System such as Moodle, it gives rise to the problem statement:

There is limited prescriptive guidance on how to create a meaningful learner profile from Moodle logs that can inform differentiated learning design choices in Moodle, leading to inadequate instructional designs.

1.3 Rationale for and Significance of Proposed Solution

Learning objects are digital educational resources used in online instruction. In differentiated instruction, the instructional offering is typically tailored on three levels: content, presentation and sequencing of Learning Objects. Adaptive e-learning provides significant advantages, since it personalises the learning experience with little extra effort from the teacher. A dynamically adaptive education system based on learner traits frequently builds a student profile with psychological tests (Brusilovsky, 1996). We can overcome inherent flaws in these psychological tests by using a method to analyse learner online behaviours to infer learner characteristics (Popescu, Badica and Moraret, 2010). A number of studies have developed proprietary systems that adapt to learner profiles based on learner characteristics (Carver, Howard and Lane, 1999; Papanikolaou, Grigoriadou, Kornilakis and Magoulas, 2003; Popescu et al., 2010; Wolf, 2003).

However, as more education institutions are moving towards a blended learning strategy (Lu, 2012), they are moving away from developing proprietary systems and adopting existing Learning Management Systems. Examples of widely used Learning

Management Systems include Blackboard, WebCT and Moodle. Learning Management Systems provide tools that teachers can use to create, administer and manage online learning (Falvo and Johnson, 2007). Despite all the tools provided in Learning Management Systems, they lack in the adaptivity department (Despotović-Zrakić, Marković, Bogdanović, Barać and Krčo, 2012). In particular, Moodle does not provide native support for learner modelling that can keep track of learner characteristics. Moodle does, however, keep track of what learners do in the system. Lecturers are provided with log files containing data such as a timestamp of access to Learning Objects. From these timestamps, one can calculate how many times a learner viewed a particular resource and how long he or she spent on a certain activity. This data can be analysed using a variety of techniques to build a profile of relevant learner characteristics. Knowing these characteristics can help the teacher make changes to the online course based on how the learners prefer to access the resources and activities. The challenge is to build an accurate learner profile and use this profile to inform learning design choices.

Four learning design related terms need disambiguation: Personalised learning, individualised learning, adaptive learning and differentiated learning. All four terms refer to tailoring instruction in some way. The underlying premise behind providing tailored instruction is a belief that a strategy that works well for one student, may frustrate another. Tailored instruction is necessary for all learners to keep them engaged and suitably challenged (Manning, Stanford and Reeves, 2010).

Differentiation: Differentiated instruction occurs when pre-set pathways towards the same objectives are created for learners based on their learning needs, goals and characteristics (Tomlinson et al., 2003). Learners are grouped together based on shared traits or needs and the curriculum tailored towards the group's characteristics. According to Nicolae (2014), differentiated instruction aims to:

- Improve learning outcomes
- Inspire lifelong learning
- Increase self-awareness
- Improve learning efficiency
- Improve learner satisfaction, motivation and engagement

Adaptation: Adaptive learning incorporates the principles of differentiated instruction, notably the tailoring of content and individualised learning paths. Adaptive learning systems are data-driven and continually update the learner profile in real time. The system adapts learning pathways and content based on what works for learners with similar traits.

Individualisation: Individualised learning allows learners to achieve individual objectives at their own pace. In individualised learning, learners may review outcomes not yet mastered. In some cases, learners may also set their own learning agenda within defined parameters.

Personalisation: Personalised learning tailors the content, pathways and pace to learners' unique needs. Adaptive algorithms are used to lay out the individual's learning path and content or to recommend Learning Objects. Students might take an initial diagnostic test that will be fed into a rules-engine. Thereafter the profile also gets continuously updated as in adaptive learning systems. Personalised learning systems assume that each learner is completely unique and provide a higher level of personalisation than differentiated instruction, individualised instruction and adaptive learning. In personalised learning, learners take control of their learning and their learning environment.

This thesis is aimed at exploring a system that enables differentiated instruction, particularly in the context of a Learning Management System.

1.4 Research Aim, Question and Objectives

The investigation into the problem identified in Section 1.2 is guided by the primary research question given in Section 1.4.1 and operationalised by the research objectives given in Section 1.4.2.

1.4.1 Primary Research Question

What are the steps of a comprehensive, learner-centric process model to enable differentiated instruction in Moodle based on a dynamic learner profile?

1.4.2 Research Objectives

The following primary research objective operationalises the primary research question:

1) Develop and evaluate a comprehensive, learner-centric process model to enable differentiated instruction based on a dynamic learner profile

The following sub-objectives instantiate the process model in a Learning Management System:

a) Instantiate the learning design phase in a Learning Management System
b) Instantiate the modelling phase in a Learning Management System

1.5 Thesis Scope

This thesis focuses on enabling differentiated instruction through identifying and using suitable data collection and analysis tools and techniques. The thesis does not report on the impact that the differentiation will have on actual learning. Giving an impact study the proper scope that it deserves will require an in-depth study with an educational psychology focus.

While acknowledging the fascinating research themes emerging from educational psychology and neuroscience, this thesis only focuses on a subset of learner traits that can be inferred through analysing online behavioural data. An initial set of learner attributes is used as part of proof of concept evaluations for building a learner profile, but more of these attributes could be identified from education psychology in future research.

In this thesis, the learner profile will be used to establish differentiated learning experiences. Even though the model incorporates elements of adaptive education systems, it is beyond the scope to provide real-time, automated intelligent adaptive learning or personalised learning as described in Section 1.3.

1.6 Thesis Structure

This thesis is presented in six chapters that guide the reader towards a proposed solution to the problem identified in Section 1.2. The aim and main contributions of each chapter are highlighted next.

1.6.1 Chapter 1. Introduction

Chapter 1 establishes the context of the research, highlights the research problem and presents the aim, scope and significance of the study. The main points from Chapter 1 include:

- The introduction of online instructional design as the general research area
- The identification of the research problem
- The significance of the problem and justification of a proposed solution
- The research question that guided the study
- The research objectives that operationalised the research question
- The scope and structure of the thesis that reports on an emerging study of the problem

1.6.2 Chapter 2. Research Methodology

The aim of Chapter 2 is to describe and justify the approach followed to develop and evaluate a proposed solution to the problem identified in Chapter 1. The main points from Chapter 2 include:

- A discussion of the researcher's ontological, epistemological and axiological assumptions
- An overview of two scientifically rigorous research methodologies with a design focus:
    - Design Science Research (DSR) applied in information systems inquiries
    - Design-Based Research (DBR) applied in education inquiries
- The synthesis of an emerging methodology – Design Research for Technology Enhanced Learning (DeRTEL) based on best practices in DSR and DBR

1.6.3 Chapter 3. Instructional Technology and Learning Design

Chapter 3 examines relevant issues from the general research area of instructional design and instructional technology. Issues that led to identification of the research problem and the research question that guided an enquiry into the problem, are explored. The chosen topics represent a conceptual framework of the problem domain within which a solution is designed. A conceptual framework emerges from a focused literature review that synthesises the "state-of-the-art" discourse on the following topics:

- Instructional Design Models
  - Procedural and Conceptual Models that guide online instructional design
- Online Learning Design
  - Learning Objects (Characteristics, Types, Granularity and Metadata to describe the Learning Objects)
  - Provision of Differentiated Instruction by tailoring fine-grained Learning Objects
- Learning Management Systems
  - Moodle resources and activities to present Learning Objects (Learning Design)
  - Moodle tools to collect and analyse data (Learner Modelling)

1.6.4 Chapter 4. Educational Data Mining

The aim of Chapter 4 is to examine relevant topics around the issue of the use of learner data to optimise the learning environment. The chosen topics represent a conceptual framework of best practices in the solution domain. The conceptual framework emerged from a focused literature review that synthesised the "state-of-the-art" discourse on the relevant topics. The main points from Chapter 4 include:

- Learning Analytics
  - Potential uses of and techniques used in learning analytics
  - An ethical code of practice for learning analytics
  - Reflection on existing Learning Analytics conceptual frameworks and process models
- Learner Modelling
  - Techniques for building a learner profile
  - Learner characteristics recorded in a learner profile

1.6.5 Chapter 5. Iterative Development and Evaluation of Proposed Solution

Chapter 5 describes the iterative refinements of the proposed solution towards the problem identified in Chapter 1. This solution chapter is structured according to the iterations in the prototyping phase of the DeRTEL methodology and the objectives stated in Chapter 1. The main points from Chapter 5 include a delineation and elaboration of the process through which the solution components are iteratively designed and evaluated according to the DeRTEL methodology synthesised in Chapter 2. The model incorporates the conceptual frameworks synthesised in Chapter 3 and Chapter 4 and attempts to solve the problem identified in Chapter 1.

1.6.6 Chapter 6. Discussion and Conclusions

The aim of Chapter 6 is to highlight the main contributions of this thesis, examine the broader implications of the findings and plan a way forward. The main points from Chapter 6 include:

- A narrative discussion of the research process through which the proposed solution, described in Chapter 5, is derived and evaluated. The summary outlines how the solution:
    - Relates to the research question and objectives established in Chapter 1
    - Is informed by the conceptual framework from Chapter 3 and Chapter 4
    - Is a result of following the research design outlined in Chapter 2
- A critical evaluation of the proposed solution in terms of:
    - The type and level of knowledge contribution
    - Theoretical contribution: A summary of the main findings presented in this thesis in the form of an emerging process model
    - Practical contribution: Instantiation of phases of the process model and implications for the use of the proposed solution in teaching practice
    - Methodological limitations that constrain the proposed solution
- Recommendations for policy and practice, refinement of the current solution and further research

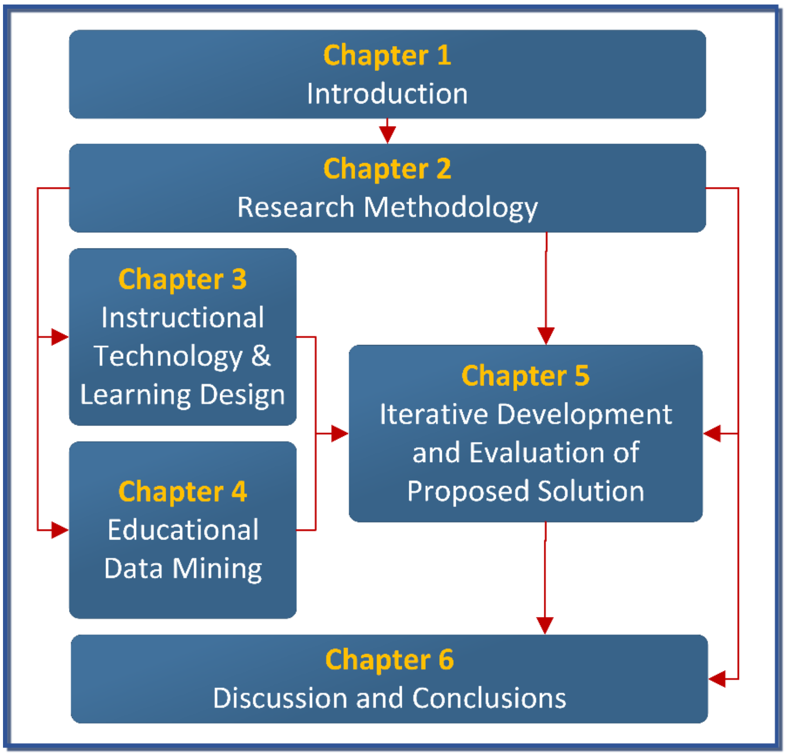


Figure 1.1 Thesis Overview

Chapter 2. Research Methodology

Chapter 2
Research Methodology

2.1) Introduction

2.2) Research Philosophy

- 2.2.1) Research Paradigm and Ontology (Reality)
- 2.2.2) Epistemology (Knowledge)
- 2.2.3) Axiology (Values)

2.3) Research Design

- 2.3.1) Philosophical Assumptions in Design Research
- 2.3.2) Design Science Research (IS) and Design Based Research (Education)
- 2.3.3) Design Research for Technology Enhanced Learning
  - **2.3.3.1) Preliminary Phase**
  - **2.3.3.2) Prototyping Phase**
  - **2.3.3.3) Assessment Phase**

2.4: Conclusion

2.1 Introduction

A research methodology includes the research design and underpinning research philosophy that guides the selection of research methods. This chapter makes explicit the research philosophy (Section 2.2) that underpins the methodological and research design choices made in this study (Section 2.3).

This study entails developing a learning analytics process model that describes a data-driven approach to online instructional design. The model and its subsequent instantiation are technological artefacts as conceptualised in Design Science Research. At the same time, it will be an intervention in an educational context, hence the need for principles to guide methodological choices suitable for instructional technology. Guidelines from Design-Based Research and information systems Development theory for online education are merged with Design Science Research guidelines (Section 2.3.2) to produce a research design that can guide this inquiry into Technology Enhanced Learning (TEL) (Section 2.3.3).

2.2 Research Philosophy

Researchers need to take a position and make known the philosophical underpinnings that inform their choice of research paradigm. Each paradigm is associated with an interrelated ontology, epistemology, axiology, methodology and method (Scotland, 2012).

2.2.1 Research Paradigm and Ontology (Reality)

Ontology is the study of the nature of reality. It addresses the researcher's assumption of what is real and how reality works. Related to the concept of an ontology, the term **paradigm** is defined as a set of beliefs that influence how someone views the world. The contemporary definition of a research paradigm, as conceptualised by scientific historian Thomas Kuhn (1977), identifies a scientific community of practitioners with shared goals at a particular point in time. Ever since Kuhn's seminal work that linked paradigms to communities of practice, several groupings have formed exhibiting the following philosophical views of reality (Saunders, Lewis and Tornhill, 2007):

- Positivists – a community of researchers who believe in a single observable reality that can be measured quantitatively

- Interpretivists\Constructivists – a community of researchers who believe reality needs to be qualitatively interpreted
- Subjectivists – a community of researchers who believe that reality is created from perceptions and actions of “social actors”
- Critical theorists – a community of researchers who believe that reality is socially constructed through internal relations of a society
- Realists – a community of researchers who believe that reality is independent of the mind and can be observed through the senses
- Pragmatists – a community of researchers who believe reality can be renegotiated and interpreted through any method that solves the problem at hand

In practice, choosing a particular paradigm does not exclude using principles from another (Saunders et al., 2007). Researchers exhibiting this view are generally classified as pragmatists. In pragmatism, the research question informs the choice of research approach. Morgan (2014) asserts that researchers should move beyond emphasising the problem-solving ability of pragmatism, to avoid downplaying other aspects of pragmatism as conceptualised by earlier philosophers such as Peirce and Dewey. Dewey (1997) in particular defines experience as a cycle between the source of beliefs that result in action and reflecting on these actions to determine the beliefs (Figure 2.1).

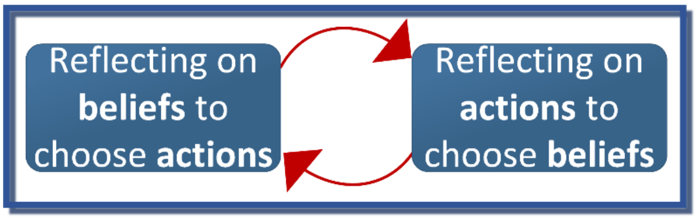


Figure 2.1 Dewey's Model of Experience (adapted from Dewey (1997))

Prior experience alone does not guide actions, context plays a role. Regarding choosing a research paradigm, context refers to the community of practice that will be interested in the outcomes of the research. This study, in particular, is aimed at those who subscribe to the pragmatist philosophy. The rest of Chapter 2 highlights this researcher’s set of beliefs and actions based on beliefs rooted in pragmatism. In particular, Section 2.2.2 clarifies the epistemological and Section 2.2.3 the axiological underpinnings of this study.

2.2.2 Epistemology (Knowledge)

Epistemology refers to the study of the nature of **knowledge** and the relationship between the knower and knowledge. Regarding research, epistemology drives how and why researchers choose the way they do research.

The choices made by the researcher during a research enquiry is based on experience as conceptualised by Dewey (Figure 2.2).

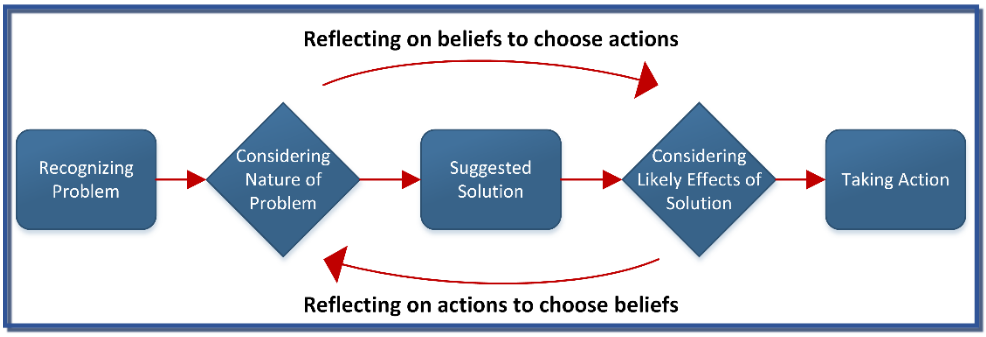


Figure 2.2 Dewey's Model of Enquiry (adapted from Dewey (1997))

Beliefs that have become problematic are examined and resolved through action (Morgan, 2014). Reflection on the researcher's beliefs and actions follows a continuous cycle throughout the enquiry. Some beliefs may be unknowable at the start of the inquiry and may only become clear at some point during. It begins by recognising that a situation is problematic and reflecting on the nature of the problem. Potential actions are suggested, and their consequences are considered before action is taken.

2.2.3 Axiology (Values)

Axiology refers to the study of values. A researcher's values affect what we value in the results of their research. Currently, the value that resonates most in terms of the outcome of this research is the notion of utility. The technological artefacts contributed through this research must be usable in the context in which they will be applied. The emphasis is therefore on problem-solving ability and providing an effective way for instructional designers to monitor learners' online behaviours through interrogating targeted metrics. The prescriptive level knowledge must enable positive change in the way instructors use this online data to influence their design choices for the online learning environment.

At the same time, this researcher also values the notion of ethics in research. As a result, care will be taken with any participant data that may be used as part of this study. This will involve removing all identifiable data points from datasets before analysis and reporting. In addition, a code of practice for learning analytics is included in this thesis.

2.3 Research Design

Having made explicit the ontological, epistemological and axiological lenses through which this study should be judged, the next step is to select a research design that aligns with these stated philosophical underpinnings. There is a wide choice of research designs, but the two that most commonly align with a pragmatic view of reality are Design Science Research (DSR and the closely related Design Based Research (DBR). Section 2.3.1 outlines the ontological, epistemological and axiological position of DSR and DBR. This is followed by the synthesis (Section 2.3.2 and Section 2.3.3) of a pragmatic research design that can drive an enquiry into the development and evaluation of a technology enhanced learning intervention.

2.3.1 Philosophical Assumptions in Design Research

Design Research integrates theory through argumentation with practice through constructing experimental solutions in applied contexts (Mehlenbacher, 2012). The ontological assumption of those who favour a design oriented approach to research view reality as being socio-technologically enabled (Vaishnavi and Kuechler, 2008). Due to the iterative nature of design research, the researcher's ontological viewpoint may shift throughout each cycle (Barab and Squire, 2004). This shift occurs as a result of action based on the researcher's beliefs, and then a reflection on these actions that in turn affects choice of beliefs. This is congruent with the view of pragmatists (Morgan, 2014).

The epistemological perspective of design science research is that knowledge is produced by iteratively building solutions towards problems within a specific context (Vaishnavi and Kuechler, 2008). Iivari (2007) defines three types of knowledge contributions: conceptual, descriptive and prescriptive. Conceptual knowledge includes terms, concepts and conceptual frameworks. This type of knowledge is used for analysing and describing the field of research.

Conceptual knowledge can also be used as a framework for prescriptive artefacts. Knowledge at the descriptive level includes theories, hypotheses and observational data (the what). Descriptive knowledge is analysed to form a conceptual model of a particular research field and provide input to prescriptive artefacts. Prescriptive level knowledge (the how) includes artefacts and recommendations for practice. Empirical research from artefacts produces descriptive knowledge and provides input to conceptual frameworks. Prescriptive knowledge is closely aligned to a pragmatic view of reality. Since this researcher prescribes to the notion of pragmatism, the main goal of this enquiry will be to produce prescriptive knowledge. Conceptual and existing prescriptive knowledge will be interrogated and applied in the production and application of technological artefacts to solve a real problem.

Axiologically, the design science researcher does value conventional research concerning descriptive truth or conceptual understanding of phenomena, but the creative control of the environment is often more highly valued (Vaishnavi and Kuechler, 2008). The design science researcher appreciates the importance of shaping phenomena in the real world through ethically creating artefacts with utilitarian value (Aljafari and Khazanchi, 2013). The value of the artefacts rests in its ability to have a positive impact on the context in which it is applied (Barab and Squire, 2004).

2.3.2 Design Science Research (IS) and Design Based Research (Education)

This Section derives a methodological design process that is scientifically rigorous and appropriate for technology enhanced learning.

Simon (1996), in one of the earliest works introducing intellectual rigour to design activities, distinguishes research in the natural sciences with research in the "science of the artificial". While the focus of research in natural science is on describing and explaining how objects in nature or society behave and interact, research into man-made objects focuses on how they are designed to meet predefined goals.

Building on the ideas of Simon (1996) and design research in other fields, Hevner, March, Park and Ram (2004) developed guidelines for conducting, evaluating and presenting design science research in the Information System (IS) discipline. DSR produces technological artefacts as relevant solutions to problems identified in a specific context. These artefacts can take the form of a construct, model, method or instantiation. The artefacts must be iteratively developed and evaluated through rigorous methods. DSR must contribute towards an existing knowledge base (Figure 2.3).

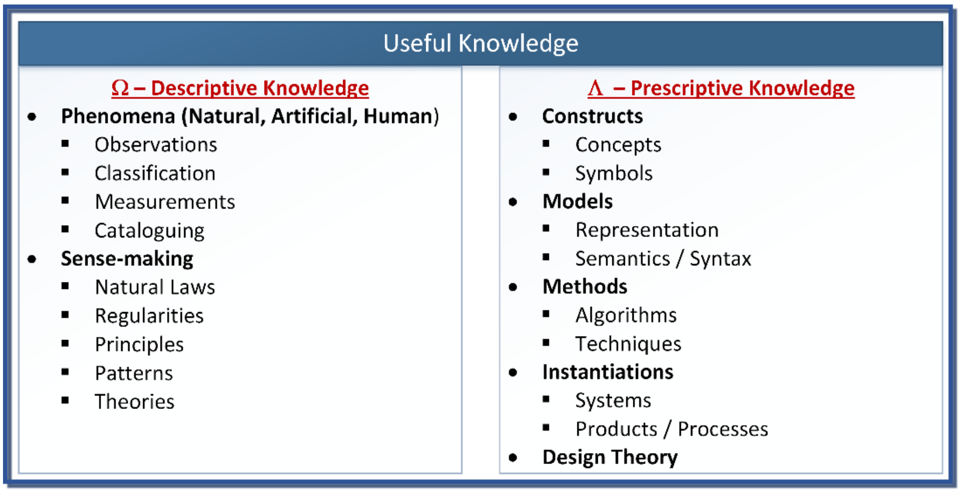


Figure 2.3 Design Science Research Knowledge Base (adapted from Gregor and Hevner (2013))

These contributions can be towards foundational knowledge such as theories, frameworks, instruments, constructs, models, methods and instantiations. Contributions can also be made towards methodologies such as data analysis techniques, formalisms, measures and validation criteria. Contributions must be shared with multiple relevant stakeholders such as fellow researchers, practitioners and management.

Although originally conceptualised for IS research, Gregor and Hevner (2013) acknowledge that the DSR principles apply equally well to the design of any technological invention. This study focuses on technology applied to an educational context, hence there is a need to incorporate principles of theories applicable to instructional technology in education.

Independent from Hevner et al. (2004), Jones and Gregor (2006) formulated an Information Systems Design Theory (ISDT) for online learning. The principles described in ISDT are derived from the development and improvement of an online learning system previously used at the Central Queensland University, called Webfuse. In 2010 the university abandoned the Webfuse system in favour of the Moodle Learning Management System.

The ISDT of Gregor and Jones (2007), however, still provides principles that apply to design science research. These are incorporated into a knowledge contribution framework for Design Science Research (Gregor and Hevner, 2013). This framework essentially combines the knowledge contribution from the Information Systems Design

Theory for web based education (Gregor and Jones, 2007) and Design Science in IS Research (Hevner et al., 2004).

Design Based Research (DBR) in education is defined as the systematic study of analysing, designing, developing and evaluating educational interventions (Plomp and Nieveen, 2013). DBR research must be rooted in a real educational context (Anderson and Shattuck, 2012) and focus on designing and testing an education intervention. Typical interventions include, but are not limited to:

- A novel learning or assessment activity
- Changes in an institution's administrative process
- Technological tools to enhance learning

DBR aims to develop solutions for problems for which no "how-to" guidelines exist (Figure 2.4). A DBR approach is an iterative enquiry that starts with observing problems identified through a collaborative partnership between researchers and education practitioners. Solutions to these problems are iteratively developed and evaluated. Prescriptive design principles are abstracted from the iterative development and evaluation and shared with the research community and practitioners.

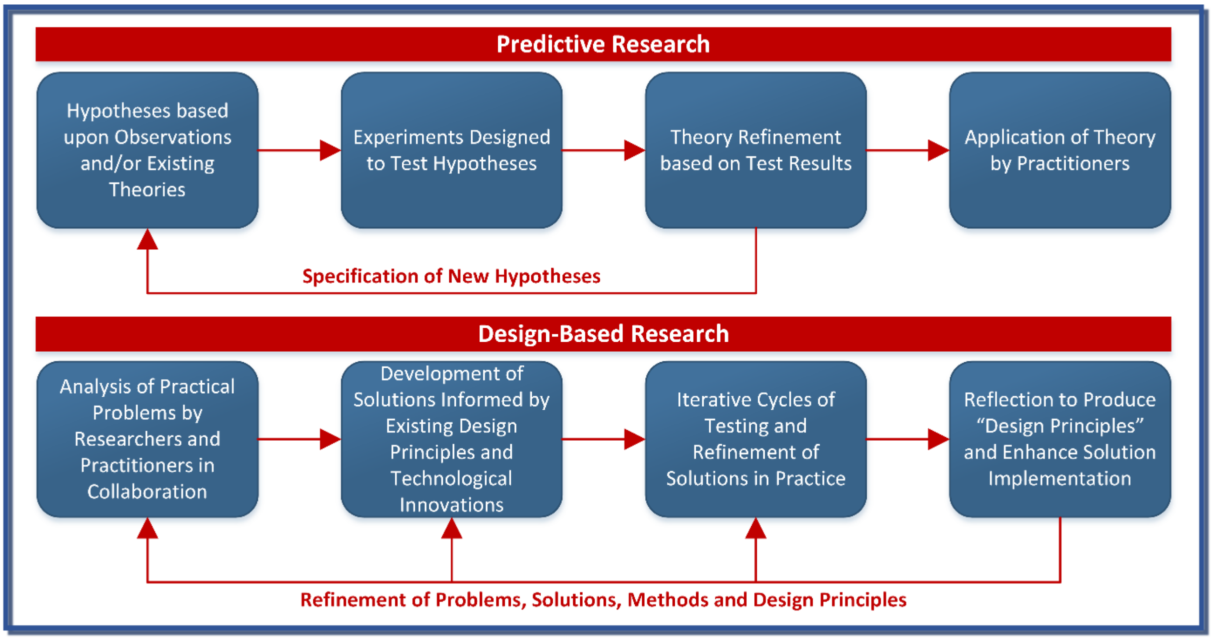


Figure 2.4 Predictive and Design Based Research (adapted from Reeves (2006))

DBR contrasts with traditional predictive research that aims to prove or develop theories based on hypotheses (Figure 2.4). Experiments are conducted to test hypotheses and the experimental results are used to further refine the theory. The theory is shared with practitioners who can apply it in their practice.

A model developed through design research can describe how to solve a complex problem in education practice (Van Den Akker, 1999). To distinguish design research from systematic educational design, it also aims to contribute to the scientific body of knowledge. It achieves this through applying research rigour in three main phases: Preliminary phase, Prototyping phase and Assessment phase. During the preliminary phase a needs analysis and content analysis is performed through techniques such as site visits, conducting interviews and review of state-of-the-art literature. The preliminary phase helps develop the theoretical or conceptual framework for the study. The prototyping phase is iterative and involves the development and formative evaluation of educational interventions. The assessment phase is semi-summative and concludes a particular study. Due to the nature of design research, further refinement of the intervention will remain ongoing.

DBR is aimed at systematically designing, developing and evaluating solutions for complex educational problems or to validate theories of educational practice (Van Den Akker, Bannan, Kelly, Nieveen and Plomp, 2010). The main aim is to enhance the utilisation of education research (Anderson and Shattuck, 2012) in the form of solutions to problems identified and described through “traditional” research approaches. Where traditional approaches focus on gaining a theoretical understanding of the domain, Design Based Research focuses on practical application in context (Herrington, McKenney, Reeves and Olive, 2007).

While DBR is particularly useful in research with an education perspective (Anderson and Shattuck, 2012; Design-Based Research Collective, 2003; Plomp and Nieveen, 2013; Van Den Akker et al., 2010), DSR is useful for IS research (Gregor and Hevner, 2013; Hevner and Chatterjee, 2010; Iivari, 2007; Jones and Gregor, 2006; Thakurta, Mueller, Ahlemann and Hoffmann, 2017). The current study represents the first step towards establishing a data driven solution to educational problems. A combination of education oriented DBR and information systems oriented DSR is therefore deemed suitable as a research design.

2.3.3 Design Research for Technology Enhanced Learning (DeRTEL)

Table 2.1, Table 2.4 and Table 2.6 compare steps and activities from DSR and DBR and concludes with a consolidation of the two design approaches into an emerging methodology, Design Research for Technology Enhanced Learning. This methodology (DeRTEL) as applied to the current study is described in Section 6.2. Each phase of DeRTEL will be elaborated on next and is summarised in Figure 2.7 on page 34.

2.3.3.1 Preliminary Phase

The preliminary phase (Table 2.1) consists of interrogating the Application Domain and the two main components of the Knowledge Base (Problem and Solution Domains):

- Context (Application Domain):
  - General – Contextual Analysis of the general research area, delineated down to area of interest, concentration, focus and topic
  - Deployment Site – Contextual Analysis of the institution or institutions where the solution will be deployed and assessed
- Problem Domain: Identification of a problem in education that a technological intervention may address
- Solution Domains: Identifying the goals and scope of a potential solution

The preliminary phase lay the foundation for a technological artefact to be iteratively developed and applied in an educational context:

Step 1. Identify an educational problem that a technological intervention may address

In step one, the researcher must conduct a general literature review and/or consult with educational practitioners or researchers in order to identify potential problems or opportunities in the educational context. Once a relevant problem has been identified, a conceptual or theoretical framework of the problem domain must be established through a literature review and/or site evaluation.

Step 2. Conduct a requirements analysis to determine the goals of the solution

In step two, an exploratory evaluation of existing interventions in the problem domain should yield best practice guidelines for potential solutions. The initial evaluation can include:

- Recent insights by experts or practitioners
- Prior interventions to similar problems
- A focused literature review
- The practical context in which the problem has been identified

Step 3. Perform contextual analyses

Step three is a two-part process that changes scope throughout the project lifecycle. Initially, before the solution is developed during the prototyping phase, the focus is on establishing the general background of the study. This involves delineating the area of interest, and further scoping down to the concentration, focus and a specific topic. The results are combined with the problem statement and solution goals and scope to form a conceptual framework as input into the prototyping phase.

Once the solution is developed and ready to be deployed, a more in-depth site analysis is necessary. The results may require the solution to be adapted to the site's requirements. The context within which the solution will be applied must be evaluated to determine the intervention's scope and relevance. This includes:

- The roles, capabilities and characteristics of people that will interact with the solution in some way, to enhance acceptance of the solution.
- The education institution's vision, mission, values, strategies, teaching and learning policies. This will give an indication whether management and stakeholder support can be counted on.
- The current Information and Communication Technology (ICT) infrastructure and support in place at the educational institution. This could include hardware, software, communication networks, data and support services and software development capabilities of the institution within which the solution will be implemented.

From the best practice guidelines and the contextual analysis, the researcher must define the objectives of the solution. These goals should motivate what the proposed intervention must accomplish that previous interventions could not.

Table 2.1 Comparing Preliminary Phase of Design Research Frameworks

<table>
<tr><th colspan="2">Summary of Existing Design Frameworks</th><th rowspan="2">Proposed research steps and activities based on DSR and DBR</th></tr>
<tr><th>Design Science Research (DSR used in IS)</th><th>Design Based Research (DBR used in education)</th></tr>
<tr><td>Awareness of problem
• Problem identification and motivation
• Define objectives of technology solution</td><td>Problem Identification and Needs Analysis
• Identify problem in context
• Conduct needs analysis</td><td>Develop Conceptual Framework
• Context: Describe background to general field of study and deployment site
  ○ Delineate to area of interest, concentration, focus, topic
  ○ Investigate site where solution will be deployed
• Problem: Identify an educational problem that a technological intervention may potentially address
  ○ Practitioners or Researchers
• Solution: Conduct a requirements analysis to determine the objectives and scope of the solution
  ○ Existing approaches to similar problems</td></tr>
</table>

2.3.3.2 Prototyping Phase

The preliminary phase should yield a relevant problem statement, the envisaged goals and scope of a potential solution and a general description of the application domain in a conceptual framework. The prototyping phase involves developing and formatively evaluating a proposed solution for technology enhanced learning through multiple cycles – starting with a tentative design and culminating in a technological intervention that can be fully implemented in an educational context.

Step 4. Iteration 1: Construct and evaluate a tentative proposal

During the first cycle, a proposal for a tentative solution must be constructed based on the solution goals and scope. The solution must address a problem statement towards a situation of concern within an educational context. The tentative proposal must reflect the current discourse in the problem and solution domains. To ensure research rigour, the researcher should also classify the type and level of knowledge contribution that the study will make to the existing knowledge base. The contribution types are based on solution and application maturity, classified as high or low in Figure 2.5. A routine design exercise, in which known solutions are applied to known problems have no knowledge contribution and is therefore not suitable as a research inquiry. Based on this maturity model, knowledge contributions can be classified as improvement, exaptation or invention. A knowledge contribution is classified as an invention if a new solution is developed for a previously unknown problem. This is a highly rare form of knowledge contribution and examples in current literature are scarce (Gregor and Hevner, 2013). A knowledge contribution is deemed an improvement if it is a new solution for a known problem and an exaptation if it adopts solutions from other fields to new problems.

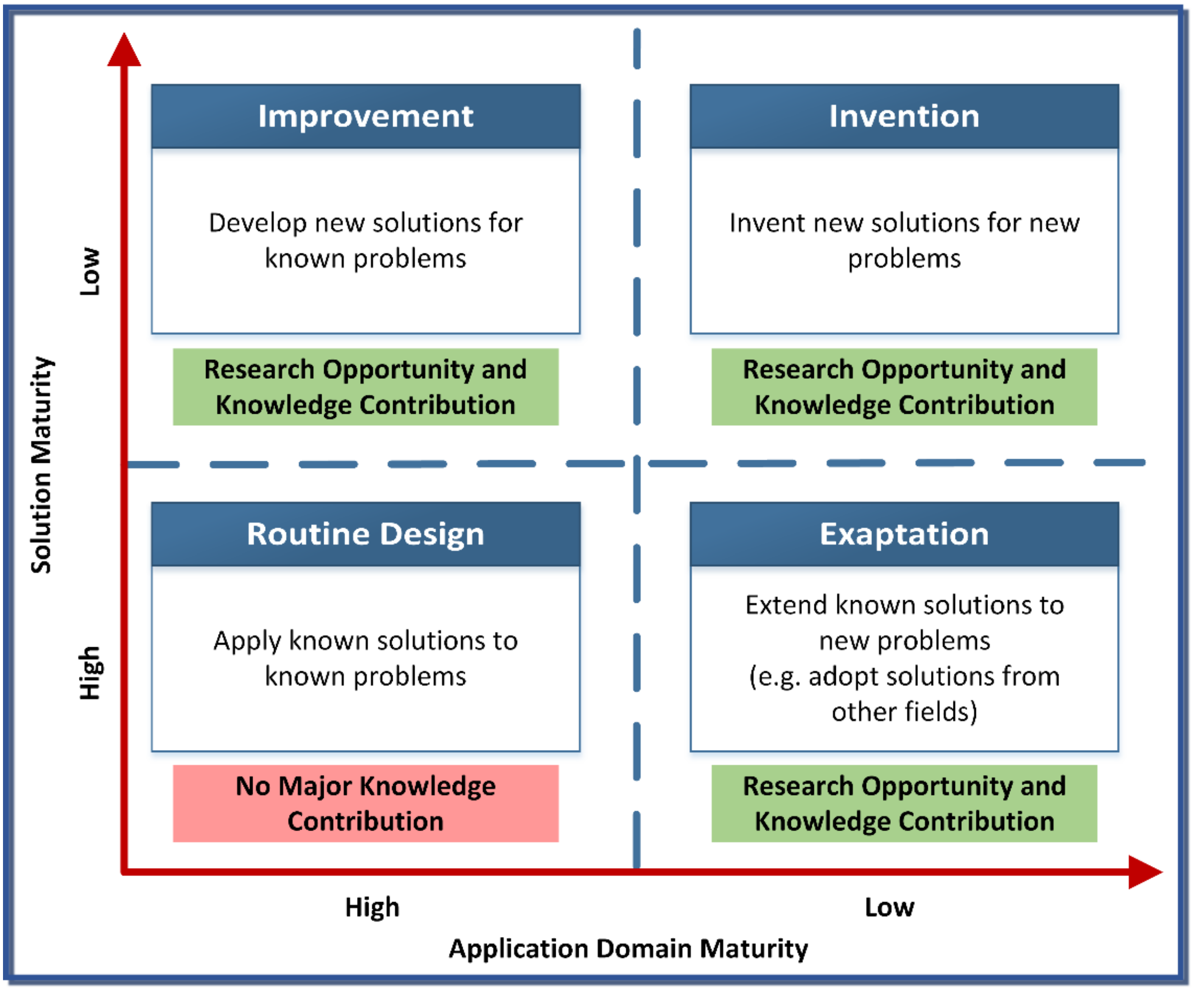


Figure 2.5 DSR Knowledge Contribution Framework (adapted from Gregor and Hevner (2013))

When classifying a contribution on sliding scales from specific to abstract, limited to complete and less mature to more mature, three levels can be identified (Table 2.2; (Baskerville et al., 2018)):

- Level 1: Situated implementation of artefact, e.g. instantiation of a software product or application of a process to develop and evaluate the product
- Level 2: Emerging design theory in the form of prescriptive knowledge, e.g. constructs, methods, models, design principles and technological rules
- Level 3: More complete mid-range or grand design theories about embedded phenomena

Table 2.2 DSR Research Contribution Types

Knowledge Levels	Contribution Types	Example Artefacts
More abstract, complete and mature knowledge	**Level 3:** Well-developed design theory about embedded phenomena	**Design Theories:** (Mid-range and Grand Theories)
	Level 2: Nascent design theory – knowledge as operational principles / architecture	**Prescriptive Knowledge:** • Constructs • Methods • Models • Design Principles • Technological Rules
More specific, limited and less mature knowledge	**Level 1:** Situated implementation of artefact	**Instantiations:** (Software Products or Implemented Processes)

Since evaluation forms an integral part of the prototyping phase, an evaluation strategy must form part of the tentative design proposal. Four types of evaluation strategies are (Figure 2.6):

- Quick and Simple
- Human Risk and Effectiveness
- Technical Risk and Efficacy
- Purely Technical

The focus of the strategies is based on two dimensions – the functional purpose of the evaluation (between formative and summative) and the way the evaluation is done (between artificial and naturalistic). The strategies involve interspersing solution design and development phases with evaluation milestones.

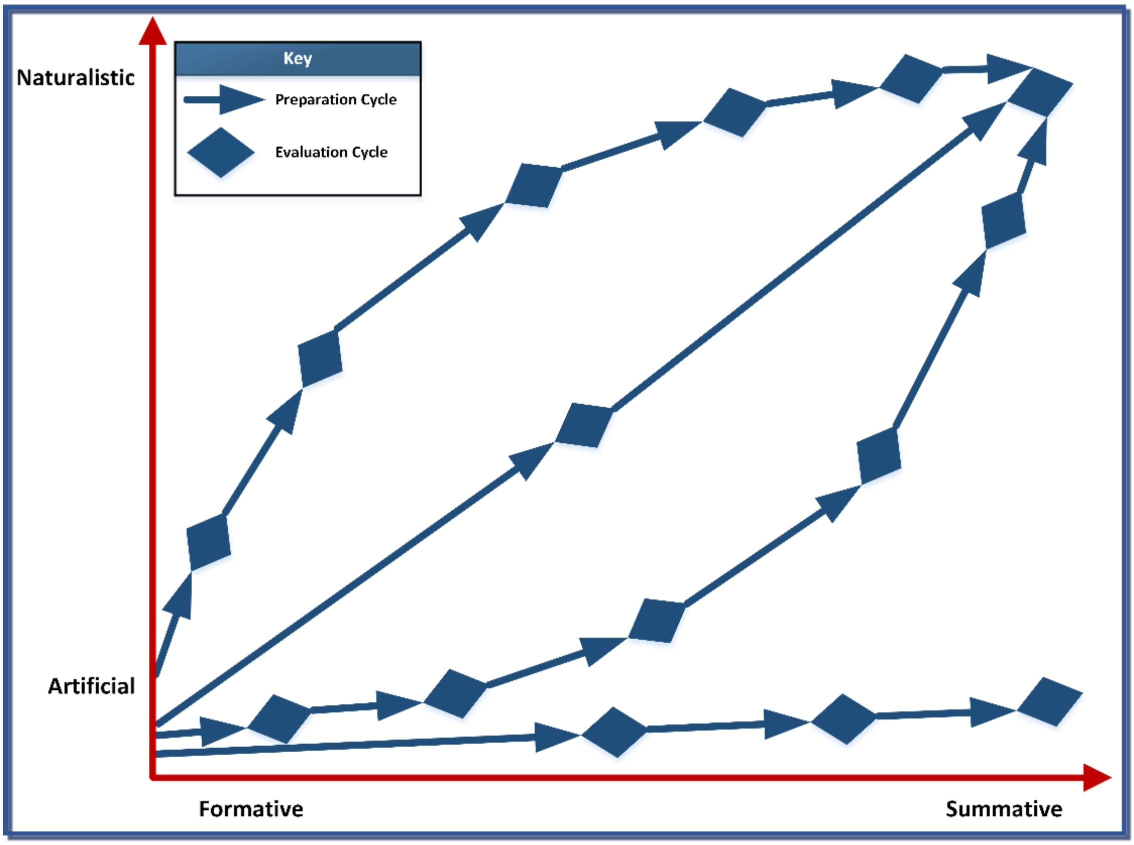


Figure 2.6 Framework for Evaluation in Design Science (adapted from Venable et al. (2016))

A quick and simple strategy proposes rapid completion of the design and development that moves evaluation quickly through a single artificial formative evaluation and single summative evaluation in a naturalistic setting. The human risk and effectiveness strategy starts early with formative evaluations, but moves rapidly from artificial to naturalistic settings, where rigorous evaluations are done to measure the effectiveness of the implemented solution. The technical risk and efficacy strategy emphasises the need for multiple artificial formative evaluations to measure the contribution of subcomponents towards the final solution. Once subcomponents are satisfactorily tested, the focus shifts to multiple summative evaluations of the solution in the naturalistic environment. A purely technical strategy implies formative and summative evaluations of the solution only artificially. The choice of strategy is based on the following aspects (Table 2.3):

- Whether design risk is technical or social
- The cost of performing the evaluations and access to real users in practice
- Goal of the solution

Table 2.3 Design Science Evaluation Strategies

Evaluation Strategy	Design Risk	Cost and Access to Real Users	Goal
Quick and Simple	Low Technical and Social	Low cost, quick access	Fast roll out
Human Risk and Effectiveness	Mainly Social	Low cost, quick access	Rigorously evaluate long term impact of solution in practice
Technical Risk and Efficacy	Mainly Technical	Expensive, delayed access	Rigorously test subcomponents of solution before impact analysis in practice
Purely Technical	Completely Technical	No or delayed access to real users	Rigorously test technical aspects of system

The researcher can use the following heuristics to select an appropriate evaluation strategy:

Select Quick and Simple if:

- The design risk has low technical and human implications
- It is relatively cheap and possible or necessary to test rapidly with real users
- A major goal of the intervention is rapid roll-out in practice

Select Human Risk and Effectiveness if:

- The design risk is mainly human in nature
- Real users can be accessed quickly and at relatively low cost
- A major goal of the intervention is the rigorous evaluation of the long-term impact of the solution in practice

Select Technical Risk and Efficacy if:

- The design risk is mainly technical in nature
- Access to users is either prohibitively expensive or it is not necessary to test the solution in practice until later
- A major goal of the intervention is the rigorous formative evaluation of parts of the solution before final roll-out

Select Purely Technical if:

- The design risk is completely technical in nature
- Access to users is not possible or not necessary
- A major goal of the intervention is to test rigorously the technical aspects of the solution, without human/social implications

In general, to ensure research rigour, at least the following evaluation goals need to be included at various stages throughout the prototyping phase (Van Den Akker et al., 2010):

- Relevance/Content validity – The solution addresses a real need, the design is based on current scientific knowledge
- Consistency/Construct validity – The solution is viable and logically designed
- Actual/Expected Practicality – The solution is/has the potential to be usable
- Actual/Expected Utility – The solution achieves/can potentially achieve its goals

Any design research would have different goals, and therefore unique evaluation criteria. The researcher needs to investigate and add other potential evaluation criteria to the above as part of the evaluation strategy. Techniques for evaluation will also differ from project to project, between different phases and even for different goals. Some examples of techniques could include, but are not limited to:

- Screening with a checklist, content analysis
- Questionnaires, focus group or one-on-one interviews
- Simulation or cognitive walkthroughs with experts or potential users
- Software unit testing/pilot study/micro-evaluation to observe all or part of the solution in a limited capacity with test users

Evaluation techniques are numerous, and the researcher should investigate and select appropriate techniques suitable to the solution and the goal of the evaluation. The primary focus of evaluation during iteration one should be to ensure relevance of the proposed design, to ensure that the solution addresses a real need.

Step 5. Iteration 2: Create and evaluate global design

Iteration two involves elaborating on the tentative design by specifying the requirements for the proposed solution. If the solution involves the implementation of a software artefact, the overall architecture of the solution should be drawn up through relevant software development models. At this point black box designs that hide inner workings will suffice. The primary focus should be on showing how components of the solution would interact with each other. Since the problem and solution domains will differ from project to project, it is beyond the scope of this thesis to go into detail with regard to design principles for all possible types of solutions. It is incumbent upon the researcher to investigate and incorporate appropriate design principles into whatever solution is proposed.

The global design must at least be evaluated for consistency to ensure the solution is viable and logically designed.

Step 6. Iterations 3 to n: Develop and evaluate sub components of solution

Step six represents several iterations that focus on different components of the solution in a stepwise fashion. In software development projects, teams typically each have a specific module that they contribute towards the overall system. They can do so concurrently, following global design principles to ensure interoperability. The same principle applies to the iterations in step six. During these iterations, some components may be implemented and unit tested using any appropriate evaluation technique. For example, if the artefact being developed is a process model, the different phases of the model should get attention in each iteration. Subcomponents must be evaluated for expected and\or actual practicality and effectiveness. This is to ensure the solution is usable in practice and able to achieve its goals.

Before proceeding to the next phase, the researcher should reflect on the results of the evaluations of each subcomponent. The results should be worked back into the constituent parts of the solution. Throughout the iterations, formative assessment findings can be shared with the context and the knowledge base.

.

Table 2.4 Comparing Prototyping Phase of Design Research Frameworks

Summary of Existing Design Frameworks		Proposed research steps and activities based on DSR and DBR
Design Science Research (DSR used in IS)	**Design Based Research (DBR used in education)**	

Suggestion and Development
• Classify knowledge contribution type and level
• Build artefacts
 	
Formative Evaluation
• Determine evaluation strategy
• Identify Evaluation Criteria
• Examine artefact against evaluation criteria, reflect and iterate back to design
	Design and Implement
	• Design proposal
	• Global design
	• Partly detailed intervention
	• Completed intervention
	
	Formative Evaluation
	• Identify Quality Criteria
	• Examine artefact against evaluation criteria, reflect and iterate back to design
		• Iteration 1: Tentative Proposal
		  ○ *Prepare*: Describe form and function of a tentative solution (process and product); evaluation strategy and criteria
		  ○ *Evaluate*: Motivate relevance of tentative solution
		    ▪ Reflect on area of interest, problem statement, solution goals and scope, type and level of knowledge contribution
		• Iteration 2: Global Design
		  ○ *Prepare*: Elaborate on tentative proposal by specifying requirements and\or technical specifications for solution
		  ○ *Evaluate*: Motivate consistency of global design
		    ▪ Reflect on design against existing solutions and best practices
		• Iterations 3 to n: Develop and Evaluate Subcomponents
		  ○ *Prepare*: Refine components of solution design
		    ▪ Instantiate and/or pilot test components
		  ○ *Evaluate*: Assess practicality and effectiveness of units
		    ▪ Reflect on expected\actual practicality and effectiveness

2.3.3.3 Assessment Phase

Step 7. Develop deployment plan and stabilise solution

The assessment phase starts with the development of a deployment plan that needs institutional approval before implementing the solution at the deployment site. A contextual analysis of the application domain is necessary to complete the deployment plan. The contextual analysis should investigate the site readiness in terms of the people that will engage with the solution, the existing infrastructure and the strategic management of the institution. The solution may need revision to ensure it matches the site's requirements. Once the solution is stable, the initial roll-out of the complete solution can commence.

Step 8. Demonstrate and evaluate solution in practical context

Once the solution is live, it needs to undergo semi-summative evaluation in the practical context. This step involves impact studies of the solution. The researcher must observe whether the solution produced the desired goals. At the very least, the actual practicality and effectiveness of the entire solution must be tested. Any feedback from actual use must be incorporated into subsequent improvements. It is also conceivable that new opportunities or problems may arise through continual use, which means that a new research design project should be initiated.

Step 9. Communicate findings to researchers and practitioners

To ensure that the knowledge contributions (product and process) are open to professional scrutiny and critique, the semi-summative assessment research finding must be shared on any relevant forum. The practical and theoretical contributions of the solution must be shared with the research community, institutional management and practitioners.

The practical contribution or implemented artefact must serve as an example outcome of what can be achieved by applying the theoretical contribution. The implemented artefact should also be open to scrutiny and be available to test any claims made with regard to the outcome of the intervention. As a result, the researcher needs to clarify the context within which the implemented artefact is meant to be used. In addition, the intended users (e.g. lecturers and/or students) of the solution and the ICT staff that will ultimately support the solution must be trained in its intended use.

Table 2.5 Framework for a Theoretical Contribution

Component	Description
1. Purpose and scope	Describes the type of artefact and the goals that an artefact of this type aims for Describes the boundaries of the theory
2. Constructs	Representations of the entities of interest in the artefact
3. Principle of form and function	The architecture that describes the artefact (product or process)
4. Artefact mutability	The anticipated degree of change expected in the artefact
5. Testable propositions	Truth statements about the theory, e.g. predictions about outcomes achieved through use of the artefact
6. Justificatory knowledge	The kernel theories that provide the basis for explaining the design
7. Principles of implementation	Processes for implementing the theory (product or method) in a specific context
8. Expository instantiation	Physical implementation of the artefact used to represent and explain the theory. Can be used to test claims made in the theory.

A theoretical contribution can take the form of a number of different types, each capable of answering different types of questions (Jones, Gregor and Lynch, 2003). A theory that is meant to analyse and describe can answer “What is?” type questions. A theory that is meant to aid understanding of some phenomena can answer “How and why?” type questions. A predictive theory is used to predict “What will happen if …?”. A theory of design and action describes how to do something. Gregor and Jones (2007) describe the components of a design theory (Table 2.5). The theoretical contribution of an inquiry following DeRTEL steps can be described using this framework.

Table 2.6 Comparing Assessment Phase of Design Research Frameworks

Summary of Existing Design Frameworks		Proposed research steps and activities based on DSR and DBR
Design Science Research (DSR used in IS)	**Design Based Research (DBR used in education)**	

Demonstrate artefact in context
• Consider practical and theoretical contribution
• Communicate findings
	Summative Evaluation
	• Determine extent to which intervention lead to desired outcomes.
		Summative Evaluation
		• *Deployment Site Context*: Prepare and evaluate deployment plan
		  ○ Examine site where educational technology will be implemented (willingness to change, conditions, vision, mission, values, resources available, stakeholders)
		• *Impact Analysis*: Evaluate actual practicality and effectiveness of entire solution in context
		  ○ Observe whether solution produced desired goals
		  ○ Keep improving solution based on continual evaluation
		• *Dissemination*: Communicate findings to researchers and practitioners
		  ○ Reflect on practical and theoretical contributions of solution (formative and summative)

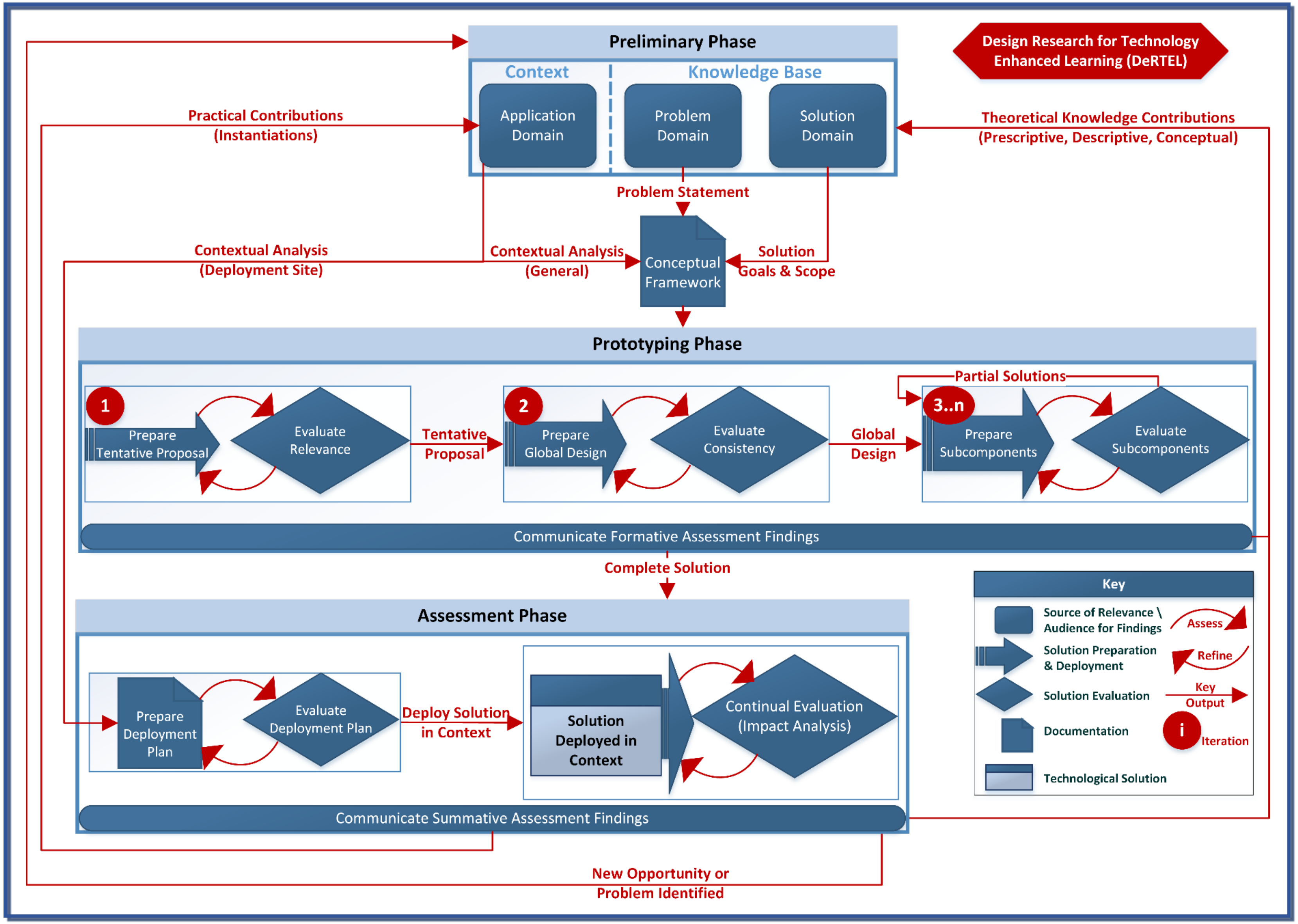


Figure 2.7 Design Research for Technology Enhanced Learning (DeRTEL) (Own Construction)

2.4 Conclusion

The aim of this chapter is to give a broad overview of the design that guided the study on which this thesis is based. An emerging methodology, Design Research for Technology Enhanced Learning (DeRTEL), is synthesised by combining Design Science Research (DSR), used in Information System and Design Based Research (DBR), used in education. DeRTEL can be used to iteratively develop and evaluate an artefact to be used for Technology Enhanced Learning. DeRTEL proposes three phases: preliminary, prototyping and assessment (Figure 2.7).

The preliminary phase establishes a conceptual framework of the application, problem and solution domains. The prototyping phase consists of several cycles through which the solution is iteratively developed and evaluated.

- The first cycle develops a tentative proposal for the artefact and is evaluated for relevance.
- The second cycle develops the requirements for the global design of the artefact and is evaluated for consistency.
- The remaining cycles focus on implementing subcomponents of the global design until the solution is completely developed. Each subcomponent is evaluated for expected or actual practicality and/or utility.

The final assessment phase starts with an in-depth analysis of the deployment site, and once approval is granted the solution is deployed at the targeted institution. The solution undergoes a semi-summative evaluation to determine the impact of the artefact in practice. All formative and summative findings are shared back to the application domain and the knowledge base. The application of each step in DeRTEL is described in different thesis chapters as indicated in Table 2.7.

Detailed methods and techniques are discussed in Chapter 5, where each iteration of the artefact is described in detail. Whenever a theoretical approach is used towards analysis, the following methods apply:

- Selection of appropriate sources relevant to the objective of this study
- Discussion, comparison and critical review of existing models in context

When a technological approach is used, subcomponents of the proposed model is instantiated in a standard Moodle installation and evaluated for practicality and utility.

Table 2.7 Thesis Chapters linked to DeRTEL Phases

Phase	Step	Thesis Chapter	Analysis Approach
Preliminary	Describe General Context	• Section 1.1	• Theoretical
	Describe Problem Domain	• Section 1.2 • Chapter 3	• Theoretical
	Describe Solution Domain	• Section 1.3 • Chapter 4	• Theoretical
Prototyping	<u>Iteration 1</u>: Describe Tentative Proposal	• Section 5.2	• Theoretical
	<u>Iteration 2</u>: Describe Global Design (Sub-objective 1: Process model for differentiated instruction using learning analytics)	• Section 5.3	• Theoretical
	<u>Iteration 3</u>: Refine solution component (Sub-objective 2: Differentiated learning design in Moodle)	• Section 5.4	• Technological
	<u>Iteration 4</u>: Refine solution component (Sub-objective 3: Learner modelling in Moodle)	• Section 5.5	• Technological
	Describe Deployment Site Context	• Future Work	• Theoretical
Assessment	Describe Impact Analysis	• Future Work	• Technological

Chapter 3. Instructional Technology and Learning Design

The aim of Chapter 3 is to examine relevant issues from the general research area of instructional design and instructional technology. Issues that led to identification of the research problem and the research question that guided an enquiry into the problem, are unpacked. The chosen topics represent a conceptual framework of the problem domain for which a solution was designed. The conceptual framework emerged from a focused literature review that synthesises the state-of-the-art discourse in the relevant topics. The main points from Chapter 3 include:

- Instructional Design Models
  - Procedural Models
  - Conceptual Models
- Online Learning Design
  - Learning Objects
  - Differentiated instruction through tailored Learning Objects
- Learning Management Systems
  - Moodle resources and activities
  - Moodle data sources

In order to achieve the primary research objective of enabling differentiated instruction, there is a need to first investigate current instructional design models. In line with the pragmatic philosophical stance underpinning this study, the focus is on instructional design models popular among practitioners. This includes procedural models ADDIE, Dick & Carey, Gerlach-Ely, ASSURE, Backward Design and SAM (Section 3.2.1). In addition, the conceptual models of Kemp, the 3P model, the conceptual framework of high-quality online learning environments, Mehlenbacher's instructional design principles and Kirkpatrick's model for evaluating training effectiveness are also investigated (Section 3.2.2). The purpose of the review into these procedural and conceptual models is twofold:

- To motivate the relevance of the research problem and subsequent solution goals and scope
- To derive guidelines and procedural steps for creating student-centric learning experiences

Since the learning design addressed in this thesis is instantiated in an online environment in general, the characteristics of learning objects used to deliver online content are explored. To guide the discussion on learning objects in this and subsequent chapters, the seminal work of Wiley (2002) and Hodgins (2006) are adopted as the guiding definitions (Section 3.3.1.1). The pragmatic approach of CISCO Systems (2001) provides structure to the delivery of online content (Section 3.3.1.2), while in Section 3.2.1.3 the language to describe this content is provided by the IEEE 1484.12.1 – 2002 Standard for Learning Object Metadata (IEEE 1484.12.1, 2002b).

The primary objective is to enable differentiated instruction, hence, the work of Tomlinson et al. (2003) is examined. Differentiated instruction necessitates the tailoring of granular learning objects and within this context, the work of Brusilovsky (1996), Brusilovsky, Wade & Conlan (2007) and Battou et al. (2011) on real-time adaptive systems are used to extract an architecture for providing tailored learning objects (Section 3.3.1). Applying principles of real-time adaptive systems to a study on differentiated instruction further ensures the model can be adapted to other forms and levels of adaptation in subsequent cycles.

This study narrows the scope of the online environment to Learning Management Systems in general, and the instantiation of the model is evaluated within Moodle in particular. Consequently, a description of Moodle resources and activities (Section 3.4.1) are extracted from Moodle documentation (Moodle Docs, 2018b, 2018a). Finally, the development of a learner profile for this study relies upon learner data captured from a Learning Management System, hence, Moodle data sources are examined (Section 3.4.2).

3.1 Introduction

The view of instructional design (ID) taken in this thesis mirrors that of Richey, Klein and Tracey (2011), who define ID as “the science and art of creating detailed specifications for the development, evaluation and maintenance of situations which facilitate learning and performance (p. 3)”. The view of instructional design as a science is congruent with the definition of Smith and Ragan (1999) that perceives it as a systematic and reflective process that translates learning principles into the design of instructional materials and activities. The view of instructional design as an art resonates with (Kemp, Ross and Morrison, 1998), whose model encourages manipulating learning principles and design elements in imaginative ways that may result in unanticipated learning experiences, yet still be backed by sound educational principles.

Kemp et al. (1998) identify seven principles underlying successful instructional design processes. They:

(1) *Are systematic and precise:* The process applies an orderly method of identifying, developing and evaluating strategies to attain an instructional goal. Each element of the process requires rigorous attention to detail and has a sound basis in theory.
(2) *Are applicable to course development:* Instructional design applies to constructing units of learning for a particular course. This happens after institution-wide curriculum planning.
(3) *Are used by instructors as a planning tool:* Instructional design processes are used by instructional designers to plan instructional materials as part of units of learning. The planning documents are usually not presented to learners.
(4) *Enable planning for satisfactory achievement of all learners:* Instructional design processes should include a feedback loop whereby learner achievement can influence refinements to the learning material.
(5) *Are informed by quality data:* At each stage of the instructional design, data that are relevant, consistent, practical and effective must be obtained and worked into the design.
(6) *Focus on learner instead of content:* Learner characteristics, goals and achievement must influence any design choice made in the construction of learning materials.
(7) *Are not exclusive in their ability to design quality instruction:* While instructional design processes can reduce the dependence on intuition and curtail unsubstantiated development of instructional materials, there is no single guaranteed path to success. The success of an instructional design process is its ability to guide instructors towards preparing instructional materials that support satisfactory learning in an acceptable timeframe.

The model developed in this study (Chapter 5) is not intended as a complete instructional design model but will focus on collecting and analysing learner data that can be used to improve the instructional design. This study is guided by principles five and six above. While acknowledging the need for identifying appropriate content, the delivery of this content is driven by data produced directly by the learner. A data-driven approach to instructional design will ensure all instructional design choices are based on learner characteristics and goals. While instructional design focuses on the systematic process of preparing engaging learning experiences, learning design provides a way of describing learning activities. A key principle in learning design is the reusability of learning activities. To accomplish reusability and enable communication of the learning design, Conole (2014) proposes the seven Cs of learning design:

(1) *Conceptualise*: Envision the module in terms of the learning objectives and learner characteristics.
(2) *Create*: Identify existing learning materials from Open education Resources or create new materials.
(3) *Communicate*: Facilitate learner-teacher, learner-learner and learner-community communication.
(4) *Collaborate:* Enable collaboration and group work.
(5) *Consider:* Promote demonstration of learning achievement through reflection.
(6) *Combine:* Reflect on the design process.
(7) *Consolidate:* Implement the design and evaluate its effectiveness.

In this thesis, learning design is defined as the practice of creating effective learning opportunities intended to achieve defined learning objectives in a particular context (Mor, Ferguson and Wasson, 2015). Learning design (LD), therefore, focuses on the resources and activities required to accomplish learning.

LD and instructional design (ID) are complementary concepts with a slightly different focus. ID focuses on the overall process and LD focuses on the product, i.e. the preparation of learning materials as part of the instructional design process. Instructional design is the global view from the teacher's perspective, while learning design focuses on the learning process from the learner's perspective.

3.2 Instructional Design Models

A model is a simplified representation of reality that aims to provide some structure and order to a complex process (Richey et al., 2011). The basic framework for any instructional design model should include the following four elements (Figure 3.1): Learner Characteristics, learning objectives, Instructional Strategies and Assessment Procedures (Morrison, Ross, Kemp and Kalman, 2010). Recall from Section 1.3, the proposed solution is aimed at providing differentiated instruction. For differentiated instruction the learning objectives will remain the same for all learners, but the instructional strategies, formative and summative assessments will be tailored based on learner characteristics.

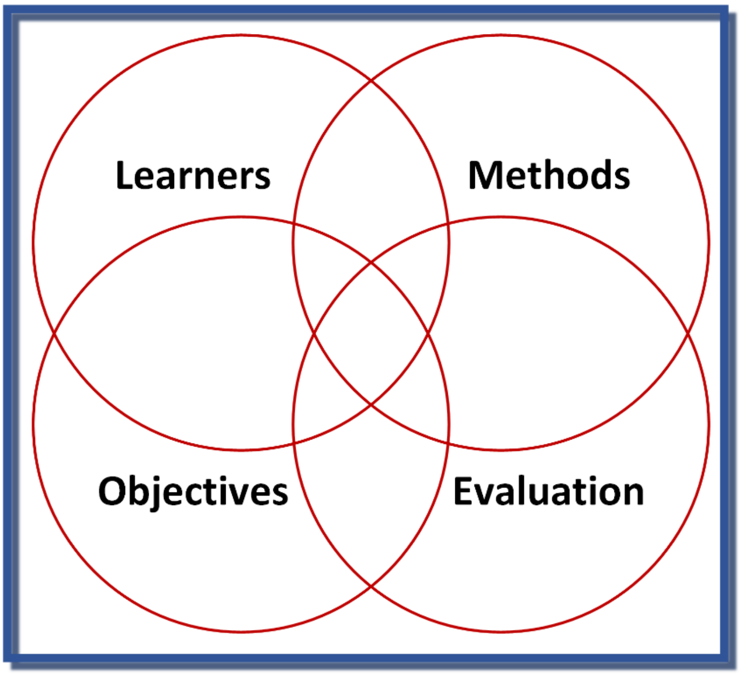


Figure 3.1 Components of Instructional Design (adapted from Morrison et al. (2010))

Instructional Design Models typically add additional components to the four basic elements, each element having some effect on learning outcomes. The methodology adopted for this study (Figure 2.7) requires the solution goals and scope to be evaluated for relevance. This study models an approach to analyse data generated by learners in an online environment, in order to inform instructional design choices. Consequently, procedural (Section 3.2.1) and conceptual (Section 3.2.2) instructional design models are examined next to motivate the relevance of the proposed solution and establish current discourse in the domain in which the solution will be applied.

3.2.1 Procedural Models

Procedural models explain how to perform a complex task (Lee and Jang, 2014; Richey et al., 2011). Instructional design models classified as procedural are derived from either theory or practice or a combination of both. Typically, they are visually represented as a flowchart of sequential steps that guide the creation of instructional interventions or artefacts.

3.2.1.1 ADDIE

ADDIE is an acronym for the five stages proposed to design and develop instruction: Analyse, Design, Develop, Implement and Evaluate.

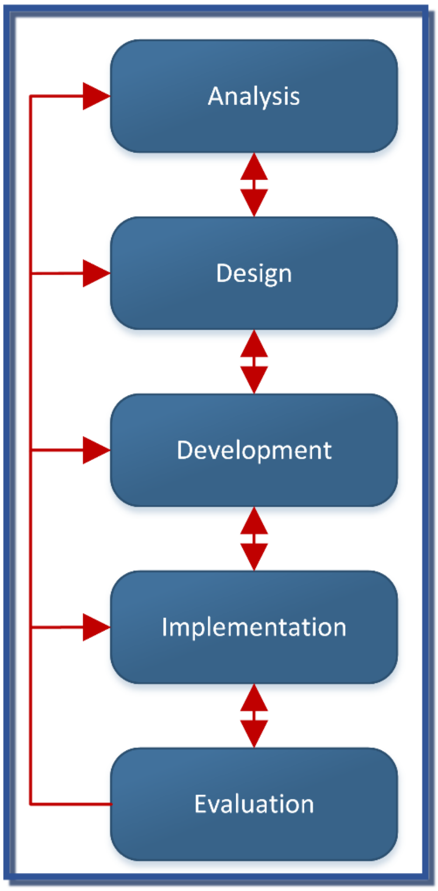


Figure 3.2 ADDIE ID Model (adapted from Molenda (2003))

While ADDIE has been used as the basis for many procedural instructional design models (Gustafson and Branch, 2002), the greatest value of ADDIE is as a conceptual model for instructional systems development (Bichelmeyer, 2005; Molenda, 2003). Even though it is one of the most frequently cited instructional design models, it does not adequately describe the processes used by instructional design practitioners and educators. This can likely be attributed to one of the biggest criticisms levelled at ADDIE: its inability to recognise the flexibility, creativity and ingenuity required in the practice of instructional design (Allen, 2012). In other words, in the definition of instructional design adopted in this thesis[1], ADDIE is good at the "science" part of instructional design but fails to acknowledge the artistic side. The sequential approach

[1] "the science and art of creating detailed specifications for the development, evaluation and maintenance of situations which facilitate learning and performance"

prescribed by ADDIE is also prone to lengthy development cycles, leaving little time for proper testing. The lack of an early prototype poses further communication problems with stakeholders who often cannot visualise the final product (Allen, 2012).

In a search for the original author of ADDIE, Molenda (2003) came to the conclusion that the acronym is merely a colloquial term that describes a systematic approach to instructional systems development. Each stage in ADDIE as implemented by various authors will be described next.

Analysis: The focus of this phase is a performance needs analysis, learner analysis and contextual analysis (Dousay and Logan, 2010). The purpose of the performance analysis is to determine the gap between what the learners know and the competencies they need to exhibit on conclusion of the instruction (Peterson, 2003). The performance analysis should yield educational objectives in the form of learning outcome statements that can be drafted using Bloom's revised taxonomy for learning (Krathwohl, 2002).

The learner analysis requires an investigation into the attributes that influence learning needs. The content should be directed towards the learner achieving the defined outcomes and should be presented in such a way that learners can efficiently navigate through the material. For the contextual analysis, the context in which instruction will take place and the context where the skills will be used warrants attention (Dousay and Logan, 2010). The context will influence the availability of resources during instruction and the performance that is expected from the learners on conclusion of the instruction.

Design: Using the educational objectives derived from the contextual analysis, a task analysis must be constructed outlining steps to achieve each objective. Each task can be expressed as Learning Objective statements that together will guide the learner towards the objective identified during the analysis phase (Dousay and Logan, 2010). Associated with each objective should be an associated learning activity and assessment (Anderson and Krathwohl, 2001). It is during the design phase that the alignment between objectives, activities and assessment is recorded in a task inventory list (Dousay and Logan, 2010).

Development: The development phase is responsible for preparing the instructional strategies and materials. Strategies can be recorded using Gagné's nine events of instruction as a framework (Dousay and Logan, 2010). The materials should cover all resources and activities necessary to implement the strategies determined in the Design phase.

Implementation: The implementation phase puts into practice everything produced during the development phase. Instructional materials are presented according to the stated strategies.

Evaluation: Formative evaluation and summative evaluation are conducted throughout the ADDIE cycle. Formative evaluations are necessary to ensure relevance and consistency of the instruction during the design and development phases (Van Den Akker et al., 2010). Summative evaluations are typically conducted after implementation. In a summative evaluation, the learners' reaction to the instruction and the impact on the performance after the instruction can be assessed. Any feedback from these evaluations must lead to a refinement of the instruction.

3.2.1.2 Dick and Carey

The systematic instructional design model developed by (Dick, Carey and Carey, 2014) was derived from Robert Gagné's "*The Conditions of Learning".* It is based on the premise that learning is complex and controlled by a learner's internal mental processes. The instruction should be developed to support these existing mental models. The instructional design model pragmatically combines techniques from behaviourist, cognitivist and constructivist learning theories with actual instructional practice.

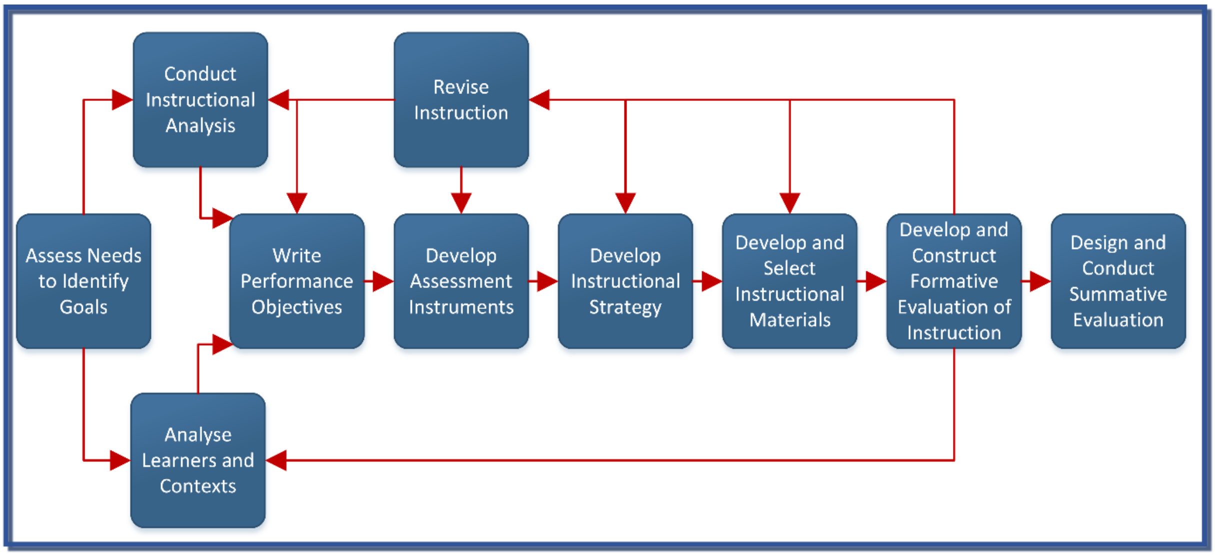


Figure 3.3 Dick and Carey ID Model (adapted from Dick et al. (2014))

The model of Dick et al. (2014) starts with a needs assessment to identify instructional goals. The source of the goals can be from a problem identified with current instruction or from the expectations industry places on new or existing qualifications. After the goals have been identified, each goal must be broken down into steps followed to achieve the goal. So-called “entry behaviours” in the form of knowledge and attributes must be identified to ensure that the prerequisites for the instruction can be published. At the same time learners and their context must be analysed.

The focus of the contextual analysis is twofold, the context in which learners will receive the instruction and the context where they will apply the newly-learnt skills. The learner analysis must focus on the current knowledge and attributes exhibited by the learner. Together the instructional, learners and contextual analysis will feed into the writing of performance objectives. These objectives must describe the exit outcomes that must be demonstrated by the learners. Performance objectives must identify the skills to be learnt and the criteria to assess mastery. Following on from the performance objective statements, assessments instruments must be developed that can evaluate proficiency. Instructional strategies must be devised that can guide the learner towards the stated performance objectives. The theory-backed learning strategy must align with the delivery medium, the content and the learner characteristics. The strategy needs to record:

- Pre-instructional activities
- Presentation of content, practice activities and formative feedback
- Assessment
- Follow-through activities

The instructional strategy will inform the preparation of instructional materials. The material consists of resources that present the content and activities to reinforce the content. It can be created from scratch or constructed from a variety of existing resources. The tentative design of the instruction must be assessed through formative one-on-one, small group or field evaluations. The data collected in the formative evaluation must feed back into the relevant previous steps where shortcomings have been identified. Summative evaluation is conducted once instruction has completed. The final evaluation is meant to determine the impact of the instruction.

3.2.1.3 Gerlach-Ely

The instructional design model of Gerlach and Ely (1971) was developed in response to situations where teachers do not have plenty of time, money and resources to develop instruction.

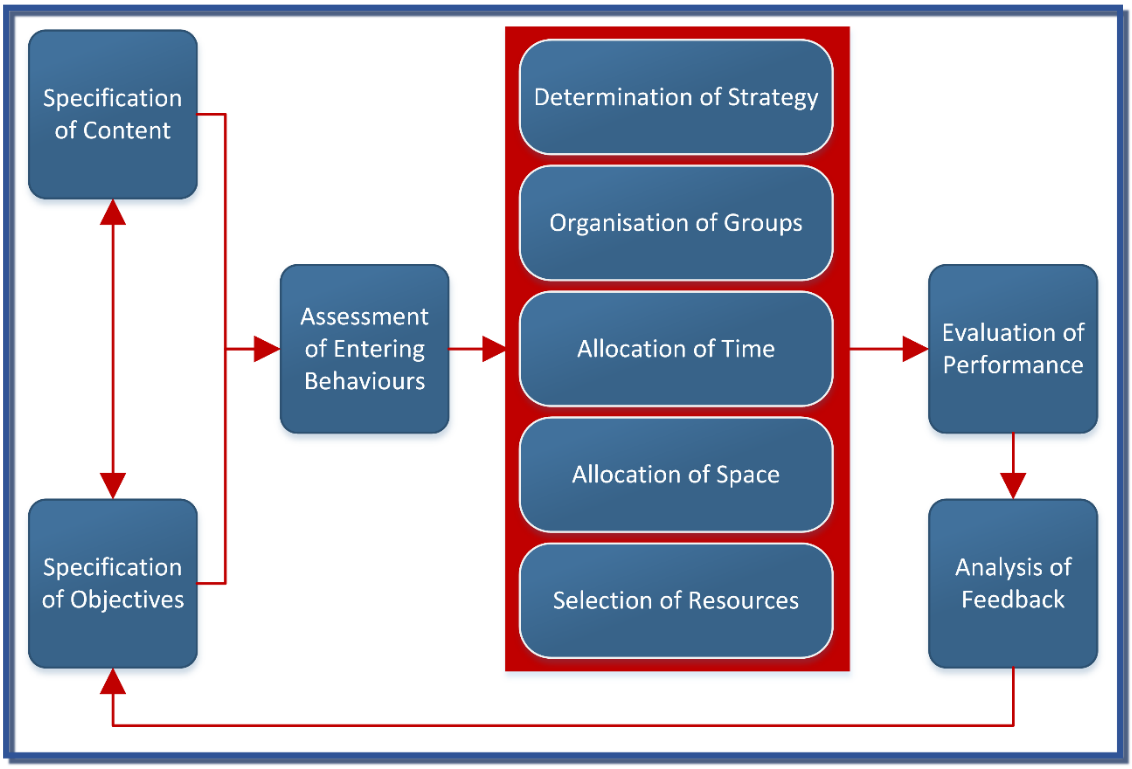


Figure 3.4 Gerlach and Ely ID Model (adapted from Grabowski (2003))

Usually performance objectives are specified before content (Dick et al., 2014). The Gerlach-Ely model, in contrast, acknowledges that this may not always be possible. In some cases, objectives have already been set by a higher body and/or content already specified (Grabowski, 2003). The fact that instructional design can start with specification of objectives or specification of content is depicted by a double-sided arrow between these two steps. Once content and objectives have been specified, the entering behaviours of learners must be assessed. This can include prior knowledge through a pre-test or identification of relevant attributes that affect learning. The next phase is a series of five interconnected steps that can be performed near simultaneously. These steps involve determining a suitable learning strategy, organising learners into groups, allocating the necessary time and space and selecting instructional resources. The theoretically grounded strategies must match the content and learner preferences. The strategy selection must be accompanied by a decision whether the instruction will be self-study, small groups or the entire class. Three pertinent questions can assist in group organisation (Grabowski, 2003):

- Which objectives can learners reach without teacher intervention?
- Which objectives require peer interaction?
- Which objectives require teacher intervention?

The answer to these questions will also affect the decisions regarding time and space. Depending on time constraints under which the teacher operates, the instructional materials can be acquired and modified as needed or created if no suitable materials are found. In practice, though, most teachers follow the approach of modifying existing materials due to time constraints in the classroom. Assessment of learner performance achievement and their attitude towards the instruction are conducted after the instruction. This feedback must be analysed and worked back into the instruction.

3.2.1.4 ASSURE

The ASSURE model (Heinich, Molenda, Russell and Smaldino, 2002) is a procedural instructional design model that incorporates Gagné's events of instruction (Gagné, 1965). ASSURE is an acronym for the six steps a teacher can follow to prepare a technology enhanced lesson plan.

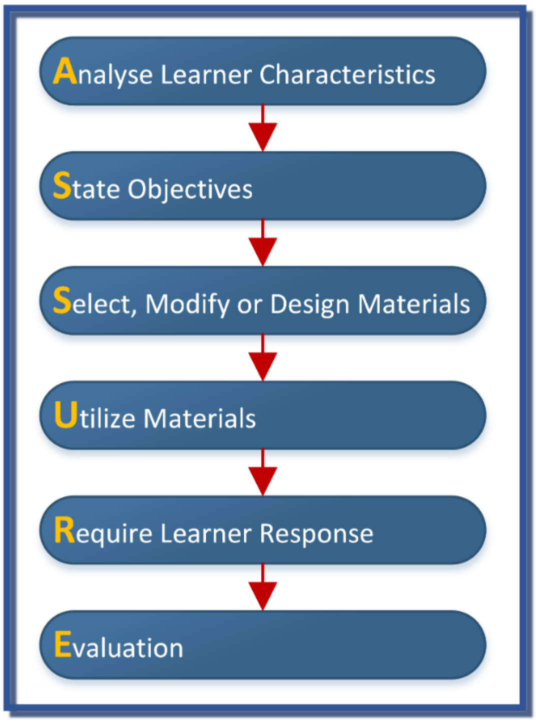


Figure 3.5 ASSURE ID Model (adapted from Heinich et al. (2002))

The process starts with a learner analysis. This involves determining learner characteristics, prior knowledge expected of the learners and their preferred learning styles. After the learner analysis, the teacher should state the objectives using the ABCD format – specify the *audience*, *behaviour* to be demonstrated on completion of the instruction, *conditions* for performance assessment and the *degree* to which the performance will be deemed acceptable. Next, the methods, media and materials must be selected and utilised.

The teacher has three options: (1) select available resources, (2) modify existing resources, or (3) create materials from scratch. Before providing learners with the materials, they must be previewed by the teacher and prepared for the lesson. The teacher must also prepare the environment and set the learner expectations. When providing the learning experience, learner engagement is advised by providing opportunities for practice and feedback. After the lesson, learner achievement must be measured, and the instruction evaluated. The feedback should be incorporated in the planning of future lessons.

3.2.1.5 Backward Design

Backward Design is a three-stage approach to planning instruction (Wiggins and McTighe, 2005). The basic premise in backward design is that learning experiences should be planned with the end in mind.

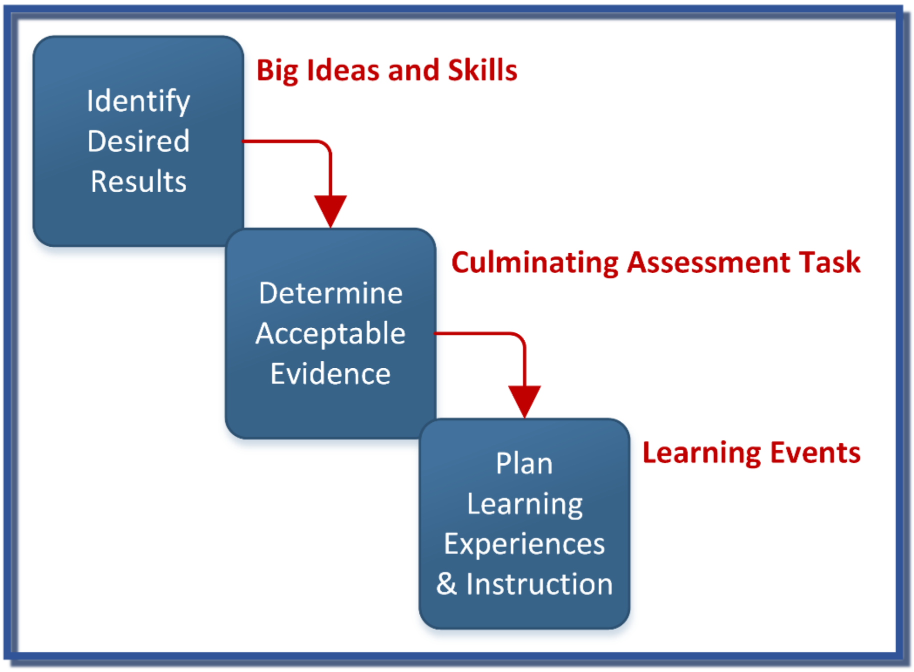


Figure 3.6 Backward Design (adapted from Wiggins and McTighe (2005))

The first step in this instructional design process is the identification of the goals in the form of content standard, program objectives or learning outcomes. The desired results should be prioritised on three levels:

- Big ideas that the learners need to understand
- Knowledge and skills that are important to possess
- Knowledge that is worth being familiar with

Once the goals have been stated, the focus should move to assessment and the setting of criteria for establishing whether the required performance is achieved. Authentic assessment opportunities must be created to collect evidence of performance achievements. The type of assessment must match the priority level assigned to the desired results. Students must also be given self-assessment and reflection opportunities.

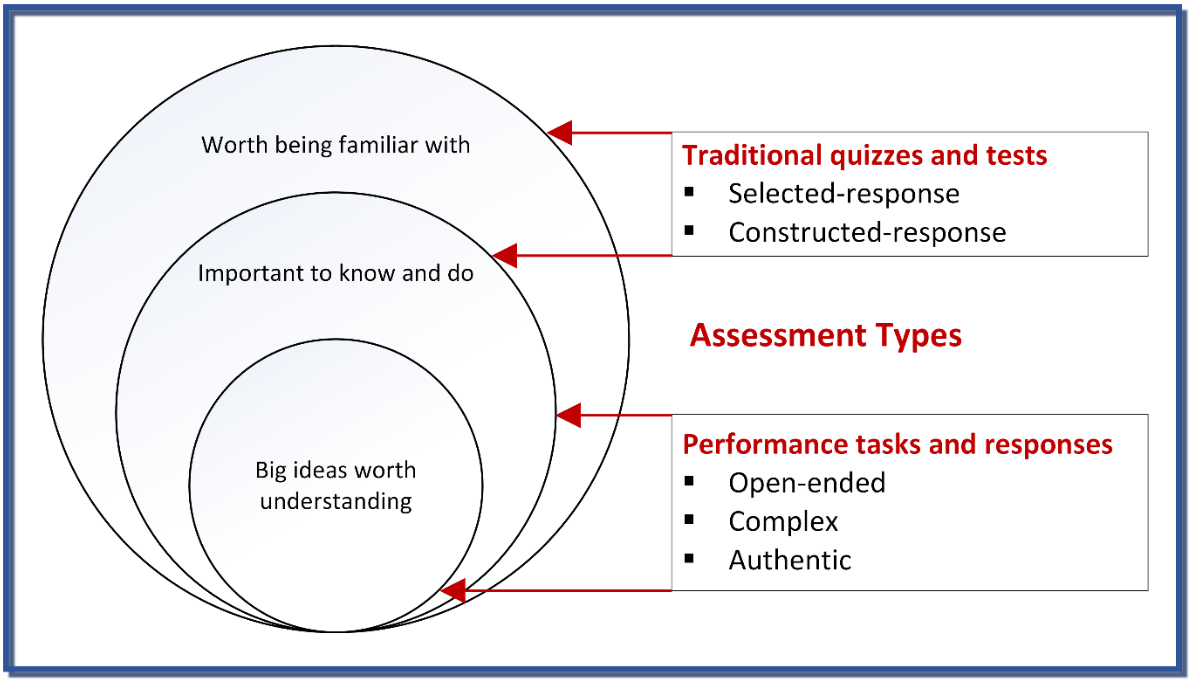


Figure 3.7 Assessment in Backward Design (adapted from Wiggins and McTighe (2005))

After the necessary assessments and rubric have been specified, a learning plan must be created that aligns with the desired result and assessment opportunities. Wiggins and McTighe (2005) propose the acronym WHERETO to help guide the preparation of the lesson plan.

The designed lesson must include:

- **W**here the unit is going, **w**hat is expected and **w**hat the learners' prior knowledge and interests are
- **H**ooks to grab the learners' attention and ways to **h**old their interest
- **E**xploration of issues towards the key ideas
- **R**eflection and **r**evision opportunities
- **E**valuation opportunities for learners on their own work
- **T**ailored instruction based on the diverse needs, interests and abilities of learners
- **O**rganising principles to enhance initial and ongoing engagement

3.2.1.6 SAM

The SAM instructional design model is an acronym for Successive Approximation Model (Allen, 2012). SAM is an agile model for rapidly building quality learning experiences. SAM can be scaled to fit the needs of the instructional designer and the project. It can go through a rapid cycle of analysis, design and development with an early prototype that can be released to stakeholders and gradually improved over time.

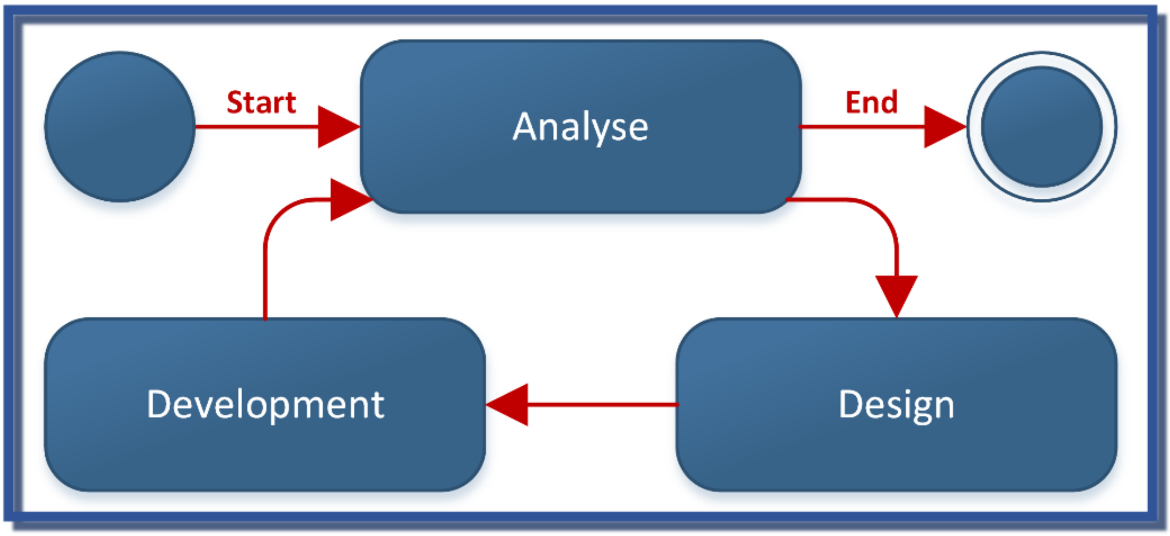


Figure 3.8 Successive Approximation ID Model (adapted from Allen (2012))

For bigger, more complex projects, the extended version of SAM can be followed (Allen, 2012). This involves a preparation phase, iterative design phase and iterative development phase.

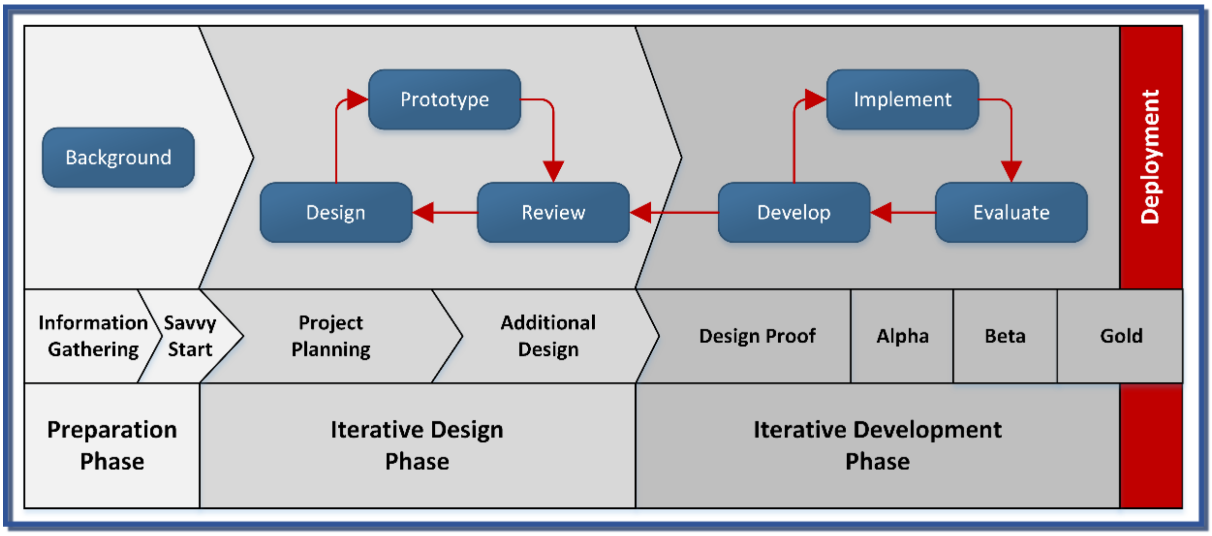


Figure 3.9 Extended Version of SAM (adapted from Allen (2012))

The extended version of SAM starts with a brainstorming workshop with a group of relevant stakeholders where job performance requirements are debated. The output from the preparation phase is used to sketch out ideas and draft a prototype. The prototype is reviewed by the stakeholders and feedback worked back into the design. When the design is approved, the components are developed and implemented in the desired context where authentic evaluations are performed. The development phase rapidly progresses from design through to an alpha and beta prototype before the stable release is deployed.

3.2.2 Conceptual Models

Conceptual models are a class of instructional design models that describe a particular abstract view of reality (Richey et al., 2011). Examples of conceptual models include taxonomies like Kemp's model (Morrison et al., 2010), the 3P model of Freeth and Reeves (2004), the conceptual framework of high quality learning environments (Shea, Pickett and Pelz, 2003) and heuristics as synthesised by Mehlenbacher (2012). Conceptual models, frameworks or heuristic guidelines are not bound by the sequential nature of procedural models. They provide factors that instructional designers must consider when creating effective online instruction.

3.2.2.1 Kemp's Model

Kemp's ID model presents nine core interrelated elements influencing learning outcomes (Morrison et al., 2010). The model does not prescribe a clear starting point, which provides a degree of flexibility to instructional designers. This non-linear structure also implies that some aspects can be addressed simultaneously. Initially it is advised that an instruction designer starts with an instructional problem and moves clockwise through the cycle. Adjustments to the course can then be made to any element over time. The nine elements that need to be considered by instructional designers include (Figure 3.10):

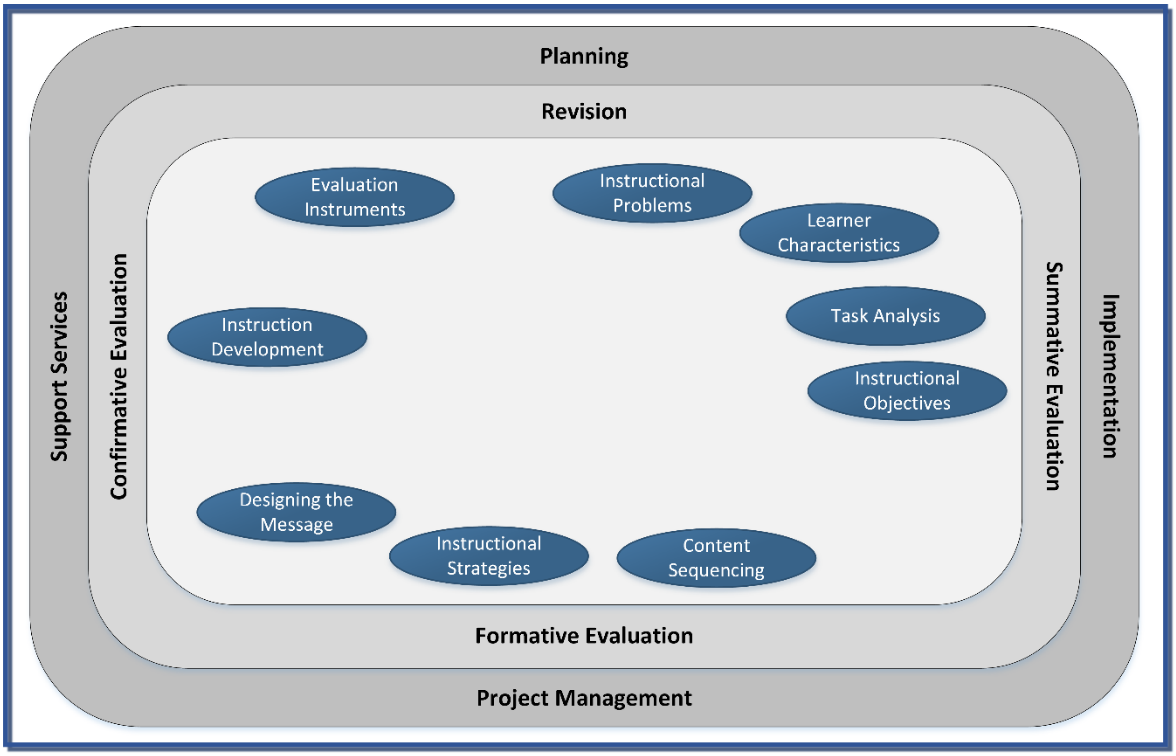


Figure 3.10 Kemp Instructional Design Model (adapted from Morrison et al. (2010))

- Instructional Problems: The focus of this element is on the identification of instructional programme goals by means of a needs assessment, i.e. what does the student need to be able to do at the end of the training?
- Learner Characteristics: This element addresses learner needs in terms of their attributes that influence their learning.
- Content and Task Analysis: This element focuses on the subject content and task components related to the goals identified.
- Instructional Objectives: This element is concerned with learning objectives that lead to the overall instructional program goal.
- Content Sequencing: The Learning Objects must be ordered logically within the units of learning.
- Instructional Strategies: This element prescribes that the instructional designer develop strategies to aid recall and knowledge integration.

- Designing the Message: Supplementary resources needed to support the instruction must be developed.
- Instructional Delivery: For this element the instructional designer must develop and deliver the instruction.
- Evaluation Instruments: Formative and summative evaluations of achievements as well as behaviours and attitudes must be conducted

3.2.2.2 3P Model

The conceptual model of Biggs (2011) presents factors that affect learning on a timeline (Figure 3.11), starting with before learning (presage), during learning (process) and the outcome after learning (product).

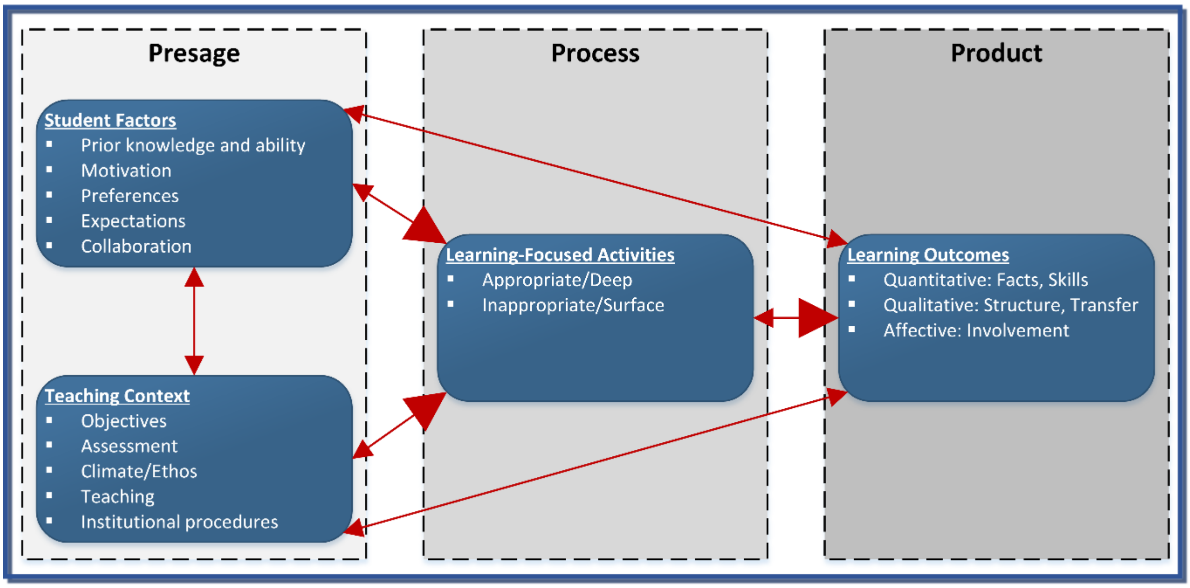


Figure 3.11 3P Instructional Design Model (adapted from Biggs (2011))

Originally, the presage phase in the instruction development was meant to focus on the learner and the teaching context. Others prefer to divide teaching context into two factors, teacher characteristics and context (Freeth and Reeves, 2004). Student factors such as prior knowledge, ability, preferred approach to learning, collaboration, expectations and motivation all have an influence on the way learners approach course activities. The educators' teaching philosophy, expertise, perceptions of their learners and enthusiasm, combined with contextual factors such as learner numbers, time constraints, competing curricular demands and the political climate, also have an influence on the learners' approach to learning. In turn, the way learners approach instruction will determine the outcome. None of these contextual factors are to be seen in isolation, but as part of a system where every factor has some level of influence on the other factors throughout the presage, process and product points in time. In Figure 3.11, the arrows in between the factors denote this influence, with the larger arrow heads indicating the general direction of the influences. The arrows pointing in all directions point to the interconnected nature of all factors that influence instruction.

3.2.2.3 Conceptual Framework of High-Quality Online Learning Environments

Shea et al. (2003) proposed a framework consisting of multiple perspectives in the development of online learning environments:

- Learner
- Knowledge
- Assessment
- Community

These perspectives are represented as a Venn diagram of overlapping and integrated elements within the broader community. These elements together affect the learners' experience of the online learning environment. At its core is the notion that any learning environment (online or face-to-face) will be considered effective if it is learner centred, knowledge centred, assessment centred and community centred. To create a knowledge centred learning environment, we need to carefully consider the foundational knowledge, skills and attitudes needed in prospective graduates.

A learner centred environment accounts for relevant learner attributes, knowledge and preconceptions. A learning environment is deemed assessment centred if learners are able to provide the necessary evidence of their progress and if regular feedback is given on this progress. Finally, a community centred learning environment encourages collaboration and produces lifelong learners who contribute successfully to society at large. These integrated elements correspond to overlapping lenses through which a learning environment can be examined (Figure 3.12, Table 3.1).

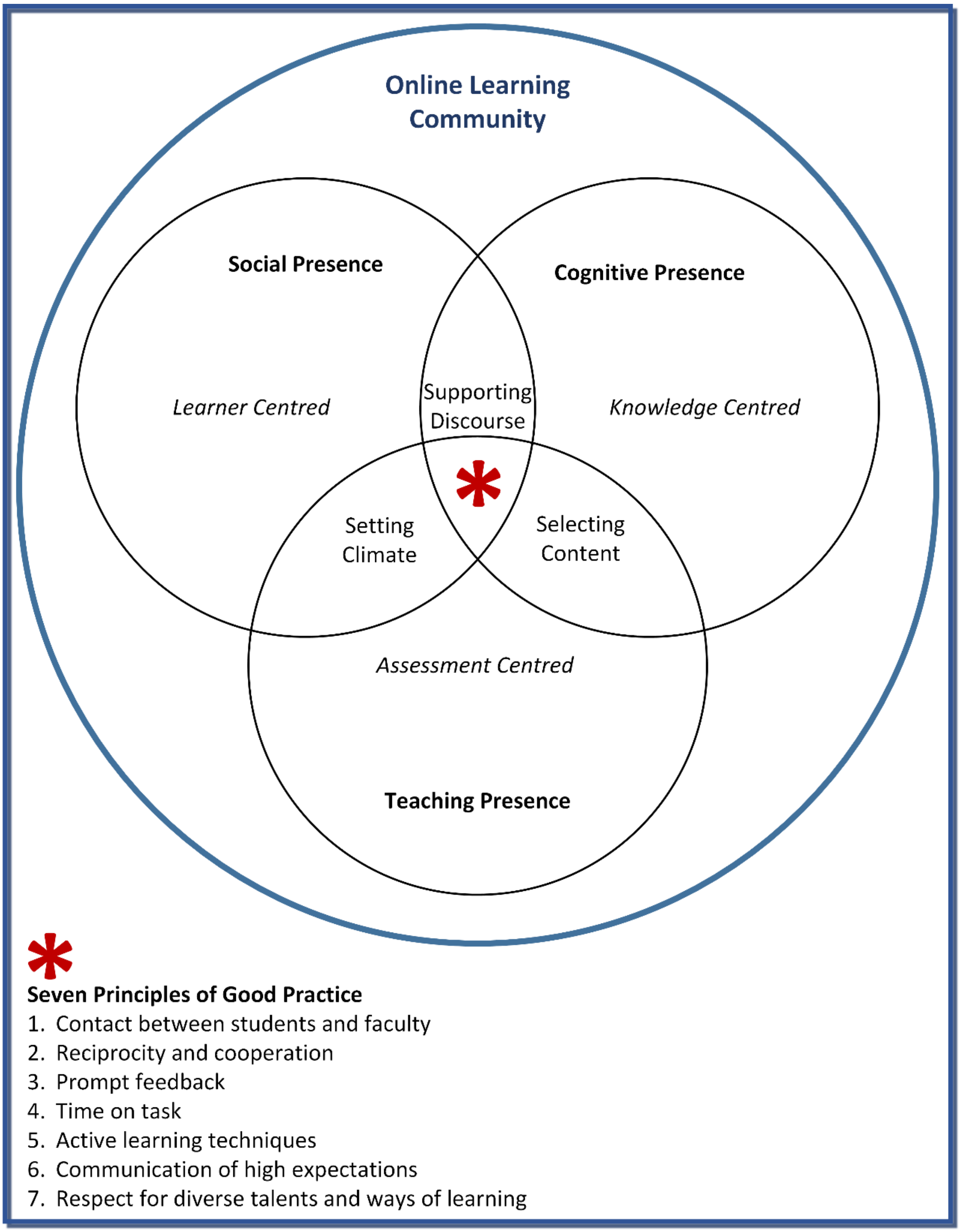


Figure 3.12 Conceptual Framework for High Quality Online Learning Environment (adapted from Shea et al. (2003))

The overlapping circles are indicative of the influence of each focus area on the other elements. For example, the ability of learners to immerse themselves socially in a community of enquiry (social presence) is influenced by the supporting discourse and contextual climate. The ability of learners to master the discourse within the community of enquiry is influenced by the supporting discourse and content provided by the teacher. Finally, the ability of educators to effectively design and develop socially meaningful learning experiences are influenced by the contextual climate and the content available to present to learners.

Table 3.1 Lenses to Examine an Online Learning Environment

Lenses	Description	Interfaces
Social Presence	Ability of learners to socially immerse themselves in a community of inquiry	• Learner-Knowledge (Supporting Discourse) • Learner-Assessment (Setting Climate)
Cognitive Presence	Ability of learners to master the discourse of the community of inquiry	• Knowledge-Learner (Supporting Discourse) • Knowledge-Assessment (Selecting Content)
Teaching Presence	Ability of educators to design and develop socially meaningful learning experiences	• Assessment-Knowledge (Selecting Content) • Assessment-Learner (Setting Climate)

The conceptual framework further proposes seven principles of good practice to facilitate student success in higher education. These include:

- Contact between students and faculty
- Reciprocity and cooperation
- Prompt feedback
- Time on task
- Active learning techniques
- Communication of high expectations
- Respect for diverse talents and ways of learning

These principles are best practice guidelines on the best way to facilitate success in higher education. They form the centre of the framework where learner, knowledge, assessment and community overlap.

3.2.2.4 Mehlenbacher's Instructional Design Principles

Based on a review of instructional design research papers, (Mehlenbacher, 2012) synthesised design guidelines organised by five dimensions of technology enhanced learning situations. These dimensions include:

- Learner Background and Knowledge
- Learner Tasks and Activities
- Social Dynamics
- Instructor Activities
- Learning Environment and Artefacts

While these dimensions are not represented diagrammatically, they have the same function as Kemp's nine elements (Morrison et al., 2010) and the factors in Biggs's 3P model (Biggs, 2011) highlighting conceptual factors that influence learning design and ultimately learners' experience of the instruction. Mehlenbacher (2012) provides the guidelines in a table along with citations to the study that produced the guidelines. He further advises that these guidelines are not meant to be implemented in a particular design, without first studying the context from which the principle was produced. He concedes that the guidelines run the gamut of being too general or too specific to apply to new situations. The usefulness of these design principles, however, is that they provide a glimpse into interventions that worked in particular contexts. The principles also act as a useful resource to guide continual evaluation of particular instructional artefacts with the view to refine these solutions deployed in a particular context.

3.2.2.5 Kirkpatrick Model for Evaluating Training Effectiveness

The Kirkpatrick model (Figure 3.13) is used for evaluating the effectiveness of training on four levels (Kirkpatrick, 1996). The first level measures how learners responded to the training. The second level evaluates learners' understanding of the content. The third level measures whether learners are applying their new knowledge and skills. The fourth level determines whether learners' change in behaviour had an impact on their place of employment.

Figure 3.13 Kirkpatrick Model for Evaluating Training Effectiveness (adapted from Kirkpatrick (1996))

The questions asked for a level one evaluation should focus on learner satisfaction and how they experienced the instruction. The feedback is used to improve the instruction for future use. The course evaluation is mostly done immediately on completion, but spot checks can be built in throughout instruction.

The level one evaluation can be formal or informal. Level two evaluations focus on the learners' knowledge development. Typically, improvement is measured through a pre- and post-assessment. To measure impact of training, the treatment group that received instruction in a specific way can be measured against the achievement of a control group that received instruction in the usual way. A wide range of formal and informal techniques exists to measure achievement during and after the instruction. Level three evaluations should preferably be conducted between three to six months after the instruction. For level three, the focus is on trying to determine whether the learner is applying what he or she has learnt from the previous instruction. Surveys, observations and interviews are standard techniques for performing this evaluation. Multiple techniques are necessary to ensure valid results. The final level four shifts the focus from the learner to the organisation where knowledge and skills are applied.

Performance metrics should be identified in order to determine what impact the instruction had on the business goals. For example, these metrics could include factors such as reduced costs, higher quality products/services, fewer incidents and efficient production times. This level is highly complex, since there may be compounding variables outside the scope of the instruction that played a role in any performance improvement.

3.2.3 Reflection on Instructional Design Models

Most procedural instructional design models are variants of the ADDIE model that describes a systematic approach to development of instructional systems. While different terms are used to describe the steps, they can all be mapped to the Analysis, Design, Develop, Implement and Evaluation phases of ADDIE (Table 3.2).

Table 3.2 Relevance of Procedural Instructional Design Models to This Study

ADDIE	Dick and Carey	Gerlach-Ely	ASSURE	Backward	SAM
Analyse	• Needs Assessment • Instructional Analysis • **Learners* and Contexts**	• **Entering Behaviours***	• **Learner Attributes***	• Desired Results • Acceptable Evidence	• **Information Gathering*** • Savvy Start
Design	• Performance Objectives • Instructional Strategy	• Objectives • Instructional Strategy	• Objectives	• Instruction **(WHERETO)***	• Project Planning • Design
Develop	• Assessment Instruments • Instructional Materials	• **Groups*** • Time and Space • Content • Resources	• Learning Materials		• Design Proof • Alpha • Beta • Gold
Implement			• Utilise Materials		• Deployment
Evaluate	• Formative and Summative Evaluation	• Performance	• **Learner Response*** • Evaluation		• Iterative Design and Development Phases

***All steps highlighted in bold font refer to the need for learner analysis as part of instructional design.**

Most procedural instructional design models are cyclical in nature, with results from evaluations feeding back into the optimisation of instruction. The **SAM model** (Section 3.2.1.6), in particular, emphasises the need for stakeholder input throughout the instruction design and development. The Analysis phase of the **ADDIE model** (Section 3.2.1.1) suggests learning can be optimised by analysing learners' knowledge, beliefs, preferences and values and the context in which learning occurs prior to instruction. This is echoed in the **ASSURE model** (Section 3.2.1.4), that also recommends analysis of learner characteristics, prior knowledge and learning styles prior to instruction. The learning environment must, therefore, be continuously optimised through analysis before instruction and evaluations during or after instruction.

According to the **Dick and Carey model** (Section 3.2.1.2), instruction must be developed to support learners' mental models. In the "Analyse learners and contexts" step, during which so-called "entry behaviours" are identified prior to instruction. Learning strategies must align the delivery medium, content and learner attributes.

Assessment of entering behaviours is also recommended by the **Gerlach-Ely model** (Section 3.2.1.3). These behaviours must feed into the teaching strategy, allocation of space and time and selection of resources. The Gerlach-Ely model also recommends organising learners into groups based on their entering behaviours. Groups are organised based on the learners' need for teacher intervention and also around peer interaction.

Backward Design (Section 3.2.1.5) recommends aligning different types of assessments with three levels of outcomes. Relevant to this study is the recommendation that the lesson plan must tailor instruction based on the diverse needs, interests and abilities of learners. The lesson should be constructed in a way that ensures learner attention is focused and engagement sustained throughout.

The elements proposed in conceptual instructional design models can also be mapped onto the ADDIE phases (Table 3.3).

Table 3.3 Relevance of Conceptual Instructional Design Models to This Study

ADDIE	Kemp	3P	High Quality Online Learning	Mehlenbacher	Kirkpatrick
Analyse	• Instructional Problems • **Learner Attributes*** • Task Analysis • Instructional Objectives	• **Student Factors*** • Teaching Context	• **Social*** • **Cognitive*** • Teaching	• **Learner Background*** • **Social Dynamics*** • Learner Tasks • Instructor Activities • Learning Environment	
Design	• Message • Instructional Strategy • Sequence	• Learning Activities	• Setting Climate • Selecting Content		
Develop	• Instruction • Evaluation Instruments		• **Supporting Discourse***		
Implement	• Implement				
Evaluate	• Formative, Summative, Confirmative Evaluation	• Learning Outcomes	• Best Practices **(7. Respect for diverse talents and ways of learning) ***		• Levels to measure teaching effectiveness

***All steps highlighted in bold font refer to the need for learner analysis as part of instructional design.**

Kemp's model (Section 3.2.2.1) explicitly includes identifying relevant learner characteristics as one of the elements that require attention during the analysis phase of instructional design.

Bigg's 3P model (Section 3.2.2.2) emphasises the interconnection between student factors, their approach to learning activities and learning outcomes. Some of the learner factors highlighted include prior knowledge and ability, motivation, preferences and expectations. The 3P model asserts that once identified, the learner factors should be considered when creating learning activities. Factors such as motivation have an influence on how engaged learners will be with the course material. In turn, the outcomes and learning activities impact learners' motivation, expectations and abilities.

The **conceptual framework for high quality online learning environments** (Section 3.2.2.3) identifies "respect for diverse talents and ways of learning" as one of the best practices in online instructional design and development. Affordances must be developed into the online environment to support learners' social and cognitive development. Learner background knowledge and social dynamics are two of the five dimensions included in **Mehlenbacher's learning design heuristics** (Section 3.2.2.4) for their influence on learning design. The tasks, activities and learning environment should consider the individual learners and their preferences towards peer interaction.

The **Kirkpatrick model** (Section 3.2.2.5) focuses on evaluating the effectiveness of the training. While this study focuses on the tailoring of instructional delivery, the Kirkpatrick model implies a need for measuring the impact of the instruction on completion.

The instructional design models are relevant to the traditional classroom environment and to online learning. This study is concerned with the online learning environment that imposes unique needs in terms of learning design (Section 3.3). Delivery of online learning through Learning Management Systems (Section 3.4) is increasingly becoming the status quo. Consequently, online learning design and Learning Management Systems, specifically Moodle, are discussed in the rest of this chapter.

3.3 Online Learning Design

Learning design is the creative and deliberate act of preparing resources and activities aimed at guiding learners to specified learning objectives in a particular context (Mor, Craft and Hernández-Leo, 2013). Online learning design choices are informed by subject knowledge, pedagogical considerations, technology and classroom practice.

As can be seen from Section 3.2.1, all procedural instructional design models include, at some point, the selection, design or development of instructional materials. Instructional materials are resources that organise and support learning. In this thesis, a technique is proposed to differentiate Learning Objects in Moodle (Section 5.4). Learning objects are the building blocks of instructional materials that deliver the content in an engaging manner (Section 3.3.1). Differentiated Instruction involves tailoring these Learning Objects to the unique needs of the learners who will interact with the resources (Section 3.3.2).

3.3.1 Learning Objects

The concept of a Learning Object (LO) emerged as a means to reduce the cost of the development time of digital educational content (Wiley, 2002). While no single universally accepted definition exists, there is some consensus of the intention of LOs and some of its features. A working definition applicable to this thesis and an overview of general LO features are described in Section 3.3.1.1. In order to link LOs to instructional design theories, it is necessary to describe different types and size of Learning Objects (Section 3.3.1.2). To enable reusability of LOs and to make differentiated instruction possible, metadata that describe Learning Objects must be addressed (Section 3.3.1.3).

3.3.1.1 LO Definition and Characteristics

The term Learning Object is attributed to Wayne Hodgins, through his efforts as chair of a working group with the vision to establish interoperable standards for technology enhanced learning (Hodgins, 2006; Ritzhaupt, 2010). Hodgins (2006) initially conceptualised Learning Objects as being analogous to LEGO blocks that can be assembled in different ways, meeting the unique needs of those who use it. In an attempt to create a richer analogy to describe Learning Objects, Hodgins (2006) borrowed concepts from the construction industry. In the same way that a building is assembled from premanufactured components, following industry standards, Learning Objects developed according to industry standards can be reused in different contexts.

Wiley (2002) suggests atoms as an alternative metaphor to Hodgins's simplistic LEGO analogy. Unlike LEGO blocks, the atom cannot combine with every other atom and specialised training is required to join atoms together in a useful manner. In the same way, for a Learning Object to be educationally useful, it must be developed and combined in a meaningful way based on sound instructional design theories and pedagogical principles. Using the atom analogy, educational expertise is emphasised in the selection or development, presentation and sequencing of Learning Objects.

In contrast to the broad definition of the Learning Technology Standards Committee of the IEEE Computer Society that expresses Learning Objects as reusable digital or non-digital entities that can be used for technology enhanced learning (IEEE 1484.12.1, 2002a), Wiley's (2002) definition focuses on only reusable digital learning resources. This thesis also adopts the Wiley definition, with the emphasis on the resources being digital and reusable.

Reusability, the ability of a Learning Object to be applied in different contexts, is one of the fundamental characteristics of Learning Objects. Çinici and Altun (2018) argue for LOs to be reusable and purposeful, they should be developed according to a suitable pedagogic model. Reusability will be improved if the Learning Object is developed at an appropriate level of granularity and appropriately tagged with pedagogic metadata. A Learning Object will be **interoperable** if it is independent of delivery mechanism. In other words, a Learning Object must be usable on any device, platform, operating system, hardware, or browser (Gürer, 2013). Furthermore, a Learning Object is said to be **durable** if it is unaffected by any hardware upgrade or software update. Reusability is achieved through standardised metadata support (IEEE 1484.12.1, 2002a). Metadata makes the Learning Object **discoverable** and the use of central repositories makes Learning Objects **accessible**.

Granularity, or size, is another fundamental characteristic of Learning Objects. The granularity of the Learning Object affects how reusable it can be in different contexts (Gürer, 2013). The size of a Learning Object is affected by its content or functionality. The more building blocks are aggregated, the less reusable the Learning Object becomes (Hodgins, 2006). Smaller Learning Objects are more reusable, since they are more context independent (Kramer, 2005). Granularity is achieved because of the **modularity** of Learning Objects. Modular building blocks can be combined to create bigger building blocks. Granularity further enhances the **manageability** of Learning Objects – i.e. they can be updated, revised or aggregated based on learner needs. The modular nature of Learning Objects enables **adaptability** – the ability to customise learning – and **generativity** – the ability to aggregate Learning Objects automatically based on learner needs.

Linden and Lederman (2015) derived a set of principles to guide the design of building blocks for Learning Objects:

- Building blocks should be generic and context independent to ensure they are reusable
- Building blocks should be independent of other building blocks, so that a change in one has little impact on others they are coupled with
- It should be possible to aggregate multiple building blocks to create more complex Learning Objects
- It should be possible to convert building blocks between different modalities

Following these principles, the creation of building blocks will improve the reusability of Learning Objects, thereby reducing the time and financial implications of developing online instructional design (Linden and Lederman, 2015). To further enhance reusability, it is important to aggregate the building blocks at an appropriate level. Section 3.3.1.2 describes various taxonomies for Learning Object types and granularities.

3.3.1.2 Learning Object Types and Granularity

The granularity of a Learning Object refers to the size of the elements aggregated together. Several authors have described different levels of Learning Objects. Wiley (2002) describes five different Learning Object types, based on various characteristics:

- Number of media elements aggregated together to form the new Learning Object
- Type of objects combined into the new Learning Object
- Whether the components of the new Learning Object are reusable in new contexts
- The usual way in which the new Learning Object is customarily used
- Reliance of the new Learning Object on information about other Learning Objects

- The purpose of algorithms and procedures within the new Learning Object
- The potential of the new Learning Object to be reusable in other contexts
- The potential of the new Learning Object to be reused in the same problem domain

Based on these characteristics, Wiley (2002) defines five Learning Object types (Table 3.4);

- Fundamental – A single, basic building block that presents a fact or an example, e.g. a graphic or text.
- Combined-closed – A limited number of digital resources aggregated at design time for instruction *or* practice. Individual constituent components are difficult or impossible to reuse, e.g. a video combined with audio.
- Combined-open – A larger number of digital resources aggregated in real time for instruction *and* practice. Individual constituent components are easier to discover and reuse, e.g. graphic, video and text combined in a single web page.
- Generative-presentation – Digital resources must be automatically aggregated for instruction, practice and assessment. These Learning Objects are reusable in similar contexts, but not so much in contexts not originally intended, e.g. animations that can test a learners' problem-solving ability.
- Generative-instructional – Digital resources must be automatically aggregated for instruction, practice and assessment. These highly reusable Learning Objects must infer learning strategies from learner interactions and choose a suitable instructional strategy in real-time.

Table 3.4 Taxonomy of LO Types

LO Types → / LO Characteristics ↓	Fundamental LO	Combined-Closed LO	Combined-Open LO	Generative-Presentation LO	Generative-Instructional LO
No of elements combined	One	Few	Many	Few-Many	Few-Many
Type of objects contained	Single	Single, Combined-Closed	All	Single, Combined-Closed	Single, Combined-Closed, Generative-Presentation
Reusable component objects	N/A	No	Yes	Yes / No	Yes / No
Common function	Exhibit, Display	Pre-designed instruction/practice	Pre-designed instruction and practice	Exhibit, Display	Computer-generated instruction/practice
Extra-object dependence	No	No	Yes	Yes / No	Yes
Type of logic in object	N/A	None / answer sheet scoring	None / Domain-specific instruction and assessment	Domain-specific presentation	Domain-independent instr. and assessment
Potential for inter-contextual reuse	High	Medium	Low	High	High
Potential for intra-contextual reuse	Low	Low	Medium	High	High

Hodgins (2006) describes a five-level taxonomy for Learning Objects;

- Media Elements – The smallest element or content asset presented in some modality, e.g. Audio, Text, Illustration, Animation, Simulations
- Information Objects – A collection of media elements to create a media independent block of information, e.g. Procedure, Principle, Concept, Process, Fact, Overview, Summary
- Application Objects (Learning Objects) – Information Objects are combined towards a single objective.
- Aggregate Assemblies – Collection of Learning Objects combined in a lesson linked to a module outcome, e.g. Lessons, Chapters, Units of Learning
- Collections – Represents whole curricula of Courses.

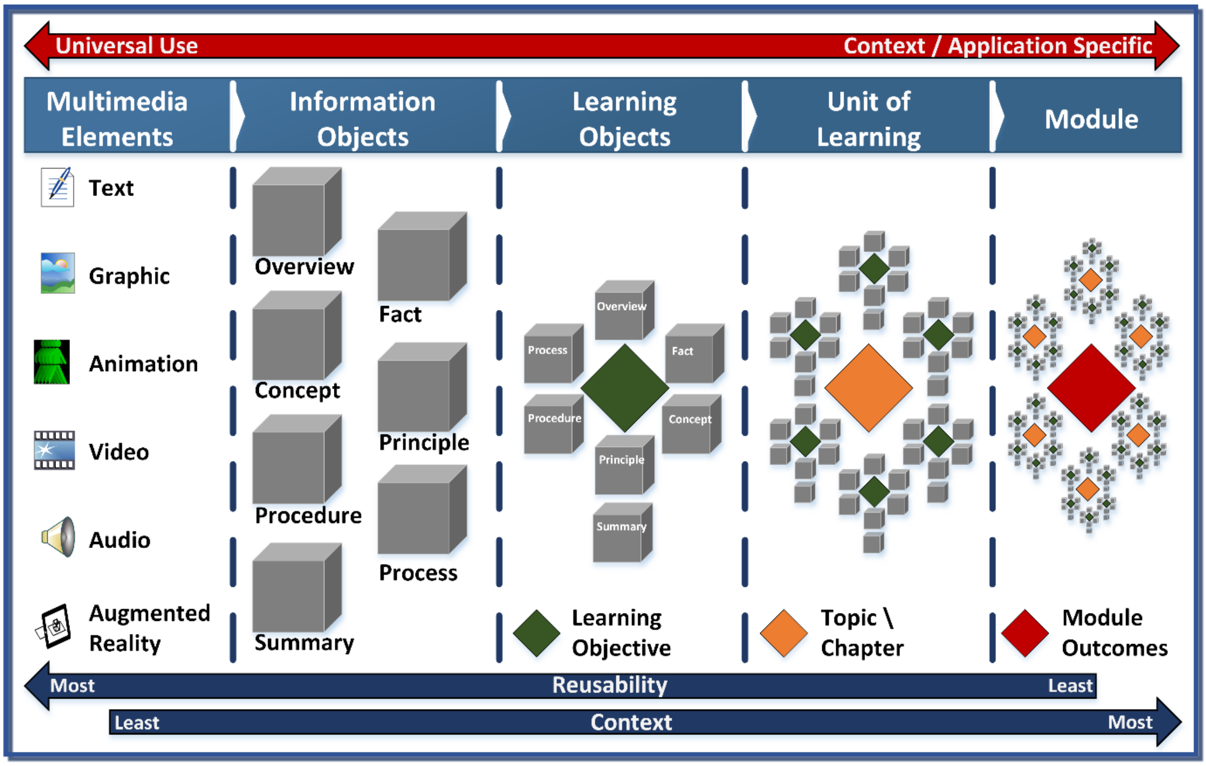


Figure 3.14 Content Object Taxonomy (adapted from Hodgins (2006))

CISCO Systems (2001) delivers training courses using a strategy of database-driven Reusable Learning Objects (RLOs) discovered, reused and repurposed. Their Reusable Information Objects (RIOs) are classified according to a hierarchy of seven information types on the same level as Hodgins' (2006) Information Objects. Two are specialised RIOs at the start of the RLO (Overview) and at the end of the RLO (Summary):

Overview RIO is an advanced organiser used to introduce an RLO:

- Introduction – explain the purpose of the RLO
- Importance – relate the RLO objective to job functions
- Objectives – the skills that should be mastered through the lesson
- Prerequisites – the skills needed at the start of the lesson
- Scenario (optional) – example from practice of how skills may be applied
- Outline – the title of each RIO

Summary RIO concludes the RLO before assessment:

- Review – recap of what was learnt
- Next steps (optional) – suggestions of other RLOs that could be completed next
- Additional resources (optional) – resources beyond the RLO with extra relevant information

Five content RIOs present the information and can be of the form concept, fact, procedure, process or principle. Each content RIO has an introduction and optional instructor notes. Introduction sections explain the purpose of the RIO and instructor notes provide optional guidance to instructors not involved with designing the Learning Objects. The purpose and structure of the other RIOs are:

Concept RIOs used to transfer information about a group of ideas, symbols or objects. Concept RIOs are identified by statements like "What is a…" or "What are the types of…". Concept RIOs include:

- Definition – describe the concept and related characteristics through text, graphic or animation
- Example\Analogy – describe an instance of the concept in various contexts and cross reference with the background of the learner

Fact RIOs are used to teach unique background about a concept. Fact RIOs are presented as statements, data or graphics of specific objects. Facts can be in the format of graphics, lists or tables.

Procedure RIOs are used to teach the sequential steps to perform a specific task. Procedures are identified by instructions like "How to… ", "Configure… ", "…Operate", etc. Steps can either be presented in a table with "Step…Action… " statements or "If…Then… " statement or as step-by-step animations.

Process RIOs are used to teach the flow of events in a mechanical, business or scientific system. Processes can be presented as staged tables, block diagrams or cycle charts.

- Staged table – can be presented as a table, diagram or chart that presents stages and what happens during each stage.
- Block diagram – can be presented as a flowchart showing all stages of the process.
- Cycle chart – can be presented as a directed graph representing cycles in a process.

Principle RIOs are used to coach tasks that requires judgment or application of guidelines to continue with a task. Principle RIOs include principle statements, guidelines and examples/analogies:

- Principle statements – a statement describing accepted behaviour
- Guidelines – decisions made by professionals in certain scenarios
- Examples / Analogies – illustrations of guidelines and principles applied or violated in practice in various contexts

Five to nine RIOs are aggregated with an overview, a summary and assessment opportunities to form a RLO (Figure 3.15).

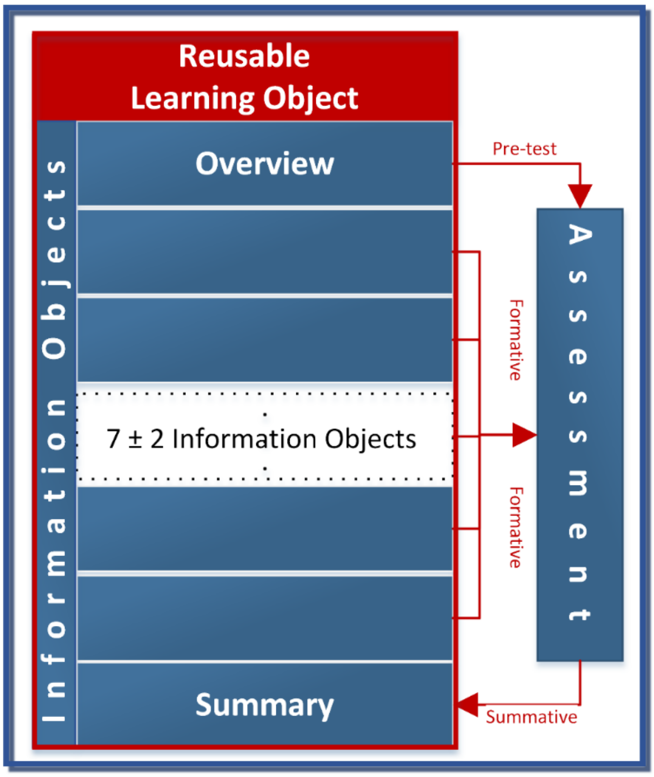


Figure 3.15 Structure of Reusable Learning Object (adapted from CISCO Systems (2001))

Limiting the number of RIOs in the five to nine range improves the granularity of the Learning Object, which in turn enhances reusability. Each RIO consists of content, practice and assessment items. Assessments can be presented as pre-assessment or post-assessment. Pre-assessments are used for recognition of prior learning to exempt or recommend learners to complete specific RIOs. Post-assessments are used to determine mastery of the content or to prescribe RIOs not yet mastered. Practice items allow learners to reinforce mastery of skills and knowledge by providing mentoring and feedback. Practice activities could take the form of a case study, simulation, practice test or hands-on practical laboratory worksheet. Each RIO should have at least one practice activity and help prepare learners for assessment.

Lessons correspond to RLOs and sections within a lesson correspond to RIOs. If necessary, the hierarchy can be expanded further: lessons combine into a module, modules into a unit and units into a curriculum.

3.3.1.3 Learning Object Metadata

Reusability of digital Learning Objects is achieved through metadata that categorises the Learning Objects stored in LO repositories (Palavitsinis, Manouselis and Sanchez-Alonso, 2014; Wang, 2008). Metadata can provide descriptive, structural or administrative information about Learning Objects (Chembrakuzhi and Haneefa, 2014).

- Descriptive metadata identifies and describes the content and usage of LOs
- Structural metadata records relationships among LOs
- Administrative metadata records the contextual information about LOs

Metadata provides a way for developers of instructional material to share information about their artefacts and to provide suggested usage guidelines. To enable communication, all stakeholders need to speak the same language. Consequently, several metadata standards and schemas have been proposed to catalogue Learning Objects in repositories (Chembrakuzhi and Haneefa, 2014; Friesen, Fisher and Roberts, 2003). These schemas provide a shared vocabulary for instructional designers. The IEEE 1484.12.1 – 2002 Standard for Learning Object Metadata (LOM) (Figure 3.16) is a pioneering schema upon which several other standards are based. IEEE LOM specifies a hierarchy of elements of a Learning Object that should be described and provides a standardised vocabulary for these elements.

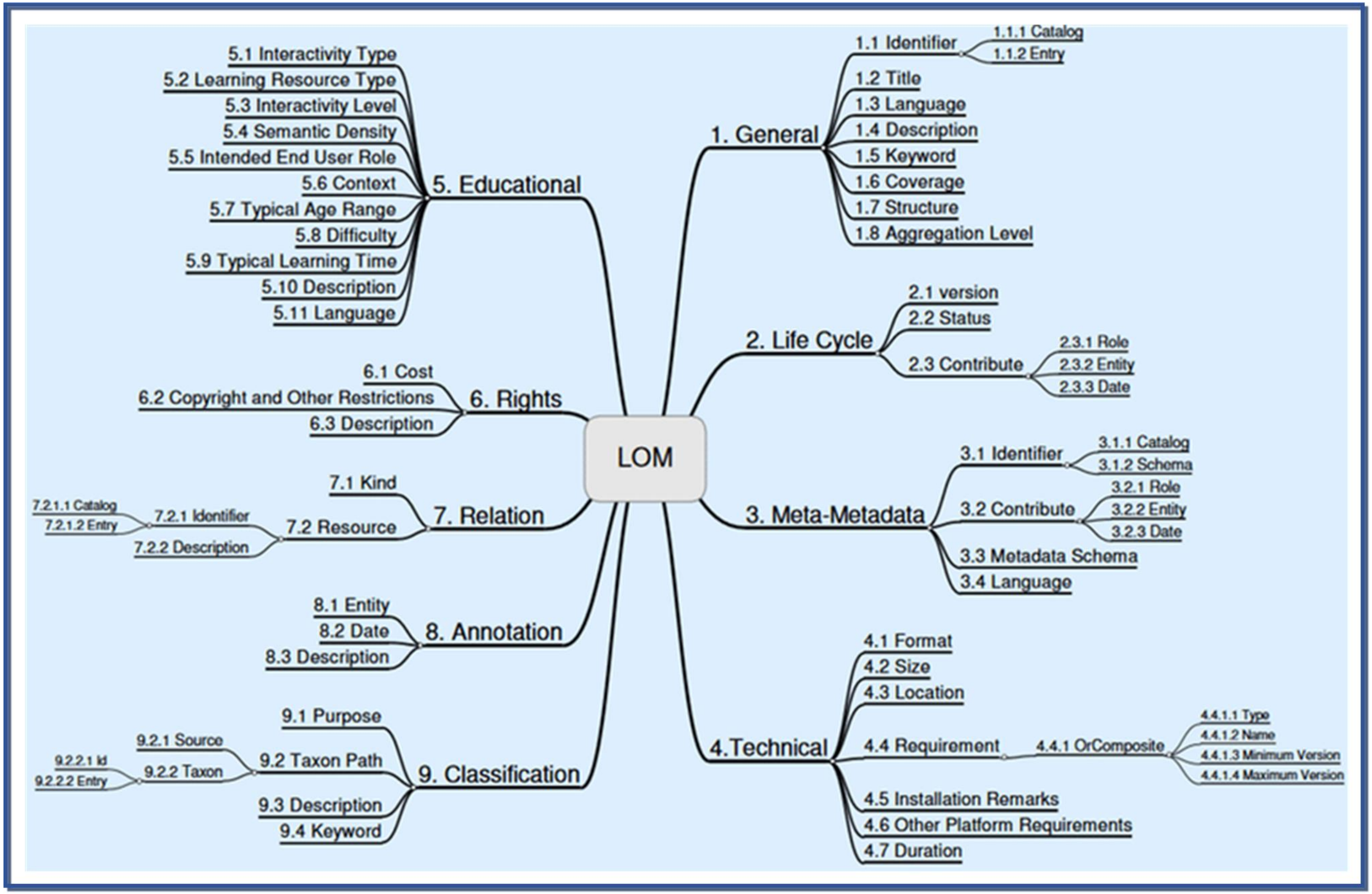


Figure 3.16 IEEE LOM Metadata Schema (IEEE 1484.12.1, 2002a)

The elements of the IEEE LOM Standard are grouped in nine categories (Chembrakuzhi and Haneefa, 2014):

- General – Introductory overview information about the Learning Object (LO)
- Lifecycle – Version control and author information
- Meta-Metadata – Information about the schema used to describe the LO
- Technical – Practical implementation details about the LO
- educational – Pedagogical attributes associated with the LO
- Rights – Intellectual property rights and other restrictions intrinsic to the LO
- Relation – Metadata that describes associations with other LOs
- Annotation – A narrative comment on potential educational use of the LO
- Classification – The classification of the LO according to a chosen taxonomy

The IMS Global Learning Consortium developed the Learning Resource Metadata (LRM) specification as part of the early drafts of IEEE LOM. A revised version (1.2) was differentiated from IEEE LOM, but the latest version (1.3) of the IMS LRM is again realigned. IMS released a best practice guide for IEEE LOM as part of version 1.3 (IMS Global Learning Consortium, 2006). Similarly, in order to improve the consistency among different organisations using the IEEE LOM specifications, the CanCore Metadata Initiative also developed best practice guidelines to standardise entering LO metadata based on the IEEE LOM schema (Friesen, Fisher and Roberts, 2004).

While these metadata schemas are broad in scope, for this thesis only educational metadata has relevance. To tailor instruction, pedagogic characteristics of the Learning Objects must be recorded. Therefore, only the educational metadata categories of IEEE LOM and the CanCore guidelines will be described in more detail.

Markup languages are used to make metadata schemas machine readable. Extensible Markup Language (XML) is the de facto standard of the IEEE LOM schema, but other languages also exist for different metadata standards. These include:

- HTML (HyperText Markup Language)
- RDF (Resource Description Framework)
- MARC (MAchine Readable Cataloguing)
- SGML (Standard Generalised Markup Language)

The educational metadata category of IEEE LOM consists of 11 fields: Interactivity Type, Learning Resource Type, Interactivity Level, Semantic Density, Intended End User Role, Context, Typical Age Range, Difficulty, Typical Learning Time, Description, and Language. To ensure interoperability and enable machine readability, Markup languages are used to record metadata tags. However, it is beyond the scope of this thesis for Learning Objects to be automatically discoverable. Therefore, descriptions and potential values of each field are only given in human readable format. IEEE LOM proposes an initial set of non-exhaustive values for each field that can be augmented by other existing vocabularies.

Interactivity Type refers to the main learning approach that is supported by the LO. The three values describing the interactivity type are:

- Active – the LO requires action on the part of the learner (not to be confused with the “active learning” instructional approach)
- Expositive – the LO provides information to the learner
- Mixed – the LO is a combination of active and expositive

Learning Resource Type describes the educational intent or content format of the Learning Object. IEEE LOM defines several possible values for learning resource types that can be extended. Popescu, Badica and Trigano (2008) further classify learning resource types as media type and instructional role.

- Media Types represent the format of the content delivered by the Learning Object (e.g. Text, Video, Image, Animation, Audio, etc.)
- Instructional Roles represent the educational intention behind the Learning Object. Instructional roles are further classified as fundamental and auxiliary.
    - Fundamental (e.g. Definition, Fact, Law, Process, Algorithm, etc.);
    - Auxiliary–Evidence (e.g. Demonstration, Proof, etc.)
    - Auxiliary–Explanation (e.g. Introduction, Overview, Conclusion, Remark, Synthesis, Objectives, Additional Info, Hints, etc.)
    - Auxiliary–Illustration (e.g. Example, Counter Example, Case Study, etc.)
    - Auxiliary–Interactivity (e.g. Exercise, Exploration, Simulation, etc.)

Interactivity Level describes the extent to which the learner can manipulate the Learning Object. The values describing interactivity level are based on a scale from very low to low, medium, high and very high.

Semantic Density describes an estimation of the succinctness of the Learning Object regarding its size, duration (Friesen et al., 2003) and number of concepts (Battou, Mezouary, Cherkaoui and Mammass, 2011). Values for semantics density fall on a similar scale as interactivity level, very low through to very high.

Intended End User Role describes the person who the Learning Object is predominantly designed for and his or her associated duties toward the object. Sample values for typical roles suggested in IEEE LOM include teacher, author, learner and manager. The vocabulary can be changed or extended to suit the context in which the Learning Object is used. Multiple user roles can be described, listed in order of dominance.

Context describes the institution and level most appropriate for the Learning Object. Values to describe the context may include school, higher education and corporate training.

Typical Age Range describes the age of potential learners targeted by the Learning Object. This age is usually linked to chronological age for school-level learners, but changes to predominantly developmental age at higher levels. Numerical values representing the age are typically used to describe chronological age and a brief textual description for developmental age. Grade levels can also be used for school-level and higher education learners.

Difficulty describes the effort the intended audience can be expected to exert when using the Learning Object. Recommended values include very easy, easy, medium, difficult and very difficult. Difficulty is commonly combined with Context and Typical Age Range.

Typical Learning Time describes an estimation of the time expected to use the Learning Object until objectives are achieved.

Description is reserved for additional comments about the Learning Object not covered in any of the other fields.

Language describes the language used in the Learning Object.

Learning objects at a finer level of granularity, tagged with educational metadata, can play a role in the provision of differentiated instruction (Popescu et al., 2008).

3.3.2 Differentiated Instruction Through Tailored Learning Objects

According to Brusilovsky, Wade and Conlan (2007), courses built with reusable Learning Objects are prone to be static in nature and follow a one-size-fits-all approach. Since the late 1980s, Intelligent Tutoring Systems (ITS) (Polson and Richardson, 2013) and later in the 1990s, Adaptive Hypermedia (AH) (Brusilovsky, 1996) started the trend to provide personalised online learning opportunities. Adaptive Learning Systems provide different support to learners based on interests, knowledge, backgrounds and learning styles (Brusilovsky and Millán, 2007). This thesis focuses on building a learning profile based on multiple learning style theories. A technique for building a learning style-based profile is proposed in Section 5.5. Educators often create, present and sequence Learning Objects based on experience and knowledge of learning strategies employed by the typical learner. Developers of Adaptive Learning Systems believe learning design should not be a one-size-fits-all endeavour, but that learning should be tailored to match the learner profile. A technique is therefore necessary to bring Learning Objects in line with learning strategies (Çinici and Altun, 2018).

This Section (3.3.2) examines approaches and best practices employed by developers of Adaptive Learning Systems in order to determine how Learning Objects can be tailored to the unique learner characteristics that influence learning approaches and strategies employed by learners. Potential learner characteristics are proposed in Section 4.3.1 and matched with suitable Learning Objects tagged with educational metadata.

3.3.2.1 Levels of Tailored Instruction

Educators who tailor instruction believe that a teaching strategy meant to help one group of learners could potentially hinder another group (Brusilovsky and Millán, 2007). Using different teaching strategies, either matched or mismatched to learners, could keep them engaged and challenged (Manning et al., 2010).

To enable **differentiated instruction**, learners sharing common characteristics receive tailored instruction unique to the group. Learning objectives are reached through a predetermined sequence of Learning Objects personalised to learners based on their learning needs, goals and characteristics (Tomlinson et al., 2003). Differentiated instruction does not tailor instruction in real time, but educators base their instructional design choices on what they know about groups of learners in their class.

Adaptivity, like differentiated instruction, tailors content and provides individualised learning paths. Profile building is done in real time and adaptation rules are used to conditionally include relevant fragments and hide or annotate navigation links (De Bra et al., 2003). In adaptative learning systems, the content is tailored in real time.

Personalised learning, like adaptivity, provides real-time profile building and adaptation. The adaptive algorithms used for personalised learning provide a higher level of personalisation than adaptive learning systems (Halim, Ali and Yahaya, 2011). Learners generally start with an initial diagnostic test to initialise instruction to the unique individual needs. Learners may also be given control of their learning and the learning environment, like in individualised or self-directed learning.

Individualised or self-directed learning enables learners to determine their own pace to work through the course material (Kop and Fournier, 2010). Objectives not yet achieved must be highlighted so that learners can revisit the appropriate content. Learners may also set their own learning agenda and must manage their own progress. A system where learners can manipulate the learning environment is generally referred to as **adaptable** (Rodriguez and Ayala, 2012).

Table 3.5 Levels of Tailored Instruction

	Profile Building	Learning Design
Differentiation	Historical, Implicit	Lecturer Control
Adaptation	Real Time, Implicit	Automated
Personalisation	Real Time, Implicit	Automated or Learner Control
Individualisation	Real Time, Explicit	Learner Control
Adaptability	Real Time, Explicit	Learner Control

This research is situated in the provision of differentiated learning. However, since valuable insights can be gained from the other closely related levels of tailored instruction, relevant concepts from each level (Table 3.5) are relevant to this study.

3.3.2.2 Learning Design Layers

At the core of this study is a belief that Instructional Design should be informed by behavioural patterns exhibited by learners as they navigate through the course material. The learning design in an adaptive learning system that adapts to learning patterns is abstracted in a layered model (Atif, 2010). This type of layered abstraction makes it easier to define differentiation goals during instruction design (Figure 3.17).

At the base is the **domain model** that represents content knowledge as an ontology of relevant concepts and semantic relationships between these concepts. The domain model can be represented as a conceptual graph with nodes representing concepts and edges representing relationships between concepts (Melia and Pahl, 2007). Domain experts are responsible for preparing and structuring learning content. At this level of the learning design, the content can be tailored to the learners' backgrounds and current competence. The domain layer should be pedagogically neutral.

The next layer, **goal and constraint model**, overlays required competencies and instructional and pedagogical constraints by applying prerequisites and postconditions in the form of learning rules to the domain ontology. Instructional constraints lead to the sequencing of concepts based on whether knowledge of one concept is needed before the learner can move on to another concept. To define pedagogical constraints, learners must be grouped based on prior knowledge and learning goals. Individual learner preferences described in learning style theories are also defined in the goal and constraint layer (Melia and Pahl, 2009).

The **learner model** layer represents the learner profile. The learner profile can be built explicitly by asking relevant questions to the learner, or implicitly through inferring relevant characteristics by analysing their behaviours (Graf, Kinshuk and Liu, 2008). The learner model can capture the knowledge progression before, during and after instruction. The learner model can also record learner goals, needs and preferences. The learner model is built while learners work through the course material.

The **resource model** focuses on identifying, repurposing or constructing Learning Objects that represent the learning content. These Learning Objects are tagged with metadata based on a standard specification such as IEEE LOM (Atif, 2010). The resource model is, therefore, the layer where the basis is set for instructional design tailored towards the characteristics defined in the learner model. The focus of the adaptation is on the content and presentation of the Learning Object.

The **course model** sequences Learning Objects based on the characteristics defined in the learner model. The learner's knowledge, goals, needs and preferences will ultimately dictate how the learner will traverse through the coursework as represented by the domain and goal and constraint models (Melia and Pahl, 2007).

The **validation model** is used to examine the instruction design prior to course delivery (Melia and Pahl, 2009). Validation ensures that Learning Objects are consistent and sequenced appropriately. Validation criteria linked to educational theories must be applied to each layer of the learning design model.

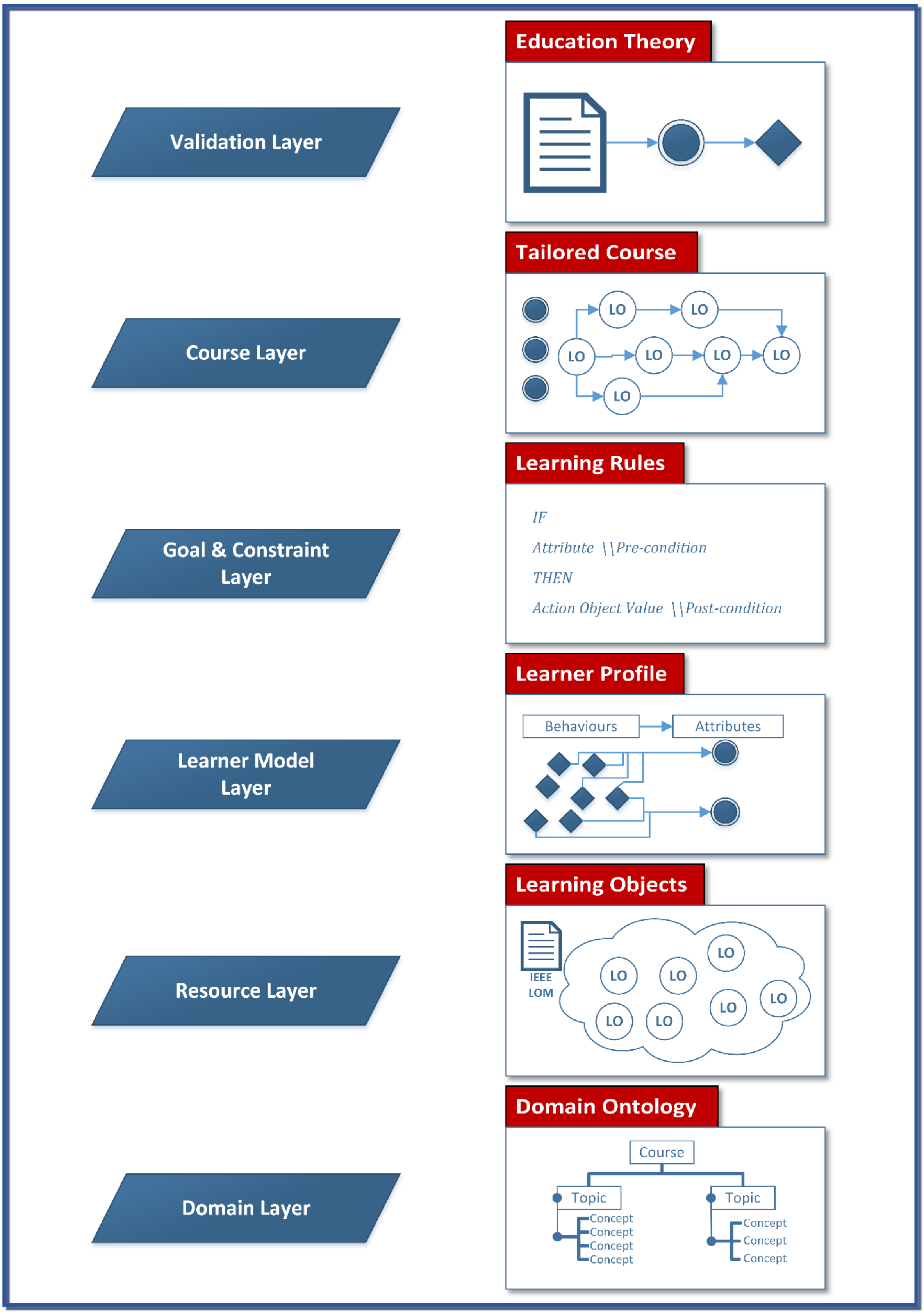


Figure 3.17 Layered Learning Design Model (adapted from Atif (2010))

3.3.2.3 Tailoring Granular Learning Objects

Fine grained Learning Objects play a vital role in each level of tailored instruction defined in Section 3.3.2.1. Battou et al. (2011) propose the structure of a fine-grained Learning Object to include:

- Concepts covered by the LO
- Time taken to use\consult the LO
- Media used to construct the LO

Content is presented on four layers (Figure 3.18):

- **Multimedia bricks** that correspond to assets as defined in a Sharable Content Object Model (SCORM), e.g. Text, Image, Video, Audio, etc.
- **Fragments** that aggregate multimedia bricks together to play an instructional role on the same level as the CISCO RIO types, e.g. Definition, Introduction, Example, Exercise, etc.
- **Documents** that consist of a combination of fragments
- **Courses** that are built by combining documents together

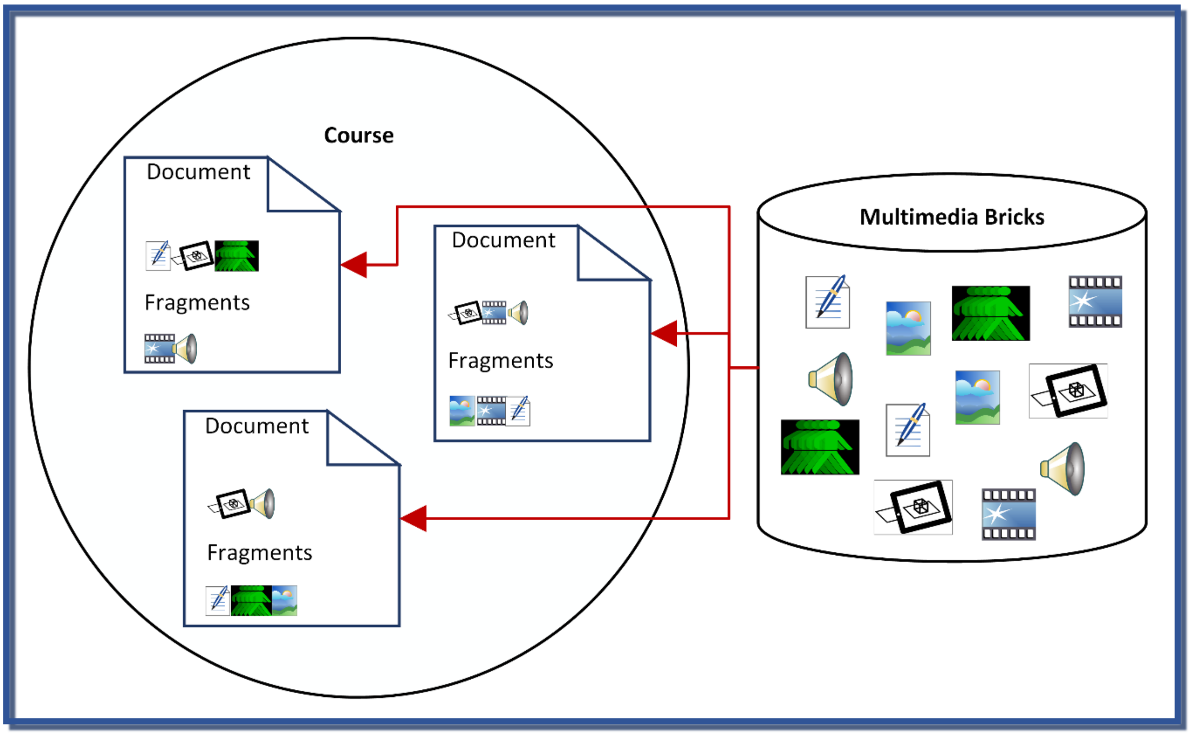


Figure 3.18 Content Model for a Learning Object (adapted from Battou et al. (2011))

Multimedia bricks and fragments are tagged with IEEE LOM educational Metadata that can be used to match different Learning Objects with different learners. For example, media bricks can be selected based on learner preferences and the sequence of fragments can be shuffled based on their cognitive needs or goals.

Battou et al. (2011) propose an architecture for implementing an Adaptive Learning System based on fine-grained Learning Objects (Figure 3.19). The architecture for an Adaptive Learning System consists of three main components, Domain, Learner and Adaptation models.

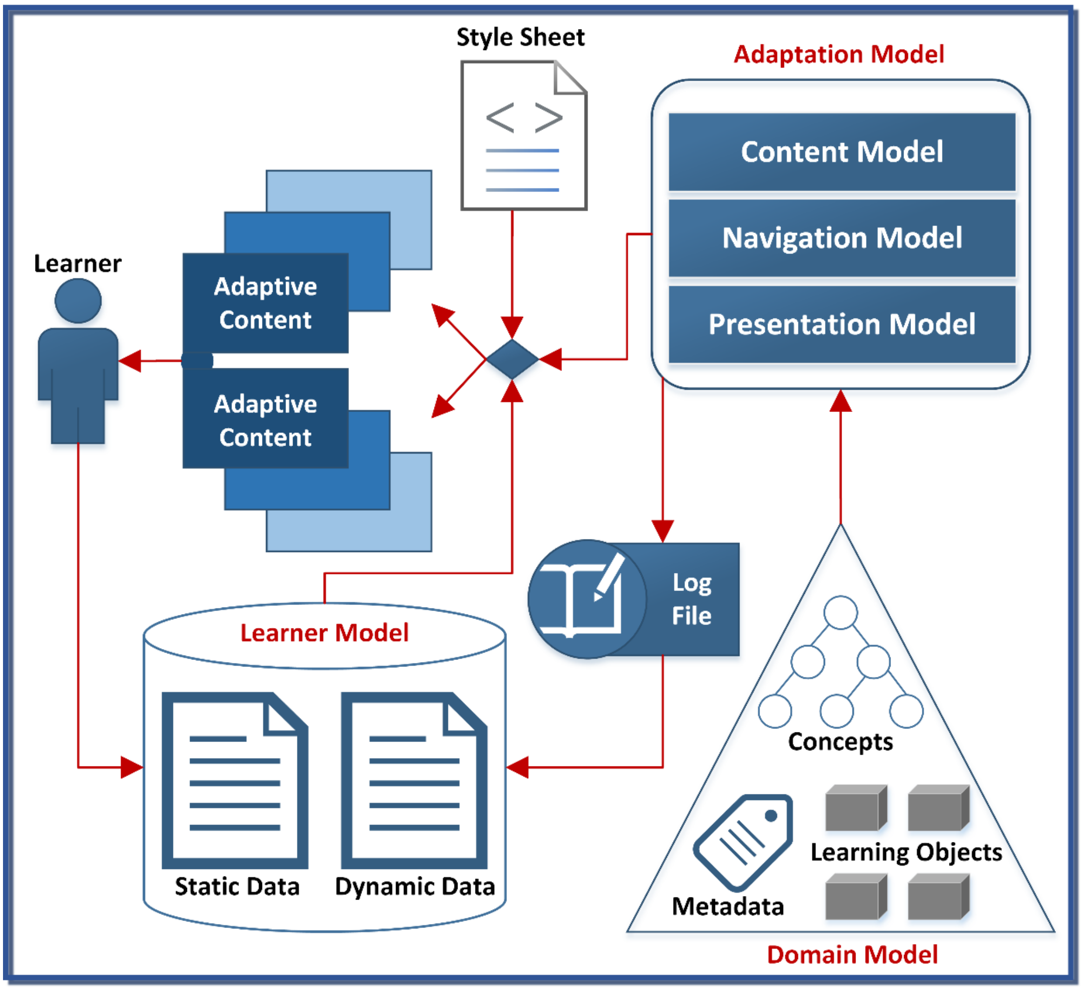


Figure 3.19 Adaptive Learning System Architecture (adapted from Battou et al. (2011))

The **domain model** consists of concepts describing the knowledge of the domain and Learning Objects to acquire the knowledge. In terms of the layered learning design model of Figure 3.17, a domain model as conceptualised in Battou et al. (2011) is analogous to the domain model, goal and constraint and resource model. In this thesis, the domain model refers strictly to the module outcomes and learning content divorced from any pedagogical strategies.

The **learner model** contains information about the learner, specifically personal details, preferences and knowledge. Some of the information in the learner model is considered static and some is dynamically updated as a result of learner progress. Both static and dynamic data is available for viewing by the learner and educator. The learner can update static data manually when needed. Dynamic data is updated based on the learner interactions with the content.

The **adaptation model** contains rules to tailor content, navigation and presentation of Learning Objects. The content model focuses on concepts covered by Learning Objects. Some learners may require supplementary content, or they may need concepts to be explained in alternative ways. The navigation model defines the sequence of Learning Objects that may be reorganised based on learner knowledge, goals or unique preferences. In the presentation model the focus is on visual media aspects of the Learning Objects.

While several Adaptive Learning Systems exist that are built on similar architecture as shown in Figure 3.19 (Brusilovsky et al., 2007), this study revolves around enabling differentiated instruction in Moodle. Consequently, Moodle as a representative Learning Management System is described next.

3.4 Learning Management Systems

Learning objects can be created using a wide variety of third-party authoring tools and presented using building blocks provided by a suitable Learning Management System. While Learning Management Systems in general do not provide adaptive learning (Popescu, Trigano and Badica, 2007a), they can be extended with adaptive functionality (Graf and Kinshuk, 2007). The choice of Learning Objects is driven by a learner portfolio created from learner behavioural data collected by the LMS. Some Moodle resources and activities and potential data sources are described in Sections 3.4.1 and 3.4.2, respectively.

The conceptual framework derived from Section 3.4 feeds into:

- A technique for building a learner profile from Moodle data (Section 5.5)
- A technique for enabling differentiated instruction using the resources and activities provided in Moodle (Section 5.4)

3.4.1 Moodle Resources and Activities

Learning content is delivered through a Moodle **resource**, a static entity with low interactivity level. Moodle provides several types of resources as part of a default installation (Moodle Docs, 2018b):

- **Book** is a module that presents information as a collection of webpages sorted into chapters. Each webpage can be loaded with multimedia elements. The book resource allows sequential navigation through Previous/Next hyperlinks or the ability to jump around using a clickable table of content.
- **File** is a module through which externally created electronic files can be uploaded. Files can be in a variety of formats such as images, Microsoft (MS) Word documents, slideshows, Portable Document Format (pdf), videos and many more. While all file types can be uploaded, learners need to have the correct software installed to be able to view the content on their devices.
- **Folder** helps to organise and categorise files of any type together under a single entity. Learners can decide to download files one by one or as a single unit.
- **IMS Content Package** allows the importing of Learning Objects and their associated metadata from other Learning Management Systems or LO repositories. The IMS Global Learning Consortium develops standards for specifying the structure of these content packages, which enables their reuse on various platforms.
- **Labels** can be used to organise the appearance of the content on the course homepage. A label can be in the format of text, graphic or embedded video.
- **Page** is a single, standard webpage that can be created using the HTML editor provided by Moodle. Multimedia content can be embedded on a Moodle page.
- **URL** is a hyperlink to any content that is stored outside the Moodle server hosting the course. The URL resource can also be used to link to web pages in your own course.

Interactivity is provided by adding a Moodle **activity** to a course. Learners can interact with the course content and follow learners or educators. Activities can be graded. A standard Moodle installation offers the following activities (Moodle Docs, 2018a):

- **Assignments** provide an opportunity to learners to complete a piece of work and upload the results for evaluation. Teachers can grade the submissions and provide formative feedback.
- **Chat** is a synchronous activity through which learners can communicate with each other or the teacher in real-time.
- **Choice** is used to collect responses from learners on a question asked by the teacher. The options are given as a multiple-choice list of radio buttons.
- **Database** provides a dynamic content repository that can be updated, maintained and searched by both teachers and learners. Database records can be in any format from files, web pages and graphics.
- **Feedback** activity allows teachers to create their own surveys. Unlike the Choice activity, the Feedback activity allows multiple choice and open-ended questions.
- **Forum** is an asynchronous activity through which teachers and learners can initiate discussions. Responses are stored in threads and may include multimedia elements.
- **Glossary** allows learners and teachers to create an online dictionary of definitions for important terms. Glossary terms are displayed alphabetically and can be grouped into categories.
- **Lesson** activities enable learners to navigate through the course content through different paths. The Lesson activity is said to be “adaptive”, however, it does not create or store a learner profile – a key element of Adaptive Learning Systems. The Lesson allows branching into an alternative track based on learner responses to prompts at various stages throughout the course.
- **Learning Tools Interoperability (LTI) External tool** allows the importing of activities and their associated metadata hosted externally. External sources must be compliant with the IMS Global Learning Tools Interoperability standard. The activities must already be deployed and configured to run on a server beyond the site hosting the course.
- **Quiz** provides several question types that a teacher can use to assess learners’ knowledge. Most question types can be automatically marked. Feedback and correct answers can be shown either immediately or deferred to a later time.

- **SCORM** packages created in external authoring tools such as Articulate Storyline can be imported into a Moodle course. SCORM compliant packages hosted externally can also be imported.
- **Survey** modules can be used to gauge learners' experience with the online course. The survey activity delivers verified questionnaires, such as Constructivist Online Learning Environment Survey (COLLES) to reflect on the online learning environment and Attitudes Towards Thinking and Learning Survey (ATTLS) to measure learners' attitudes towards thinking and learning.
- **Wiki** enables teachers and learners to collaborate on creating webpages on various topics. The standard HTML editor provided in Moodle is used to create and maintain these webpages.
- **Workshop** is a tool through which learners can submit their work for peer assessment through a rubric provided by the teacher. Learners' work can be graded by fellow learners and the teacher, and the quality of learners' assessment of others' work, can also be graded by the teacher.

These are the basic building blocks through which online course content can be created in Moodle. The concept of a Learning Object can be implemented in Moodle using any combination of these resources and activities. The open source nature of Moodle enables developers to create new resources and activities in PHP that can be plugged into the standard Moodle installation. These building blocks can be tailored to suit various learners, if we know the characteristics of individual learners. These characteristics are stored in a learner profile created from data generated by learners traversing the course material. Next, potential data sources provided in Moodle are briefly described.

3.4.2 Moodle Data Sources

Moodle enables a lecturer to incorporate a wide variety of learning resources and activities into the online instructional material. Students accessing these resources and activities will generate a large amount of data while they work through the material. This data is reported in the form of various reports, blocks or activities. These include, but are not limited to:

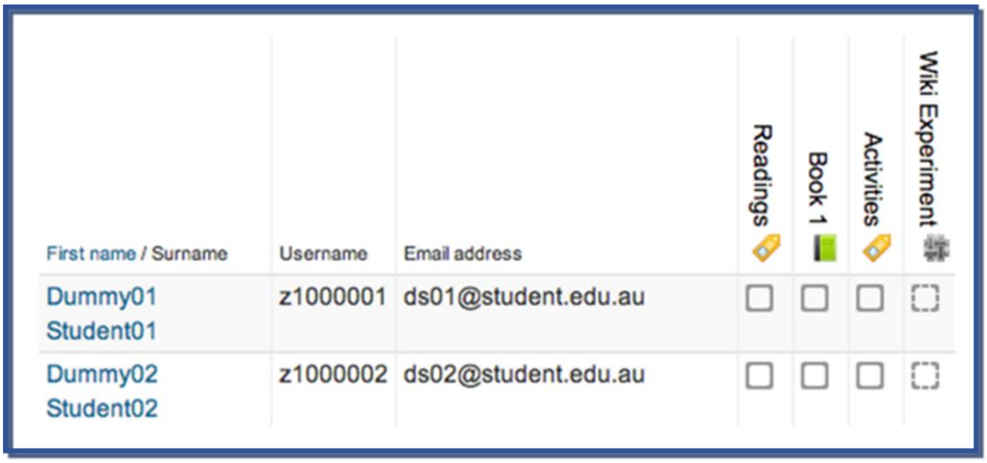
3.4.2.1 Moodle Logs

Moodle logs represent a log of events that can be filtered by student, day, activity or action. Activity refers to any resource or activity provided as instructional material. Action refers to create, view, update or delete. Each log shows a timestamp, username, activity/resource and the action that was performed on the activity/resource. The log also supplies the IP address from where the material was accessed.

The log can be downloaded as a Microsoft (MS) Excel file. Macros can be set up to filter and analyse these log files (Konstantinidis and Grafton, 2013) to discover appropriate patterns. Since the log provides a timestamp, it can be used to determine the sequence in which a learner accessed instructional material. A lecturer who wishes to, for example, distinguish between "abstract" or "concrete" learners (Gregorc, 2006) can pick up who habitually accesses theory content before examples and sequence the instructional material appropriately on subsequent topics. In addition to the timestamp, the lecturer must also be able to interpret the pedagogic intent of the specific resource or activity. One solution to this is to establish a naming convention for instructional material that describes the content. For example, when distinguishing between abstract or concrete learner a lecturer needs to also identify the content as a theoretical concept or as an example. The activity field in the log displays the name of the material, so the naming convention should somehow distinguish between theory and example. The Action field can also be used for filtering out irrelevant patterns. For patterns that involve reading the material, filter by "View" action. For patterns that involve submitting assignments or tests, filter by the "Update" action.

3.4.2.2 Activity completion

The activity completion report is a matrix of students and the activities they completed (Figure 3.20).



First name / Surname	Username	Email address	Readings	Book 1	Activities	Wiki Experiment
Dummy01 Student01	z1000001	ds01@student.edu.au	☐	☐	☐	☐
Dummy02 Student02	z1000002	ds02@student.edu.au	☐	☐	☐	☐

Figure 3.20 Sample Activity Completion Report

This report is valuable for monitoring course participation quickly. This information can be used to determine student motivation levels. For example, resistant learners can easily be identified since they will have very few checkmarks for the course material. This information can prompt the lecturer to implement some remedial action.

3.4.2.3 Course participation

Course participation (Figure 3.21) is a report that can be filtered by activity, participant (e.g. student), action (view/post) and period (days, weeks or months). The resultant log shows the name of the participant and how many times an action is completed.

First name / Surname	All actions
Gary Vasquez	No
Brenda Vasquez	Yes (5)

Figure 3.21 Sample Course Participation Report

Since this report can be filtered by activity and shows the total number of times actions are completed, it can be used to discover patterns involving the number of times a resource has been viewed, or a post has been made. For example, using this report, a student can be classified as extroverted if he or she has a high number of forum or chat posts.

3.4.2.4 Grader Report

Marks that are awarded to student assignments or completed quizzes are displayed in this report. The marks can be used for any pattern that involves some measure of performance in assessments. Certain assessments need to be classified according to type, e.g. when inferring between students who prefer organising information using synthesis versus analysis. One of the indicators of a bias towards synthesis or analysis is a student's performance in assessments requiring mastery of synthesising or analysing information (Graf et al., 2012). The lecturer would, therefore, need to have a way of identifying the type of assessment accordingly. The previously mentioned naming convention will also work in this case.

3.4.2.5 Course Dedication

This block estimates time spent by students in a course. It can show start date and time of a Moodle session, the session duration and IP address of the student (Figure 3.22). It also shows the total and average dedication time per day of each student.

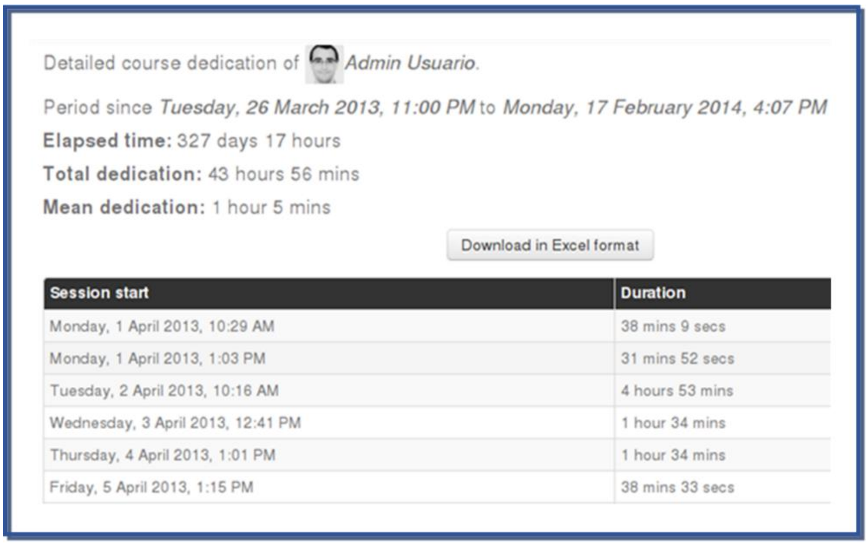


Figure 3.22 Sample Course Dedication Report

This report is useful when identifying learner motivation. For example, surface learners will spend more time close to assessment dates accessing material, while deep learners will have constant mean dedication times (Apter, Mallows and Williams, 1998).

The Moodle reports evaluated in sections 3.4.2.1 – 3.4.2.5 are all exportable in Microsoft (MS) Excel format. As evident from the evaluation of these reports, the Microsoft (MS) Excel files require further sorting, filtering and manipulation of the data to discover the necessary patterns. In most cases, the metrics provided in these reports need to be augmented by additional data identifying the pedagogical intention behind the content. The data provided in these reports will enable a lecturer to get a better profile of learner attributes in his or her class. This knowledge can be used to tailor instructional material to more closely match the needs of the class.

In some cases, simple sorting and filtering provide enough information to lecturers who wish to improve their online courses. Beyond the basics, external systems can be integrated with Moodle to analyse more complex metrics. These third-party tools include:

- **Configurable reports:** This block enables a lecturer to generate and view ad-hoc reports. The user or timeline report types will be particularly useful to monitor student activity.
- **GISMO:** A block that shows graphs of student activity participation.
- **Intelliboard:** This tool extracts statistical data from Moodle and presents it on a dashboard in the form of charts, graphs and multiple formatted reports.
- **Zoola:** Provides a library of pre-built views, reports and dashboards that are populated from Moodle data. Reports and dashboards can be scheduled, shared and embedded directly into Moodle.
- **SmartKlass:** Provides a learning analytics dashboard in a user-friendly format that can supply information to institutions, lecturers and students.
- **Blackboard Predict:** Uses Moodle data to build a predictive model that provides alerts for lecturers.
- **Loop Tool:** An open source application that provides lecturers with a course dashboard of Moodle data.

The systems mentioned in this Section provide dashboards of easy to read information to lecturers. They collect raw data from Moodle and present the data in an easy to read format for lecturers.

Should a lecturer wish to perform in-depth analysis on a large amount of Moodle data, statistical analyses can be performed using data mining tools such as Moodle Data Mining Tool, Weka and RapidMiner. Statistical analysis such as Prediction, Feature Engineering, Relationship Mining, Clustering and Factor Analysis can be performed using data mining tools. These techniques are described in Chapter 4.

3.5 Conclusion

This chapter establishes the relevance of the problem statement and provides insight into relevant issues in technology enhanced learning design.

From an examination of several procedural and conceptual instructional design models, it is evident that the learner must always be considered when designing and developing learning material (Section 3.2.3). Most procedural models advocate establishing a learner profile during the analysis phase before development of learning content and updating the profile during the evaluation phase. By only focusing on building a learner profile during the analysis phase of the instructional design, we may unintentionally create a discrepancy between the intent of the design and the actual learner experience. While the instructional design models do prescribe learner evaluation and subsequent improvements to the learning design, evaluating the design based on performance metrics alone is not necessarily an indicator that the learning design is aligned with the learners' cognitive and affective needs.

Several procedural and conceptual instructional design models can be mapped onto the analysis, design, develop, implement and evaluate phases of ADDIE (Section 3.2.3). The model developed in this study (Appendix D) is, similarly, based on the ADDIE model, but demonstrates a technique to continuously build and maintain a dynamic learner profile from data logged by a Learning Management System. Dynamically updating a learner profile will inform instructional designers whether their intended design is experienced by the learners as intended. Learning design can be tailored towards the unique needs identified in learner profiles.

A focused literature review of online learning design emphasises Learning Objects as the building blocks to create online courses (Section 3.3). One of the key principles of Learning Objects is granularity. Low-level media types are the smallest content asset and therefore the finest grain of Learning Object. These content assets are aggregated to form bigger, but less reusable Learning Objects. The task of tailoring instruction therefore becomes one of tailoring finely grained Learning Objects based on individual characteristics described in a learner profile. The learner profile is the focus of Chapter 4. For this study, differentiated learning design will be enabled through first creating, reusing and tagging learning objects during the design phase

(Section 5.4.2.2). Learning objects will be presented and sequenced according to the learner profile during the implement phase (Section 5.4.3.2).

A layered abstraction of the learning design process (Section 3.3.2.2) further simplifies the process of tailoring these Learning Objects. The model developed in this study incorporates this layered abstraction and maps each layer onto ADDIE.

An architecture for an adaptive learning system (Figure 3.19) provides the basis for a conceptual framework representing current approaches to providing differentiated instruction. The domain model applies to the design and develop phases, the learner model is the product of the evaluation phase and the adaptation model, responsible for providing tailored instruction, applies mostly to the implement layer.

Another key principle behind Learning Objects is reusability. Initially, standardisation bodies such as IEEE and IMS established a vocabulary of metadata describing Learning Objects in order to improve their interoperability and reusability. More recently, researchers and developers of Adaptive Learning Systems have been using these metadata standards to match Learning Objects with unique learner characteristics. Of particular relevance to this study is educational metadata.

Since the focus of this study is on the provision of differentiated instruction in Moodle, tools from a standard Moodle installation are briefly described (Section 3.4). This provides the general context in which two phases of the proposed solution described in Chapter 5 is implemented for evaluation.

The next chapter examines pertinent issues when maintaining a learner profile upon which differentiated learning design choices can be based.

Chapter 4. Educational Data Mining

Chapter 4
Educational Data Mining

4.1) Introduction

4.2) Learning Analytics

4.2.1) Potential uses of Learning Analytics

- **4.2.1.1) Predictive Modelling**
- **4.2.1.2) Structure Discovery**
- **4.2.1.3) Relationship Mining**
- **4.2.1.4) Distillation of Data for Human Judgement**
- **4.2.1.5) Discovery with Models**
- **4.2.1.6) Reflection on LA Goals and EDM Techniques relevant to this study**

4.2.2) Learning Analytics Ethical Code of Practice

- **4.2.2.1) Ethics**
- **4.2.2.2) Privacy**
- **4.2.2.3) Equity**
- **4.2.2.4) Reflection on Ethical, Privacy and Equity Concerns regarding LA**

4.2.3) Learning Analytics Conceptual Frameworks and Process Models

- **4.2.3.1) Elements from LA Conceptual Frameworks**
- **4.2.3.2) Steps from LA Process Models**
- **4.2.3.3) Reflection on LA Conceptual Frameworks and Process Models**

4.3) Learner Modelling based on Educational Theories

4.3.1) Techniques for Constructing a Learner Profile

- **4.3.1.1) Explicit Modelling**
- **4.3.1.2) Implicit Modelling**
- **4.3.1.3) Negotiated Modelling**

4.3.2) Content of a Learner Profile

- **4.3.2.1) Selecting Suitable Learner Attributes**
- **4.3.2.2) Mapping Online Learner Behaviours to Learner Attributes**

4.4) Conclusion

The aim of Chapter 4 is to examine relevant topics from Educational Data Mining (EDM). EDM provides potential solutions to the online learning design problem identified in Chapter 3, i.e. tailoring online instruction towards unique learner needs. The chosen topics represent a conceptual framework of best practices in the solution domain. The conceptual framework emerged from a focused literature review that synthesises the state-of-the-art discourse in the relevant topics. The main points from Chapter 4 include:

- Learning analytics
  - Potential uses, goals and associated educational data mining techniques
  - Ethical considerations
  - Learning analytics process models
- Learner modelling
  - Techniques used to build a learner profile
  - Learner characteristics stored in learner profiles

Differentiated instruction requires building a learner profile. Key sources consulted in the area of Learning Analytics and Educational Data Mining include (Baker, Shum, Duval, Stamper, & Wiley, 2012; Baker & Yacef, 2009; Siemens, 2013; Siemens & Baker, 2012). From these are derived current definitions and state of the art techniques for data collection and analysis in educational contexts. The work of (Bogarín, Cerezo, & Romero, 2018; Luna, Castro, & Romero, 2017; Romero, Ventura, & García, 2008) demonstrate educational data mining tools and techniques in the Moodle LMS. This information feeds into the learner modelling phase of the model developed in this study.

Any data-driven intervention requires careful ethical considerations. Towards developing an ethical Learning Analytics code of practice, key voices of (Bailey, Dittrich, Kenneally, & Maughan, 2012; Bull & Kay, 2007; Griffiths et al., 2016; Slade & Prinsloo, 2013; Steiner, Kickmeier-Rust, & Albert, 2016) are examined and feeds into the learner modelling phase of the proposed model. The objective of the proposed model is to be simultaneously useful in the context of instructional design and learning analytics. In this chapter, existing procedural and conceptual learning analytics models are reviewed for the same purpose as the review into instructional design models (reported in Chapter 3):

- To motivate the relevance of the solution goals and scope (i.e. a data-driven solution to instructional design)
- To derive guidelines and procedural steps for ethically creating and maintaining a comprehensive learner profile based on their online behaviours (which in turn can feed into the creation of student-centric learning experiences)

Primary sources of conceptual learning analytics models include Chatti, Dyckhoff, Schroeder, & Thüs (2012); and Drachsler & Greller (2012). Procedural models examined include Chatti et al. (2012); Clow (2012); Hundhausen, Olivares and Carter (2017); Verbert, Duval, Klerkx, Govaerts & Santos (2013) and Romero, Ventura and García (2008). These conceptual and procedural models contribute guidelines and steps for the process of building learner profiles from educational contexts.

Since the core of the learner profiles will be based on learner attributes inferred from online behaviours, a classification system is needed to suggest suitable attributes and associated behaviours. For this, the study turns to Coffield et al. (2004) to first identify influential learning style models (Section 4.3.2.1) and then deconstruct these learning style models using the metamodel of Labib et al. (2017). The deconstructed learning style models feed into the analysis phase of the proposed model.

4.1 Introduction

Technological advances are making it increasingly feasible to exploit **big data** for improved decision making (Mcafee and Brynjolfsson, 2012). Big data refers to extremely large datasets that can be analysed to detect patterns, trends or any valuable insights for the decision-making process. The need to maintain a competitive advantage is the driving force behind several organisations adopting **data analytics** to solve complex business problems. Ubiquitous data has given rise to growing research into methods for extracting meaningful information from large datasets (Provost and Fawcett, 2013). Management uses the insights provided by data analytics to take the necessary action that optimises products or services.

Data analytics borrows from multiple disciplines (Roiger, 2017), including computer science, mathematics, statistics and data warehousing. **Data mining** is used to produce actionable data analytics through discovering useful trends or patterns within the data that are imperceptible to the human eye or basic statistical techniques. Many data mining techniques use an induction-based learning process by observing known examples and building a generalised model to apply to new situations. Where no known examples exist, similarity measures are used to discover patterns in the data.

Data mining has a wide range of applications in diverse fields. For example, data mining can be used:

- In the field of emergency response to combat natural disasters (Goswami, Chakraborty, Ghosh, Chakrabarti and Chakraborty, 2018)
- In product development to recommend innovative features for products based on online reviews (Zhang, Rao and Feng, 2018)
- In logistics to improve supply chain management (Mishra, Gunasekaran, Papadopoulos and Childe, 2018)
- In cyber-security to improve malware detection (Souri and Hosseini, 2018)
- In the medical field to predict coronary heart disease (Wadhawan, 2018)

In this study, data mining is applicable to a problem identified in the field of education. **Educational Data Mining (EDM)** is concerned with developing methods to explore complex data from educational contexts (Romero and Ventura, 2010). The **Learning Analytics (LA)** research community uses EDM techniques to understand and improve learning processes and learning environments (Siemens, 2013). The vision of LA researchers is modest incremental interventions to complex educational problems (Merceron et al., 2015). The potential uses, process models and ethical considerations of LA are explored in Section 4.2. Learning Analytics is used to build a simplified model of learners with a view to optimising the learning environment (Griffiths, 2017). Learner Modelling, with a focus on the content and techniques for building a learner profile of attributes from learning style theories, is discussed in Section 4.3.

4.2 Learning Analytics

Emerging research from the EDM and LA fields is showing the promise of these related disciplines for answering educational questions. A systematic review of LA/EDM from 2008 to 2013 reveals EDM methods are typically used in education for modelling learner behaviour, predicting performance, increasing self-reflection and self-awareness, predicting learner dropout and retention, improving feedback and assessment and recommending resources (Papamitsiou and Economides, 2014). A recent systematic review of the current landscape of LA from 2012 to 2018 focuses more generally on the ability of LA initiatives to improve learning outcomes and provide support for teaching and learning (Viberg, Hatakka, Bälter and Mavroudi, 2018). The improved learning outcomes attributed to LA interventions include better knowledge acquisition, improved skill development and increased cognitive gains. The enhanced support for teaching provided by LA initiatives includes improved learner retention, optimisation of learning design, better understanding of learning processes and building learner models.

Learning Analytics and Educational Data Mining are applicable to this study, since the primary goal is to build a dynamic learner profile based on online learning behaviours and to use this learner profile to optimise the learning design. The methodology used in this study (Figure 2.7) requires the tentative proposal to be expanded into a global design and evaluated for consistency, i.e. that the solution is viable and logically constructed. The requirements for the learning design phase of the proposed model are derived in Section 3.3. The requirements for the learner modelling phase are derived in Section 4.2. In particular, Section 4.2.1 examines potential LA goals and associated EDM techniques, Section 4.2.2 addresses ethical considerations to ensure widespread acceptance of LA initiatives and Section 4.2.3 derives key elements and steps of a typical learning analytics process model by examining existing LA frameworks and LA process models.

4.2.1 Potential Uses of Learning Analytics

In learning analytics, data is analysed to build a profile of the learner, and action is taken based on the insights contained in the learner profile. Baker and Yacef (2009) present a taxonomy of potential uses of educational data mining. This taxonomy includes prediction, clustering, relationship mining, distillation of data for human judgment and discovery with models. All these objectives have one overriding goal, building a profile about the data subjects who produced the raw data through their actions or responses. Sections 4.2.1.1 to 4.2.1.5 describe the aim, information that can be modelled about the data subject, possible interventions that can be performed based on the profile and common algorithms used to build the model. Since the field is constantly growing with new techniques, the examples provided below do not constitute an exhaustive list.

4.2.1.1 Predictive Modelling

Aim: With prediction, historical data can be used to predict the values for an unknown variable. Using techniques for predictive modelling in educational contexts, features can be used to build a prediction model to estimate something that cannot be directly observed. The following are examples of information that can be predicted about learners (list not exhaustive):

- The score the learner may achieve on future tests (Conijn, Snijders, Kleingeld and Matzat, 2017)
- The time the learner may take to solve a problem
- The duration that a learner may engage with course material
- The probability that the learner is exhibiting negative behaviours, e.g. gaming the system (Baker, 2010)
- The likelihood that the learner will complete a course
- The likelihood that the learner will get the next question correct
- The likelihood that the learner mastered a specific skill

When predictions are made about the data subject, the data client can use these predictions in the following ways (list not exhaustive):

- The relationship between learner behaviour and learning outcomes can be explored
- Teachers can intervene when they see learners disengaged with the course material
- The profile can trigger automated interventions to minimise the effect of negative learning outcomes
- The learner can be allowed to progress to the next learning outcome
- Alternative material can be provided if content is not yet mastered

Techniques that are commonly applied to predictive modelling problems include classification, latent knowledge estimation and regression (Table 4.1).

Classification is used when the predicted variable (classifier) is binary or categorical. Classification methods used in educational data mining include decision trees, random forest, decision rules, step regression and logistic regression. Latent Knowledge Estimation (LKE) is a special classifier used to predict whether learners have mastered given outcomes. LKE methods used in educational data mining include Bayesian Knowledge Tracing (BKT) and Performance Factors Analysis (PFA). In regression, the predicted variable is a continuous variable. Linear regression is a method commonly applied to educational data mining.

Table 4.1 Predictive Modelling Techniques and Methods

EDM Techniques	Methods/Algorithms
Classification	• Decision Trees • Random Forest • Decision Rules • Step Regression • Logistic Regression
Latent Knowledge Estimation	• Bayesian Knowledge Tracing • Performance Factor Analysis
Regression	• Linear Regression

4.2.1.2 Structure Discovery

Aim: Structure Discovery algorithms are useful when trying to find structure in data without knowing what the outcome should be. Learning Analytics initiatives that use structure discovery techniques aim to find patterns in the data that were not clear at the start of the investigation. Structure discovery can be used in the following ways (list not exhaustive):

- Identify effective strategies used by learners
- Explore influences on learner behaviour
- Confirm hypotheses on factors influencing learner behaviour
- Discover the strength of connections between project groups to understand the differences between effective and ineffective groups
- Grouping problems together for Latent Knowledge Estimation

Techniques that are commonly applied to structure discovery problems include clustering, factor analysis, social network analysis and domain structure discovery (Table 4.2).

Clustering enables discovery of data points that form natural clusters within a dataset where the patterns are not known at the start. Algorithms used for clustering include, K-Means, Mean-Shift, Density-Based Spatial Clustering of Applications with Noise (DBSCAN), Expectation-Maximisation (EM) using Gaussian Mixture Models (GMM), and Agglomerative Hierarchical Clustering (AHC). The goal of factor analysis is to discover variables that form natural clusters of hidden factors. Exploratory Factor Analysis is used to discover patterns and Confirmatory Factor Analysis is used to test hypotheses. The Non-linear Iterative Partial Least Squares (NIPALS) method is a common algorithm used for Factor Analysis. Social Network Analysis (SNA) forms relationship models from interactions between data clients. Community Detection Algorithms are commonly used for SNA.

Table 4.2 Structure Discovery Techniques and Methods

EDM Techniques	Methods/Algorithms
Clustering	• K-Means • Mean-Shift • DBSCAN • EM-GMM • AHC
Factor Analysis	• NIPALS
Social Network Analysis	• Community Detection

4.2.1.3 Relationship Mining

Aim: The aim of relationship mining is to discover meaningful relationships between data. Relationship mining can be used to either discover or confirm strong relationships between two variables. Relationship mining can be used in the following ways (list not exhaustive):

- Discovering the correlation between certain behaviours and success
- Discovering features in a learning environment misused by students “completing” lessons without engaging with the material
- Discovering factors that will likely lead to failing a module

Techniques that are commonly applied to relationship mining problems include association rule mining, correlation mining, sequential pattern mining and causal data mining (Table 4.3).

Table 4.3 Relationship Mining Techniques and Methods

EDM Techniques	Methods/Algorithms
Association Rule Mining	• Apriori
Correlation Mining	• Pan-Correlation Mining • N-ary Schema Matching
Sequential Pattern Mining	• GSP • SPADE • FreeSpan • PrefixSpan • MAPres
Causal Data Mining	• Bayesian (CD-B or CD-H)

Association Rule Mining builds if-then rules for a set of variables strongly associated with some variable of interest. The Apriori algorithm is commonly used for Association Rule Mining. The goal of Correlation Mining is to find positive or negative correlations between variables. The Pan-Correlation Mining Algorithm and N-ary Schema Matching are examples of a method for mining correlations. The goal of sequential pattern mining is to discover relationships between variables on a particular timeline (Mudrick, Azevedo and Taub, 2018). Commonly used algorithms for sequential pattern mining include GSP algorithm, Sequential Pattern Discovery using Equivalence Classes (SPADE), FreeSpan, PrefixSpan and MAPres. Causal Data Mining is used when there is a need to confirm whether one event causes another. Bayesian algorithms like CD-B and CD-H are used to discover cause and effect relationships.

4.2.1.4 Distillation of Data for Human Judgement

Aim: In the educational context, learners generate large sets of data that needs to be prepared in a meaningful way for educators or the learners themselves.

Techniques that are commonly used to prepare data for human judgement include data visualisation and text mining (Table 4.4).

Table 4.4 Techniques and Methods for Distillation of Data for Human Judgment

EDM Techniques	Methods/Algorithms
Data Visualisation	• Heat maps • Learning Curves • Learnograms • Kinect Data • Eye-Tracking Data
Text Mining	• Sentiment Analysis • Word Clouds • Syntactic Parsing

Data Visualisation techniques provide insights into behaviours that would otherwise have been inaccessible. Heat Maps colour code user behaviour by presenting different values in different colours. "Warmer" colours usually indicate a higher concentration of a certain activity. Learning Curves visualise a learner's performance over a timeline. Learnograms show how learners switch their attention between tasks. A device such as Kinect can record human body movements that can reveal learner affect. Eye-tracking Data can be superimposed on specific areas of a screen. Eye fixations or movement paths can be represented as heatmaps or saccade paths. Text mining can sift through large amounts of discussions on forums or social media and analyse for themes. Sentiment Analysis can be used to extract attitudes from online discussions. Word Clouds can be used to indicate the frequency or importance of words in a block of text. Syntactic Parsing is useful for enabling automated marking through computationally determining the meaning of a sentence.

4.2.1.5 Discovery with Models

Aim: The aim of discovery with models is that the result of one analysis technique is used as input to another phase.

Techniques:

- Prediction > Prediction – The result of predictive modelling is used as predictor variables for a second prediction model.
- Prediction > Relationship Mining Analysis – The relationship between the results of the initial prediction model and additional variables are examined.

4.2.1.6 Reflection on LA Goals and EDM Techniques Relevant to this Study

This study is primarily focused on building a learner profile from learners' online behaviours. The primary aim of the learner modelling phase in the proposed solution is to group together learners that exhibit similar behavioural patterns. Once these learners are grouped together, the instructional design can be differentiated towards the needs of the different groupings. While all EDM techniques mentioned in Section 4.2.1.1 to Section 4.2.1.5 are relevant to capturing actionable knowledge about a learner, the primary LA goal applicable to this study is structure discovery and the chosen EDM technique is clustering. The selection of clustering in this study does not preclude future investigation into other EDM techniques to further refine the proposed model (Section 6.4.2). Any LA goal and EDM technique can be abstracted as "data analysis".

The decisions made from Learning Analytics will be more readily accepted if an ethical data collection and data analysis process is followed in any LA initiative. The requirements of an ethical LA code of practice are derived in Section 4.2.2.

4.2.2 Learning Analytics Ethical Code of Practice

One of the underlying principles of learning analytics is that it enables the optimisation of learning environments. Griffiths et al. (2016), while acknowledging the virtuous nature of these data-driven interventions to support learners, highlight some concerns with learning analytics. One potential issue to consider is the lack of clarity with regard to the scope of these interventions. Are the interventions limited to instructional support for learners or can they be expanded to include recruitment and marketing? A second concern is the transformative potential of learning analytics research in the education sector. Pushing the boundaries of education may result in unforeseen consequences that may harm instead of help. Another issue raised in Griffiths et al. (2016) is the data client's lack of control and ownership over the data once it becomes available. Once the data is part of a central repository of datasets, the original provider of the data is unable to prescribe who has access and what they intend doing with this data.

For Learning Analytics to gain widespread acceptance with learners, they need to be confident that the collection and analysis of learner data is done ethically (Section 4.2.2.1), in particular that learners' privacy is guaranteed (Section 4.2.2.2) and that equitable decisions are made based on the analysed data (Section 4.2.2.3).

4.2.2.1 Ethics

Since the establishment of the Belmont Report (NCPHS, 1979), higher education institutions have established review committees to ensure research involving human subjects is carried out ethically (Willis, Slade and Prinsloo, 2016). The principles of ethical research upheld by these review committees include **respect for persons**, **beneficence** and **justice.**

Respect for persons is shown when the individual is given adequate information and can make informed judgements based on this information. Special care needs to be taken to protect individuals with diminished capacity from harm. Informed consent by autonomous individuals or their legally authorised guardians should be sought for any ICT related research (Bailey, Dittrich, Kenneally and Maughan, 2012). The ethical use of learning analytics should go beyond informed consent. Slade and Prinsloo (2013) propose six principles for a moral learning analytics code of practice:

(1) Learning analytics must be aimed at understanding the learner and not just measuring performance. The focus of the intervention must come from a place of moral necessity and not just be based on what is effective.
(2) Learners must be collaborators of learning analytics interventions and not merely recipients thereof. Learners' participation should go beyond providing informed consent by also allowing them to interpret and reflect on their own data.
(3) Learning analytics provides a glimpse into learners at a specific point in their lives. Their performance and certain parts of their identity dynamically change over time. Data collected and analysed should therefore have an expiry clause and learners should be allowed to opt out at any time.
(4) The success of learners cannot be attributed to a single factor. Our understanding of student learning is inherently based on incomplete data and may be vulnerable to misinterpretation.

(5) Learning analytics-based interventions should be transparent. education institutions should be upfront about the purpose for the data collection and analysis and learners must be informed of possible risks.

(6) Learning analytics has the potential to improve the learning environment and learning processes. Despite the risks and constraints associated with learning analytics, education institutions would miss a valuable opportunity if they do not use the learner data available to them.

The principle of **beneficence** compels researchers to minimise risks associated with their research and maximise the potential benefits. Invasion of privacy is one of the major ethical dilemmas associated with learning analytics (Griffiths et al., 2016; Steiner, Kickmeier-Rust and Albert, 2016). For any intervention based on learner data, the potential benefits must be weighed against the privacy concerns of the learners. The issue of privacy as it relates to learning analytics is further explored in Section 4.2.2.2.

To ensure the principle of **justice**, all human subjects should have an equal chance to be selected as participants and receive equal benefits. The issue of equity as it relates to learning analytics is further explored in Section 4.2.2.3.

4.2.2.2 Privacy

To eliminate resistance to learning analytics interventions, custodians of data have an ethical and legal obligation to protect the privacy of data clients (Hoel and Chen, 2016). Learners' privacy concerns, though, should not prohibit these data-driven initiatives. In particular, the sixth principle proposed in Slade and Prinsloo (2013) maintains that it is irresponsible to ignore the potential benefits of Learning Analytics to gain insight into complex learning processes. The issue of data privacy is therefore something that deserves careful consideration to ensure acceptance of Learning Analytics.

In the information age, data protection has become a key issue related to informational privacy (Griffiths et al., 2016). This sentiment is echoed by Steiner et al. (2016) in the development of LEA's BOX, a Learning Analytics toolbox that addresses privacy concerns associated with data-driven learner interventions. The LEA'S BOX privacy and data protection framework proposes eight principles that act as best practice guidelines for learning analytics research:

(1) **Consent:** Resistance to providing informed consent can be overcome when learners are provided with relevant information presented unambiguously (Drachsler and Greller, 2012). This includes, but is not limited to, assurance that their data will be protected, a description of the type of data collected and the purpose for analysing the data.

(2) **Data protection:** Learners need reassurance that their data will be protected from abuse. Strategies implemented, such as anonymisation of data and the use of the latest encryption standards, and privacy policies should be clearly communicated to learners.

(3) **Purpose and data ownership:** The goal for collecting and analysing data should be published. Data ownership and access rights should be clearly defined and displayed throughout the entire learning analytics process.

(4) **Transparency and trust:** Transparency in learning analytics fosters trust in the process and inspires informed consent. An Open Learner Model as used by Bull and Kay (2010) has the potential to build the trust necessary to acquire informed consent.

(5) **Access and control:** While transparency provides learners an opportunity to view their data and the inferences made from this data, they should also be allowed an opportunity to modify the data where feasible.

(6) **Accountability and assessment:** Stakeholders initiating learning analytics endeavours should have clearly defined roles and accountabilities throughout the process. This is to ensure the data sources and analysis techniques are appropriate for the goal.

(7) **Data quality:** Data collected about the learner must be timely, precise, appropriate and consistent with the goal. While data quality alone will not guarantee accurate conclusions, poor data quality may certainly contribute toward incorrect inferences. All stakeholders have a responsibility to ensure the quality of the raw data collected and inferences made on the data.

(8) **Data management and security:** Policies for data management and security must be established at managerial and technical levels.

To minimise risks and maximise the benefits to be gained from learning analytics, these eight data privacy guidelines should underpin all data-driven initiatives. This supports the beneficence principle proposed in the Belmont Report (NCPHS, 1979).

4.2.2.3 Equity

To uphold the principle of justice, learning analytics must be applied fairly and equitably (Bailey et al., 2012). Unless there is a compelling reason, no learner or group of learners should be included (or excluded) from participating in data-driven interventions above others. Furthermore, if there are conflicts of interest between the educator and learner, these must be ethically managed.

The actions taken because of the data analysis should be applied consistently to all participants (Roberts, Howell, Seaman and Gibson, 2016). To this end, special care needs to be taken to ensure models developed through learning analytics are validated. Any potential for bias must be accounted for in the development of the learner models. For example, if facial recognition data is analysed, data from male and female learners must be used to create the model. Models rarely have 100% accuracy, so automated interventions must be dealt with in a sensible way (Roberts et al., 2016). Managing the reaction time based on knowledge gained from learning analytics is a fine balancing act. If, for example, an automated message is sent to learners exhibiting negative behaviour, the system should wait for multiple occurrences and adjust the model based on the severity and confidence in the accuracy of the model before reacting. However, waiting too long to intervene may also have negative consequences (Steiner et al., 2016). One possible way to avoid mislabelling a learner through inaccurate models, is the use of an open learner model as proposed in Bull and Kay (2010). Open learner models allow learners to identify potential misinterpretations made in the analysis process.

4.2.2.4 Reflections on Ethical, Privacy and Equity Concerns Regarding LA

The ethical, privacy and equity restrictions placed on learning analytics should not deter educators from using learner data towards optimising the learning environment. Instead, the learning analytics process should be accompanied by a carefully crafted code of conduct to ensure buy-in from all stakeholders involved in the process.

Section 4.2.3 explores two conceptual frameworks and five abstractions of learning analytics process models currently in use. These models represent our current understanding of the LA field and are used in this study to critically evaluate the construct validity of the proposed model developed in Chapter 5.

4.2.3 Learning Analytics Conceptual Frameworks and Process Models

The solution proposed in this study is broadly defined as a learning analytics-based process model. Consequently, existing LA conceptual frameworks and LA process models are examined in order to derive the steps and elements that should be included in the proposed solution.

4.2.3.1 Elements from LA Conceptual Frameworks

Chatti, Dyckhoff, Schroeder and Thüs (2012) describe a reference model for Learning Analytics consisting of four dimensions, What? Why? How? and Who? (Figure 4.1).

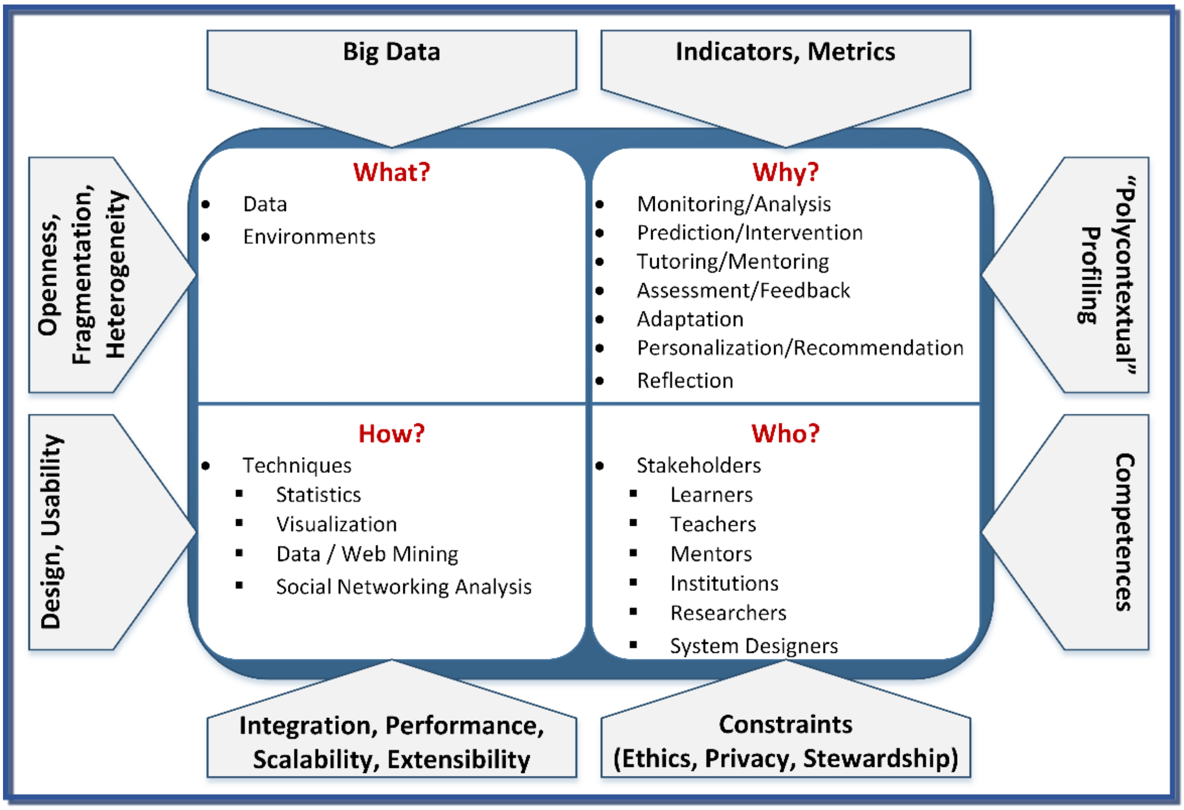


Figure 4.1 Reference Model for Learning Analytics (adapted from Chatti et al. (2012))

The "What" dimension is concerned with the type of data to be collected and analysed and the nature of the environment from which the data originates. The source of the data could be centralised from a single repository or distributed over several heterogenous systems. The "Why" dimension is concerned with the overall goal of the data collection and analysis. Section 4.2.1 provides several examples that answer the "Why?" question. The "Who" dimension refers to the stakeholders involved in the LA process. Finally, the "How" dimension is concerned with selecting relevant techniques for data analysis that match the goal of the LA initiative. This study is concerned primarily with EDM techniques, but general statistics still play a role in data analysis where relevant. The data analysis technique must always match the overarching goal of the LA initiative.

Drachsler and Greller (2012) include the dimensions of Chatti et al. (2012) but provide a design-oriented view of Learning Analytics (Figure 4.2). The “What” dimension is labelled “Data”, the “Why” dimension is labelled “Objectives”, the “Who” dimension is labelled “Stakeholders” and the “How” dimension is labelled “Instruments”.

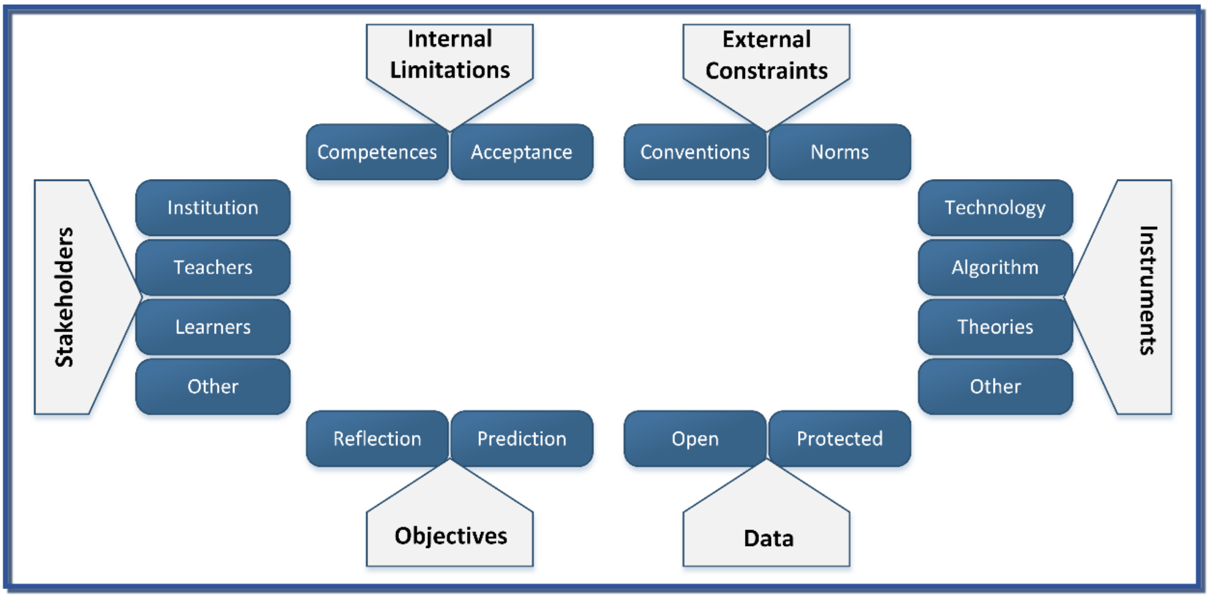


Figure 4.2 Design Framework for Learning Analytics (adapted from Drachsler and Greller (2012))

A survey based on the LA Design Framework of Drachsler and Greller (2012) reveals Learning Management Systems to be the most prominent source of learner data, with teachers and learners perceived to be the main beneficiaries of LA initiatives. Reflection on the learning process and revelation of hidden information are deemed to be the most important objectives. Trust in the accuracy and appropriateness of data analysis methods reveals highest confidence in the ability of current LA instruments to provide a comprehensive insight into learner progress and the ability to recommend relevant learning resources. Predicting future performance is at the lowest level of trust in LA instruments. The constraints that impose the highest potential barriers to LA adoption are data privacy and ownership rights. Finally, the Drachsler and Greller (2012) survey recommends that learners need additional support from teachers to benefit from LA reports adequately.

For a process model involving learning analytics to be comprehensive, the six elements recommended by both conceptual frameworks need to be addressed. These are summarised in Section 4.2.3.3.

4.2.3.2 Steps from LA Process Models

Bichsel (2012) sees analytics as a tool that uses data analysis and prediction techniques to gain insight into a strategic problem and to act upon this insight. Learning analytics initiatives involve the collection and analysis of learner data in order to optimise learning processes and the environment in which learning takes place (Siemens and Long, 2011). Several cyclical models have been proposed to abstract the steps in a typical learning analytics process.

Chatti et al. (2012) describe the process as three steps, data collection and pre-processing, analytics and action, and post-processing (Figure 4.3).

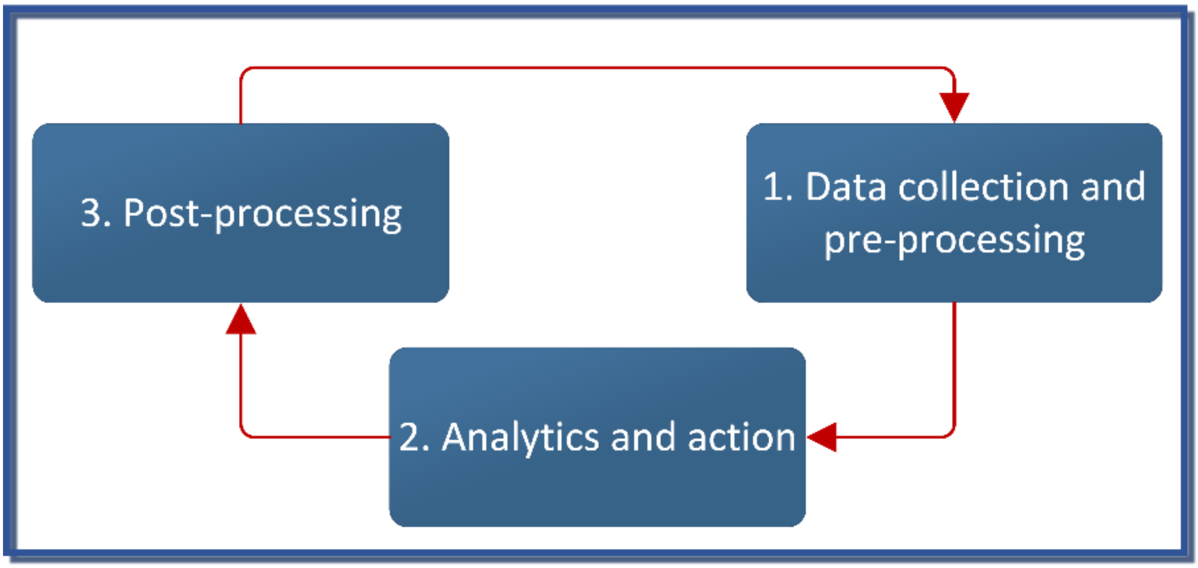


Figure 4.3 Learning Analytics Process (adapted from Chatti et al. (2012))

Data is gathered and aggregated from various educational platforms. This data is transformed into input for analysis using pre-processing techniques from the field of data mining. Learning analytics techniques are used to gain insight into strategies employed by learners navigating through online courses. The discovered knowledge of learners is used as a basis to inform suitable interventions and make informed recommendations. The final post-processing step is used to improve the analytics process.

Clow (2012) describes learning analytics as a cycle that starts with learners participating in formal or informal online learning activities (Figure 4.4). Through their actions, learners generate large amounts of data that gets logged on online learning platforms. Raw data is processed into knowledge about learning processes that can inform appropriate interventions.

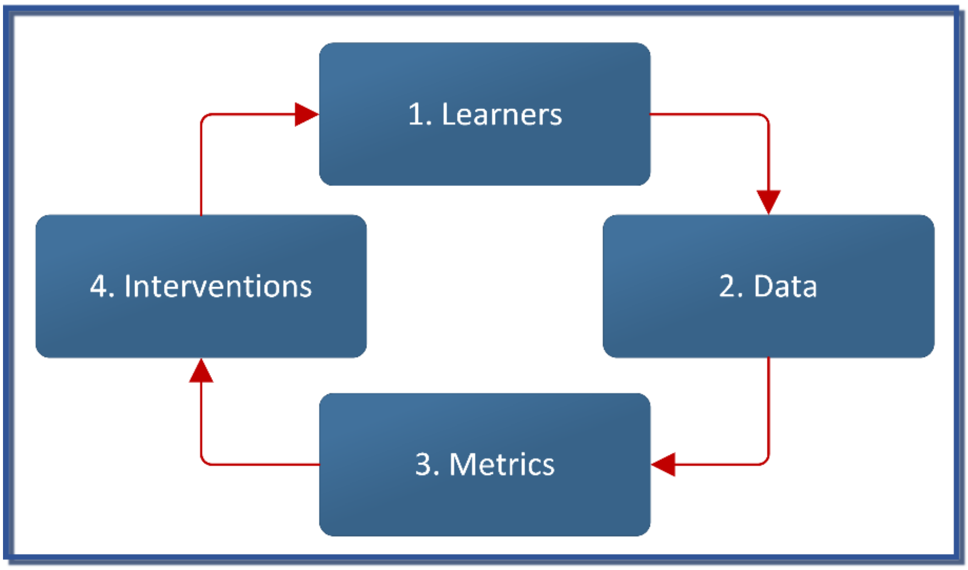


Figure 4.4 Learning Analytics Cycle (adapted from Clow (2012))

Hundhausen, Olivares and Carter (2017) describe a learning analytics process model to design an Integrated Development Environment (IDE) capable of collecting data on learning strategies while programming and intervening where necessary. The process describes four steps (Figure 4.5): collecting data from the IDE, analysing the data to discover programming behaviours, designing the intervention and establishing an automated response to scaffold learners while learning how to code.

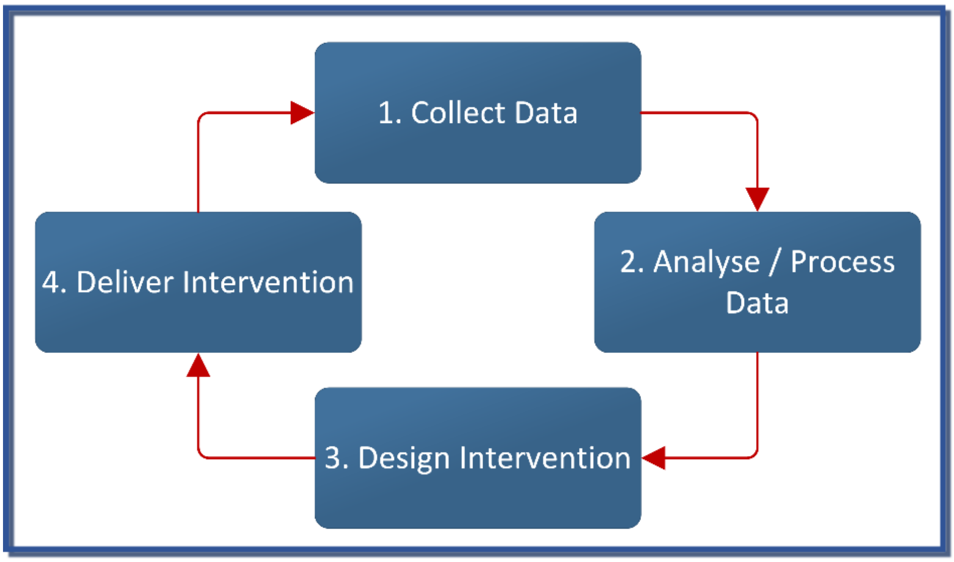


Figure 4.5 Process Model for IDE-based Learning Analytics (adapted from Hundhausen et al. (2017))

Verbert, Duval, Klerkx, Govaerts and Santos (2013) describe a cyclic model of four stages (Figure 4.6), where the focus is on the provision of a dashboard for learners to gain insight into their own learning strategies. At the first stage a dashboard will present data visually to the learner who can interrogate the data for self-reflection. After gaining a deeper understanding of their learning processes, the learners can decide whether it is in their best interest to act upon this new insight.

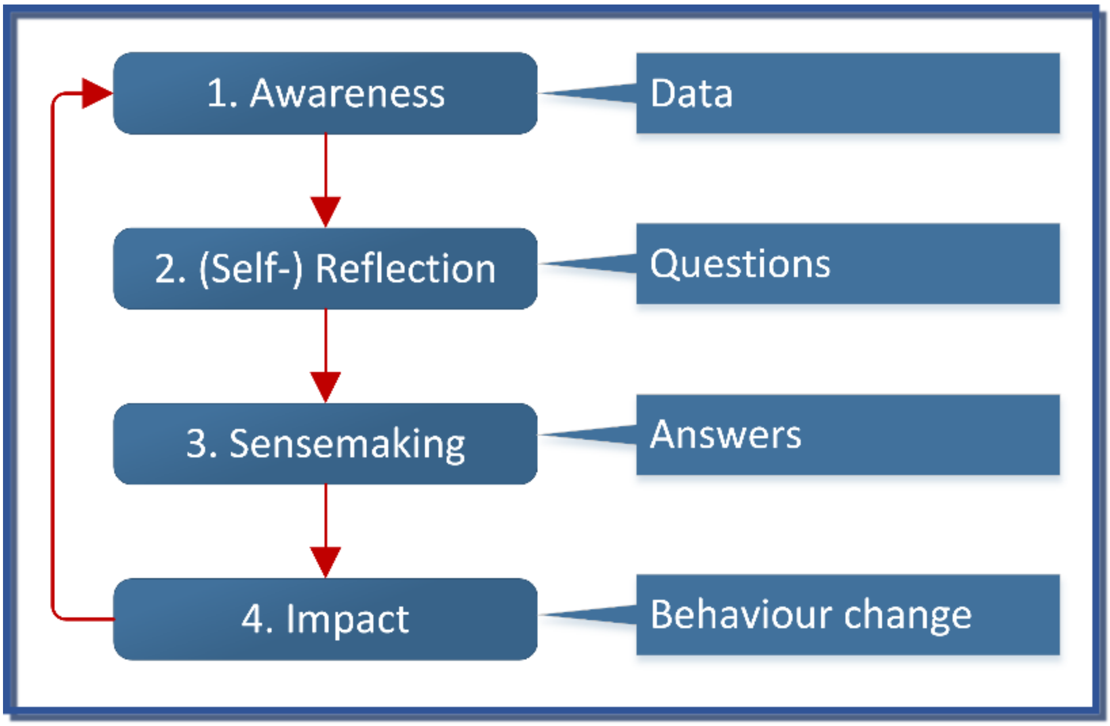


Figure 4.6 Learning Analytics Process Model (adapted from Verbert et al. (2013))

Learning analytics processes can also be used to turn raw data stored in Learning Management Systems into actionable information that can be used to enhance learning (Luna, Castro and Romero, 2017; Romero and Ventura, 2013; Romero, Ventura and García, 2008). Moodle usage data needs to undergo a pre-processing phase to transform it into a format suitable for analysis (Figure 4.7). Data mining algorithms are used on the pre-processed data to create a learner model. Knowledge represented in the learner model can be interpreted and used to make improvements to the learning environment.

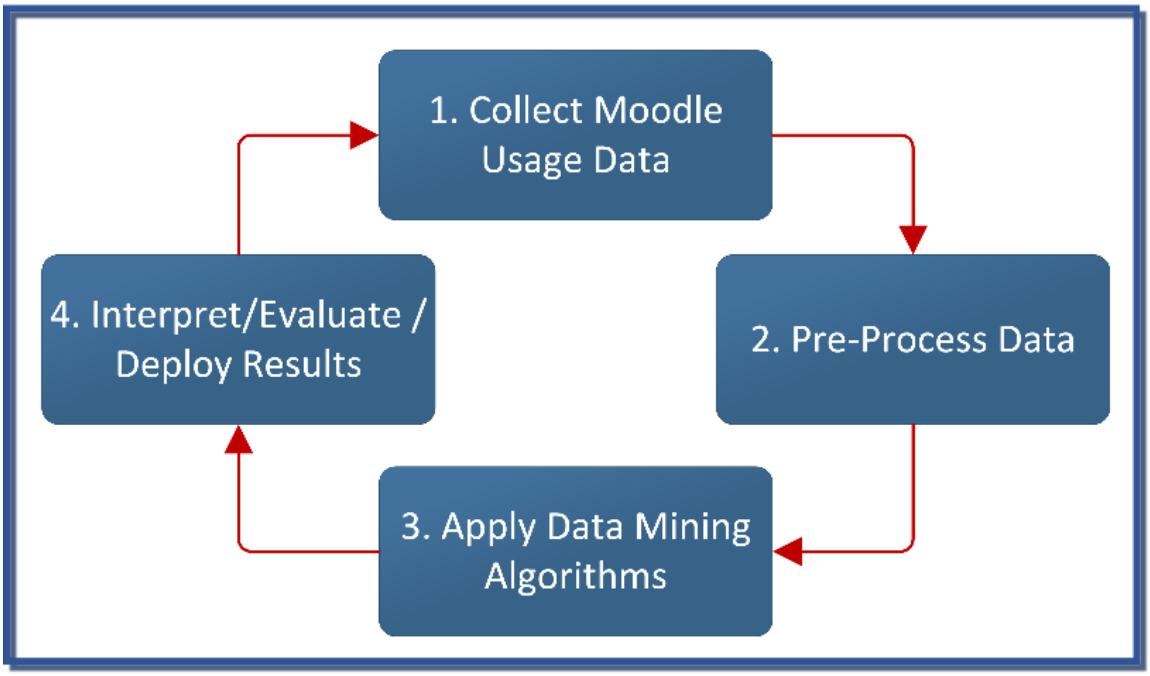


Figure 4.7 LA Process Model Applied to Moodle Data (adapted from Romero et al. (2008))

4.2.3.3 Reflection on LA Conceptual Frameworks and Process Models

The consensus between the learning analytics process models described in Section 4.2.3.2 is that they include a data collection phase, a data analysis phase and a phase where action is taken based on the results of the data analysis. What is not explicitly mentioned in the models are:

(1) An initial goal setting phase specifically linked to educational theory

(2) The form of the pedagogical intervention that can be taken based on analysis of the results

(3) An explicit reflection on an ethical learning analytics process

The first two shortcomings are echoed by Tsai and Gasevic (2017) who also identified a lack of a pedagogy-based approach to learning analytics interventions. The third deficiency in the above list concurs with a concern raised by Viberg, Hatakka, Bälter and Mavroudi (2018), who found only 18% out of 252 papers published from 2012 to 2018 on learning analytics in higher education reflected on the issue of ethics. Section 4.2.2 describes the relevant issues relating to addressing the third concern (ethics), while the next Section (Section 4.3) is aimed at addressing the first two concerns, i.e. the lack of an explicit focus on pedagogical issues to initiate and conclude a learning analytics initiative.

Based on the two frameworks of Chatti et al. (2012) and Drachsler and Greller (2012), the following six dimensions must be considered when implementing LA initiatives:

(1) **Data:** This dimension refers to the diverse range of the data sources and the type of data that are collected and analysed (Chatti et al., 2012). Drachsler and Greller (2012) add emphasis on the open and closed nature of the possible dataset that places an ethical constraint on the LA community.

(2) **Stakeholders:** This dimension refers to the people or organisations with a vested interest in the outcome of the LA initiative. Stakeholders can be classified as either data clients, the beneficiaries of the initiative or as data subjects, the suppliers of the data (Drachsler and Greller, 2012).

(3) **Objectives:** This dimension refers to the aims of the LA initiative geared towards the identified stakeholders. Drachsler and Greller (2012) identify two primary objectives for LA initiatives: reflection, providing data clients with information for self-evaluation, and prediction, providing information to data subjects that they can use for informed interventions on data clients. Example objectives of LA-based interventions are described in Section 4.2.1.

(4) **Methods:** This dimension refers to the techniques used to discover useful information from the data (Chatti et al., 2012). The methods encompass processes, technologies, algorithms, instruments and theories used in LA initiatives (Drachsler and Greller, 2012). Example EDM methods for data analysis are described in Section 4.2.1.

(5) **External Constraints:** This dimension refers to restrictions imposed on the LA initiatives. This could include sensitive issues such as legal and ethical concerns with the use of stakeholder data. Some of these issues are identified in Section 4.2.2.

(6) **Internal Limitations:** This dimension refers to the user requirements of those involved in LA initiatives. Data-driven projects applied to education require the interpretation of results and the ability to implement appropriate interventions in educational contexts.

The purpose of any Learning Analytics initiative is to know more about learners and to act upon this insight. There is, therefore, a need for a learner profile to be built. Section 4.3 explores learner modelling, with the content of a learner profile based on existing educational theories.

4.3 Learner Modelling based on educational Theories

Any type of learning environment tailored towards individual learners must maintain a learner profile (Brusilovsky and Millán, 2007). In this thesis, learner modelling refers to the process of initialising and maintaining the learner profile. In tailored instruction, the learner profile represents information about the learner that has implications for differentiated learning design. Educational data mining is especially useful for building learner profiles (Nithya, Umamaheswari and Umadevi, 2016). For the learning design to be pedagogically sound, the content of a learner profile should be backed by educational theories (Section 4.3.2). The learner profile can be constructed through an explicit modelling method (Section 4.3.1.1), whereby the learners' input is sought directly, or through an implicit modelling method (Section 4.3.1.2), which entails observing and analysing learners' behaviours (Brusilovsky and Millán, 2007, Popescu, Trigano and Badica, 2007b).

While several studies emphasise the "Technology" aspect of Technology Enhanced Learning, the "education" aspect often receives cursory attention. There is a growing movement to incorporate a pedagogical grounding in technology enhanced learning (Webb and Cox, 2004, Mishra and Koehler, 2006, De Rossi and Trevisan, 2018). In particular, adaptive education systems frequently incorporate learning style theories as the basis for learner modelling (Bayasut, Pramudya and Basiron, 2013). However, the often cited critique of educators' blind devotion to learning styles, first raised in Coffield, Moseley, Hall and Ecclestone (2004) is threefold:

(1) The abundance of learning style theories developed in silos, but often showing conceptual overlap
(2) The questionable reliability, validity and generalisability of the instruments used to measure learning styles
(3) The exaggerated claims made by proponents of learning style theories, that independent research fails to verify

Section 4.3.2 illustrates a method for populating a learner profile with knowledge about a learner in a way that addresses the issue of conceptual confusion of learning style theories. Section 4.3.1 focuses on the techniques used to build a learner profile. In particular, Section 4.3.1.1 provides a brief overview of the deficiency of using questionnaires to measure learning styles and Section 4.3.1.2 investigates a method that addresses concerns about the reliability, validity and generalisability of the measurement instruments developed by learning style theorists.

4.3.1 Techniques for Constructing a Learner Profile

Various techniques have been used to build a learner profile for learning systems that provide some level of adaptation. The most prominent technique identified in Özyurt and Özyurt (2015) is explicit modelling, which creates a static profile (Section 4.3.1.1). A method gaining prominence, in tandem with the increased sophistication of educational data mining tools, is an implicit modelling technique that maintains a dynamic learner profile (Section 4.3.1.2). A third method, negotiated modelling, proposed by Bull (2016) is a response to a growing awareness of the need to include the student voice in learning analytics processes (Section 4.3.1.3).

4.3.1.1 Explicit Modelling

A technique is deemed "explicit" when information is elicited from the learner by asking questions directly. The responses can be used to determine a range of attributes relevant to the goal of the learner profile. One can, for example, elicit knowledge, values or learning styles.

Standard assessments would fall under the explicit modelling category, since the evaluation of assessment results can be used to judge a learner's knowledge and competence. Even though gradebooks are updated throughout a course at multiple times, the update always occurs after the learner completed an assessment. Marks are assigned using standard techniques based on learners' responses, in contrast to sophisticated knowledge estimation techniques used for implicit modelling.

Learning style instruments enable teachers to evaluate their learners' characteristics, often during the analysis phase of instructional design. The instructional designer can create their own questionnaires in line with their purpose for maintaining a learner profile. Most questionnaires, though, are linked to learning style theories and inform instructional designers of learners' general abilities, proclivities, motivation, attitudes, etc. (Section 4.3.2). However, questions have been raised about the reliability of the instruments associated with learning (Coffield et al., 2004). Apart from the questionable test-retest reliability, predictive and construct validity or internal consistency of these questionnaires, they also suffer from the following weaknesses (Al-Azawei and Badii, 2014, Popescu, 2010a):

- Questionnaires rely too heavily on the metacognitive abilities of learners who may be incapable of honest self-assessment
- Excessively long questionnaires and complicated or vague questions may cause learners to arbitrarily select responses
- Learners may choose responses based on their perception of the “idealised” learner
- Questionnaire responses are static, while the constructs that they measure may be fluid and change over time or based on contextual changes
- Learners may be reluctant to complete questionnaires because they are wary of being stereotyped

Interviews can be used to triangulate data from questionnaires and learning style instruments, however, with large classes currently being the norm at higher education institutions, interviews are not a viable option in most cases.

While there is certainly still a need for explicit modelling techniques, implicit modelling has the potential to build a more sophisticated, dynamic learner model.

4.3.1.2 Implicit Modelling

Implicit modelling techniques involve surreptitiously collecting data about learners while they are studying (Popescu et al., 2008). This covert data collection and analysis intensifies the issues of ethics, privacy and equity. Consequently, the principles of ethical research (respect for persons, beneficence and justice) as demanded in the Belmont Report (NCPHS, 1979) need careful consideration (Section 4.2.2).

Learner knowledge can be estimated through the use of an **overlay model** (Brusilovsky, 2012). In an overlay model the learners’ knowledge can be computed in binary terms, as known or not-known, a weighted model that can distinguish between several levels of knowledge or a probability that the learner knows a concept.

Learners’ cognitive strategies (Graf and Kinshuk, 2008; Popescu, 2010b) and affective states (D’Mello and Graesser, 2009) can be inferred through learner **Log File Analysis**. Several learner interactions with a system get recorded in and can be extracted from log files. The data produced from these logs include time spent on Learning Objects, the order in which Learning Objects are accessed, the frequency of log ins, and the content of forums, to name a few.

Biometric data from wearable devices (De Arriba-Pérez, Santos-Gago and Caeiro-Rodríguez, 2016), facial recognition scanners (Vinay et al., 2015), eye trackers (Sharma, 2015), and cameras (D'Mello and Graesser, 2009) can also be collected and analysed to enhance a dynamic learner model. Biometric data relevant to measuring effect while learning can, for example, include heartrate, respiration, emotion or body language. Cognitive strategies may be revealed by analysing, for example, gaze patterns.

When implicit methods of profile building are used, failure to disclose the fact that data is being collected and used will be contrary to an ethical code of conduct for learning analytics. One way of obtaining the required informed consent is to open the learner model for scrutiny and validation.

4.3.1.3 Negotiated Modelling

Negotiated learner modelling leads to open learner models (OLMs). With an OLM a learner can ensure the accuracy of the content of the learner profile and enable self-reflection. An OLM can initiate a dialogue between the teacher and the learner to discuss current knowledge gaps or behaviours potentially detrimental to their progress. An OLM can also foster trust in the system, especially if data collection and analysis are automatic.

All three techniques introduced in this Section can be used together to build a learner profile. For example, the profile can be initiated through explicit modelling, dynamically maintained through implicit modelling and opened for negotiated modelling. In conjunction with the technique used, an instructional designer must decide on the information contained within the profile (Section 4.3.2).

4.3.2 Content of a Learner Profile

Educators are continuously trying to improve their teaching practice by understanding how students learn (Entwistle and Ramsden, 2015). Learning style theories are largely accepted in teaching practice, despite independent research questioning the validity of the concept of learning styles (Cuevas, 2015).

One concern is the fact that there are so many different learning style theories and that they often describe the same concept using slightly different labels. In an attempt to address the conceptual confusion, an argument can be made for deconstructing and integrating learning style theories. Popescu et al. (2007b) propose a unified learning style model that extracts learner attributes from several learning style theories. Since these characteristics represent the content of a learner profile in an adaptive education system, they must adhere to three criteria:

- The chosen characteristics must have pedagogical implications based on educational theory
- These pedagogical implications must be implementable in a technologically enhanced learning environment
- The learner attributes must be inferable by observing online behaviours

These three criteria are congruent with the framework of pedagogical practices in technology enhanced learning environments proposed in Webb and Cox (2004). According to this framework (Figure 4.8), learners will exhibit certain behaviours based on their knowledge, beliefs and values. On the other side of the spectrum, educators will combine their own knowledge, beliefs and values with pedagogical intent when creating lesson plans. Educators must have insight into learner behaviours to provide them with a high level of unambiguous affordance. Affordance of learning resources and activities will be clear if the learner is informed of the pedagogic intent behind the Learning Objects building the lesson. This affordance will aid learners in their acquisition of knowledge and skills.

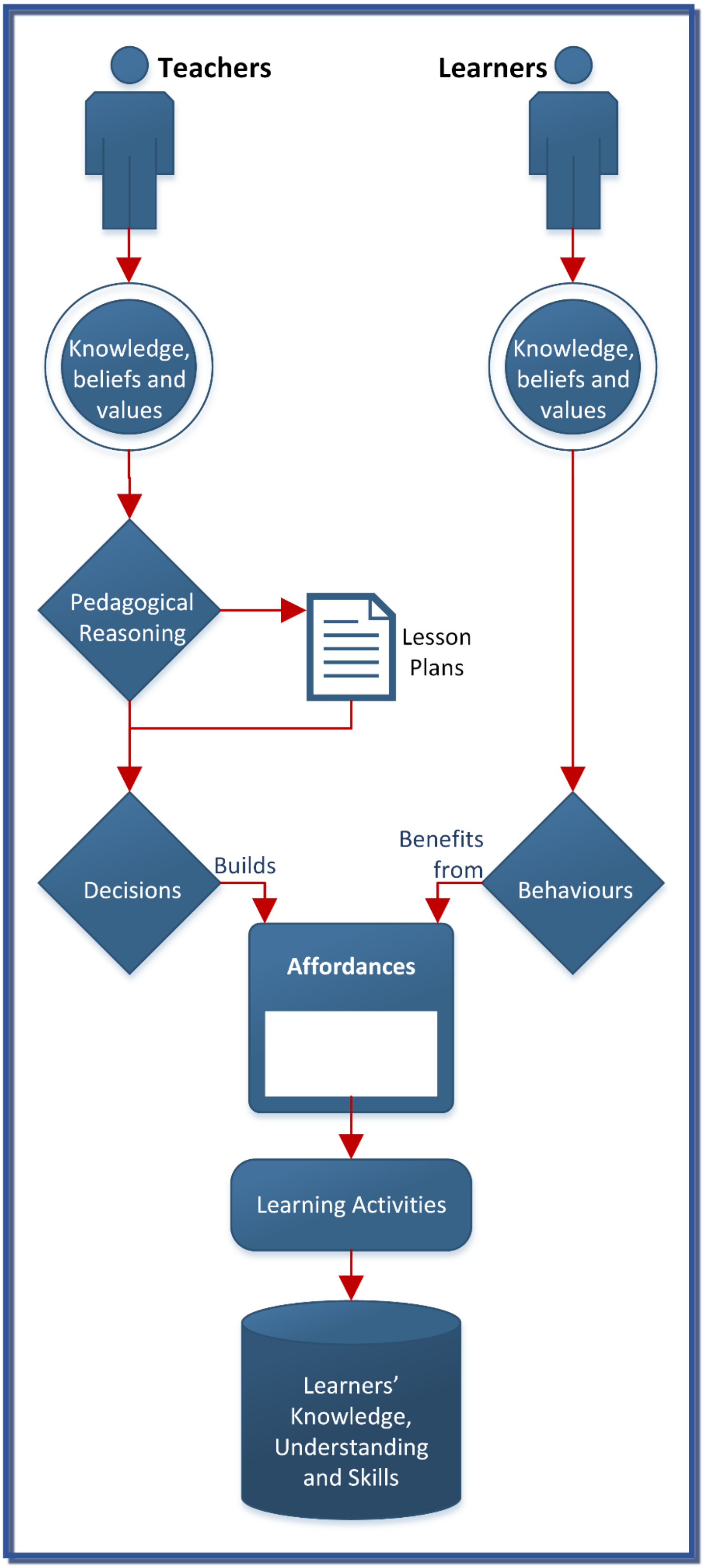


Figure 4.8 Pedagogical Practices in Technology Enhanced Learning (adapted from Webb and Cox (2004))

4.3.2.1 Selecting Suitable Learner Attributes

Like Popescu et al. (2007b), Labib, Canós and Penadés (2017) also propose building a learner profile from many learning style theories. To deconstruct learning styles, the latter developed a metamodel of the elements inherent in most learning style models (Figure 4.9). According to this metamodel, a learning style theory is composed of multiple dimensions. Each dimension describes a dichotomy of opposing poles. Each pole has one or more characteristics that can be attributed to learners. These characteristics manifest through learners online behaviours as they navigate through the course material (Graf et al., 2008, Popescu et al., 2007b, Webb and Cox, 2004). From well-known learning taxonomies, we know that learning objectives can be categorised in three domains (Anderson and Krathwohl, 2001; Krathwohl, Bloom and Masia, 1964):

- Cognitive domain, with a focus on the levels of knowledge gained
- Affective domain, with a focus on the changes in attitude
- Psychomotor domain, with a focus on manual or physical skills

While the psychomotor domain is outside the scope of this study, the cognitive and affective domains are relevant. It can be argued that since learning will bring about a change in learners' cognitive and affective domains, each learner already has innate cognitive abilities and affective states. Both cognitive and affective domains dictate how a learner will react to external stimuli. Throughout learning activities, the cognitive domain focuses on mental processes that influence how learners acquire knowledge. The affective domain dictates how a learner will instinctively react to learning tasks before cognition takes over (Zajonc, 1980).

Underpinning the origin and use of learning style theories is a wish to explain differences in learner behaviour and the impact of these differences on pedagogy. It stands to reason that the intention of learning style theories is to create a snapshot of learners' cognitive proclivities and affective states. Consequently, the metamodel of Labib et al. (2017) can be augmented by adding a learner domain component (Figure 4.9).

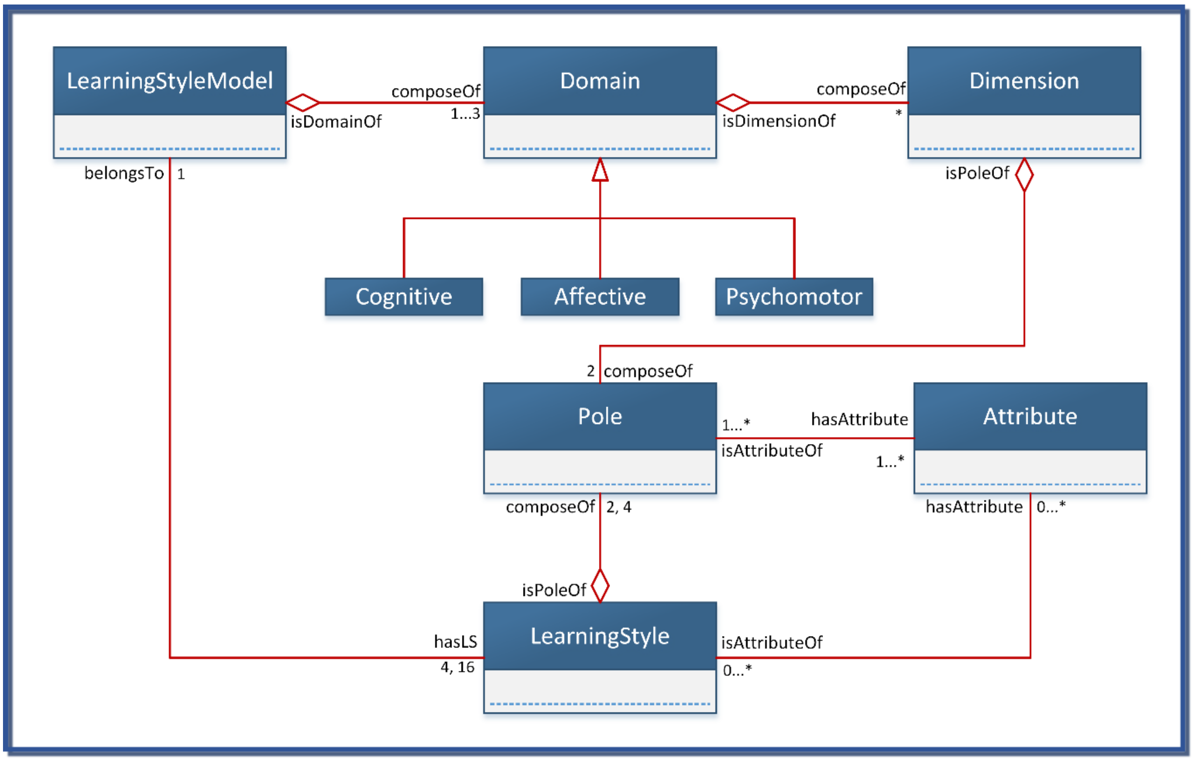


Figure 4.9 Metamodel of Learning Style Theories (adapted from Labib et al. (2017))

In a review of 71 learning style theories, Coffield et al. (2004) concluded, with reservations, that 13 influential models deserve a more rigorous examination. Each model is clustered into a grouping based on underlying theories about learning. The groupings are listed on a continuum, discussed below in terms of the proclivity of learners to change their behaviour, from most stable and unlikely to change over time (described first), to being adaptable to context fluctuations. In the discussion below, the metamodel in Figure 4.9 is used to abstract the learner domain, dimensions and characteristics from the 13 influential models identified by Coffield et al. (2004), as summarised in Table 4.5.

Constitutionally-based learning style theorists believe behaviours are biologically imposed, fixed or difficult to change. From the Coffield et al. (2004) review, influential theorists from this grouping include Dunn and Dunn (Dunn, Dunn and Price, 1981; Dunn and Dunn, 2014) and Grecorc (Gregorc, 1982; Gregorc, 2006).

- **Dunn and Dunn's Inventory of Learning Styles:** This instrument was originally derived in a primary school setting, but it has been adapted to higher age groups as well. The factors from the Dunn and Dunn learning styles model potentially relevant to Technology Enhanced Learning include motivation, perceptual modality preferences and social learning groups (Table 4.5). These factors can be mapped onto “dimensions” when using the metamodel of Labib et al. (2017). Learners exhibiting a high degree of responsibility can be said to have intrinsic motivation, while those who need parental/teacher motivation exhibit extrinsic motivation. Learners’ preferences regarding perceptual modality include visual (V), aural (A), kinaesthetic (K) (i.e. whole-body movement) or tactile (T) (i.e. touch). On the social dimension the model identifies a preference for working alone (Individual) or working with peers (group work).

- **Gregorc’s Mind Styles Model and Style Delineator:** This instrument is originally conceived for use by adult workers and later also applied in schools and universities. Gregorc describes his model using descriptors such as “Concrete Sequential”, “Abstract Sequential”, “Abstract Random” and “Concrete Random”. From this it can be abstracted that the focus is on the cognitive domain. This relates to perception, specifically an inclination towards either processing information in sequential order (Linear) or randomly (Alternating) as well as a predisposition towards making best sense from abstract or concrete information (Table 4.5).

Cognitive structure learning style theorists believe behaviours and abilities are linked to innate mental processes. From the Coffield et al. (2004) review, an influential theorist includes Riding (Rayner and Riding, 1997; Riding and Rayner, 2013).

- **Riding’s Cognitive Styles Analysis:** Riding’s theory differentiates between learning styles (stable) and strategies (variable). Riding describes his model using the labels “Holist” as dichotomous to “Analytic”, and “Linear” as the opposite of “Alternating”. These characteristics are situated within the cognitive domain (Table 4.5). Riding maps many characteristics from other learning style theorists onto the four labels and makes claims which, according to the analysis of Coffield et al. (2004), shows low reliability and validity from independent evaluations.

Stable personality type learning style theorists believe personality traits influence behaviour. From the Coffield et al. (2004) review, influential theorists from this grouping include Apter (Apter, Mallows and Williams, 1998; Apter and Desselles, 2018), and Myers-Briggs (Myers, McCaulley and Most, 1985; Myers, Kirby and Myers, 2011).

- **Apter’s Motivational Style Profile:** This theory focuses on the affective domain. It should be noted that the term “domain” as used in Apter’s theory relates to the use of dimensions as defined in the metamodel in Figure 4.9. The "Means-ends" domain maps onto the Motivation dimension, where learners with the achievement/serious orientation are extrinsically motivated and those with fun/playful orientation exhibit intrinsic motivation. The "Rules" domain maps onto the Engagement dimension. Learners classified as "Conforming" will typically show Deep engagement, while those deemed "Rebellious/Challenging" will exhibit Surface or Resistant engagement. The "Transactions" and "Relationships" domain can be reinterpreted as a Social dimension. The "Mastery" state relates to power and control, something analogous to competitiveness. The "Sympathy" state relates to care and compassion, traits shown by collaborative learners. The "Autic" state refers to those who prefer working on their own, while the "Alloic" state refers to those motivated by working with others.

- **Myers-Briggs Type Indicator:** This instrument is an attempt to attribute behaviours observed in daily interactions to the archetypes theorised by Swiss psychiatrist Carl Jung. The Extraversion – Introversion scale can be mapped onto the social dimension in the affective domain. However, the remaining three scales refer to decision making which is influenced by both affective and cognitive domains. When decisions need to be made, there is an ongoing conflict between cognitive function (sensing, thinking, judging) and affective states (intuition, feeling, perception).

Theorists who believe in **flexibly stable learning preferences** hypothesise that some behaviours remain consistent over time, while other behaviours are based on situational changes. From the Coffield et al. (2004) review, influential theorists in this grouping include Allinson and Hayes (Allinson and Hayes, 1996; Allinson and Hayes, 2012), Herrmann (Herrmann, 1991; Herrmann and Herrmann-Nehdi, 2015), Kolb (Kolb and Kolb, 2005) and Honey and Mumford (Honey and Mumford, 1992; Honey and Mumford, 2006)**.**

- **Allinson and Hayes' Cognitive Styles Index:** This instrument measures the cognitive style of the individual and classifies employees along an intuition-analysis dichotomy. Even though this dichotomy only refers to the cognitive domain, the classification of individuals based on factors influencing decision-making also touches on the affective dimension and specifically the conflict between the two when making decisions. In between intuitive and analytic, a person can be classified as quasi-intuitive, adaptive and quasi-analytic. This classifies an individual along a continuum instead of the furthest opposite sides of the dichotomy.

- **Herrmann's Brain Dominance Instrument:** This instrument is based on the principle that the brain can be divided into four quadrants: an upper cerebral region and lower limbic region, as well as a left and right hemisphere. Herrmann theorises that the majority show strong preference in two of the quadrants at a time instead of being dominant in a single quadrant. The behaviours driven by the brain's upper cerebral region focus on the cognitive domain. These are characterised by labels "Theorist/Rational (Quadrant A)" and "Innovator/Experimental (Quadrant D)". The overriding characteristic extracted from quadrant A is congruent with a reflective learner and the dominant characteristic described in quadrant D is an active learner. The behaviours driven by the brain's lower limbic system focus on the affective domain. Labels used to define affective characteristics are "Organiser/Safe-keeping (Quadrant B)" and "Humanitarian/Feeling (Quadrant C)". The descriptions of the characteristics in Quadrants B and C are not clear dichotomies and are therefore not presented in Table 4.5.
- **Kolb's Learning Style Inventory:** This instrument originated from the theory of experiential learning. According to this theory learning is best considered as a continuous process grounded in experience, instead of the outcomes that accrue through the process. In addition, knowledge is constructed through choosing between concrete experiences, reflective observations, abstract conceptualisations and active experimentations. Kolb conceived four learning styles that match the theory of experiential learning. These styles are labelled:
    - Converging: relying on active experimentation and abstract experience
    - Accommodating: relying on active experimentation and concrete experience
    - Diverging: relying on concrete experience and reflective observation
    - Assimilating: relying on reflective observation and abstract conceptualisation

The characteristics described in experiential learning are focused on the cognitive domain and information processing dimension. When mapping Kolb's learning styles to the metamodel (Figure 4.9), one can identify two dichotomies: active versus reflective and concrete versus abstract (Table 4.5).

- **Honey and Mumford's Learning Styles Questionnaire:** This instrument is influenced by Kolb's Learning Styles Inventory and is premised on a belief that learning style preferences can change to suit the needs of the individual. The learning cycle proposed by Honey and Mumford labels learners as "Activists (Stage 1)", "Reflectors (Stage 2)", "Theorists (Stage 3)" and "Pragmatists (Stage 4)". The learning cycle essentially removes opposing poles and proposes that learners move from one stage to the next in the cycle, i.e. from active to reflective to abstract to concrete.

From the Coffield et al. (2004) review, theorists who believe **learning approaches and strategies** are influenced by experience and contextual changes include Entwistle (Entwistle, Hanley and Hounsell, 1979; Entwistle and Ramsden, 2015), Sternberg and Vermunt (Vermunt, 1996; Vermunt and Donche, 2017).

- **Entwistle's Approaches and Study Skills Inventory for Students (ASSIST):** This instrument is underpinned by a belief that approaches to learning are in constant flux. Entwistle's ASSIST subscribes to the notion that learners employ different cognitive strategies in different situations and as learners' conception of learning matures. On the cognitive level, some learners exhibit a tendency towards comprehension learning and others towards operation learning. Comprehension-learning and operation-learning have conceptually similar descriptions as field-independence and field-dependence, as described in Witkin and Moore (1977). Comprehension learners display field-independence in that they are able to discover relations between topics easily. On the opposite end of the spectrum, operation learners display field-dependence in that they can easily grasp details but fail to see the interrelationships between concepts. Comprehension learners are labelled Holist, and Operations learners are labelled Serialist (Pask, 1976). On an affective level, learners fluctuate between intrinsic and extrinsic motivation throughout their formal studies. Learners' engagement levels alternate not only between deep and surface learning approaches, but also exhibit a strategic approach that combines elements from deep and surface learning. A deep learning approach is followed by learners with intrinsic motivation in seeking meaning. The strategic approach is driven by an extrinsic motivation to obtain good marks, while the surface approach is followed by those wishing to do the bare minimum to pass.

- **Vermunt's Inventory of Learning Styles:** Vermunt classifies learners as "Meaning-directed", "Application-directed", "Reproduction-directed" and "Undirected". For each of these styles, the characteristics are grouped into cognitive processing, learning orientation, affective processes, mental model of learning and regulation of learning. While learner characteristics are presented as dichotomies in other learning style theories, they are represented here in a non-mutually exclusive matrix. Elements of Vermunt's learning style are present in dichotomies derived from other theorists, e.g. "Undirected" learners sharing characteristics of Apter's "Resistant" classification, or "Meaning-directed" learners sharing traits of "Deep" learning.

While it is beyond the scope of this thesis to make any judgements regarding the merits of these learning style theories and their accompanying psychometric tests, they do provide a theoretically grounded starting point for selecting suitable learner attributes that influence their mental strategies and affective states. It is also not the intention of this thesis to explore all existing learning style theories, nor to replace current learning styles with a single unifying learning style theory. The discussion of the learning style theories, selected and based on the recommendations of Coffield et al. (2004), and the subsequent synthesis of the dimensions and characteristics in Table 4.5, illustrate a method for extracting relevant attributes from multiple theories.

This study is predicated on a belief echoed in the work of Entwistle and Ramsden (2015) that learning is a complex endeavour which no single learning style theory can adequately capture. In an effort to build rich, multi-faceted learner profiles, more rigorous analysis of a chosen learning style theory will be needed when the model for differentiated instruction proposed in this thesis is deployed in a specific context. This includes a closer examination of the measuring instrument that accompanies the description of the learning style classifications.

Table 4.5 Influential Learning Style Theories Relevant to TEL

Theorist	Instrument	Domain and Dimensions	Dichotomous Characteristics
Dunn and Dunn	Learning Style Inventory	Affective Domain • Motivation • Social Cognitive Domain • Perceptual Modality	• Intrinsic versus Extrinsic • Individual versus Group Work • VAKT
Gregorc	Mind Styles Model and Style Delineator	Cognitive Domain • Information Processing • Information Organisation	• Abstract versus Concrete • Linear versus Alternating
Riding	Cognitive Styles Analysis	Cognitive Domain • Information Processing • Perceptual Modality	• Holist versus Analytic • Verbal versus Imagery\Visual
Apter	Motivational Style Profile	Affective Domain • Motivation • Engagement • Social	• Intrinsic versus Extrinsic • Deep versus Surface\Resistant • Competitive versus Collaborative • Individual versus Group

Table 4.5 Influential Learning Style Theories Relevant to TEL (Continued)

Theorist	Instrument	Domain and Dimensions	Dichotomous Characteristics
Myers-Briggs	Type Indicator	Affective Domain • Social Cognitive versus Affective Domain • Decision Making	• Extraversion versus Introversion • Sensing versus Intuition • Thinking versus Feeling • Judging versus Perceiving
Allinson and Hayes	Cognitive Styles Index	Cognitive versus Affective Domain • Decision Making	• Analysis versus Intuition
Herrmann	Brain Dominance Instrument	Cognitive Domain • Information Processing	• Reflective (A) versus Active (D)
Kolb	Learning Style Inventory	Cognitive Domain • Information Processing	• Reflective versus Active • Abstract versus Concrete
Honey and Mumford	Learning Styles Questionnaire	Cognitive Domain • Information Processing	• Active to Reflective to Abstract to Concrete

Table 4.5 Influential Learning Style Theories Relevant to TEL (Continued)

Theorist	Instrument	Domain and Dimensions	Dichotomous Characteristics
Entwistle	Approaches to Study Skills Inventory for Students (ASSIST)	Cognitive Domain • Information Processing Affective Domain • Motivation • Engagement	• Serialist versus Holist • Field-Dependent versus Field-Independent • Linear versus Alternating • Intrinsic versus Extrinsic • Deep versus Strategic versus Surface

While learning style theorists may use different vocabulary to describe learners' behaviours, the definitions used to describe some of these labels overlap to an extent. Table 4.5 is a first step towards identifying learner characteristics representing the content of a learner profile. Combining characteristics from numerous learning style theories enables educators to create a richer learner profile than would have been possible if only a single theory was used. An overarching fusion of various learning style theories could be more effective in capturing the inherent complexity of the learning process.

Some of the most influential learning style theories are consolidated in Table 4.5, and after eliminating duplicate dimensions and characteristics, the following dimensions emerge in the **cognitive** domain:

- Information processing, organisation and reasoning – characteristics that represent inclinations towards the mental processes a learner goes through to make sense of new information and commit this new information to long term memory
- Perceptual modality – characteristics that represent the preferred means through which the learner best extracts information from the environment

Corresponding cognitive characteristics are listed in Table 4.6.

Table 4.6 Learner Characteristics in the Cognitive Domain

Information Processing, Organisation and Reasoning		Perceptual Modality
• Abstract	• Linear	• Visual\Imagery
• Concrete	• Alternating	• Verbal\Aural
• Reflective	• Analysis	• Read\Write
• Active	• Synthesis	• Kinaesthetic\Tactile
• Serial	• Inductive	
• Holistic	• Deductive	
• Field-Dependent		
• Field-Independent		

After eliminating duplicate dimensions and characteristics in Table 4.5, the following dimensions emerge in the **affective** domain:

- Motivation and Engagement – characteristics that represent the underlying reasons for learners acting in a particular way (motivation), manifesting in how they interact with the course material (engagement)
- Social – characteristics that represent how learners prefer to interact with other learners and teachers

Corresponding affective characteristics are listed in Table 4.7.

Table 4.7 Learner Characteristics in the Affective Domain

Motivation and Engagement		Social
• Intrinsic	• Deep	• Individual
• Extrinsic	• Surface	• Group Work
• High persistence	• Strategic	• Competitive
• Low persistence	• Resistant	• Collaborative
• Meticulous		• Extrovert
• Careless		• Introvert

While Table 4.6 lists cognitive characteristics that are in dichotomous poles with other cognitive characteristics and Table 4.7 lists affective characteristics that are in dichotomous poles with other affective characteristics, decision making can also be a balancing act between cerebral cognitive function and affect (feelings and emotions). Some learning style theories acknowledge this conflict between cognition and affect through characteristics listed in Table 4.8.

Table 4.8 Learner Characteristics Showing Conflict Between Cognition and Affect

Cognitively Dominant Decision Making	Affectively Dominant Decision Making
• Sensing	• Intuition
• Thinking\Analysing	• Feeling
• Judging	• Perceiving

4.3.2.2 Mapping Online Behaviours to Learner Attributes

Learners' cognitive strategies and affective states dictate their online behavioural patterns (Popescu, 2009; Graf and Kinshuk, 2008). It is, therefore, necessary to describe typical behaviours associated with characteristics identified from selected learning style theories. Table 4.5 summarises a range of dichotomous attributes proposed by influential learning style theorists that are relevant to technology enhanced learning.

The model developed in this thesis proposes log file analysis to build a learner profile of characteristics extracted from several learning style theories. Section 5.5 describes how to build a learner profile from raw data stored in Moodle log files. The behaviours described in this Section dictate which metrics and educational metadata tags are necessary to map behaviours onto learner attributes. Student interactions with the course material and each other are recorded in log files. The type of data that can be observed from log files includes:

- **Navigational metrics**, e.g. the number of times a learner returned to the same Learning Object, the order in which the Learning Objects are visited, etc.
- **Temporal metrics**, e.g. the time spent on a Learning Object, the total time spent studying, the average time spent on Learning Objects of specific types, etc.
- **Performance metrics**, e.g. the results of a quiz, number of resubmissions, etc.

These metrics alone will not be enough to describe learner behaviours and need to be tagged with educational metadata to record the teacher's pedagogical reasoning for including the Learning Objects in the lesson. The IEEE LOM metadata standard can be used in this regard (Section 3.3.1.3). For example:

- Interactivity types are useful to describe the Learning Objects as "Active", typically favoured by "Concrete" learners, or "Expositive" favoured by "Abstract" learners.
- Media type metadata is useful for describing the format of the Learning Object and can be used to record the modality through which the information is presented (visual, verbal, textual, tactile).
- Instructional role metadata is useful for describing pedagogic intention of the Learning Object. These can be used to order resources in a way suitable for:
  - Reflective learner – Fundamental before Auxiliary-Illustration
  - Active learner – Auxiliary-Illustration before Fundamental
  - Field-Dependent learner – Add more Auxiliary-Explanation where needed

When learner metrics are combined with educational metadata, an implicit modelling method can be used to build and maintain a learner profile (Section 5.5). This learner profile can be used to differentiate the learning design (Section 5.4).

4.4 Conclusions

Educational Data Mining contributes to our understanding of the learning process, which in turn enables us to optimise learning environments. In order to optimise learning, instructional designers need to build and maintain a dynamic learner profile that is used as a basis for learning interventions. The learner profile is built through data collection and analysis, which results in actionable knowledge about the learner. Example categories of learning analytics-based interventions include:

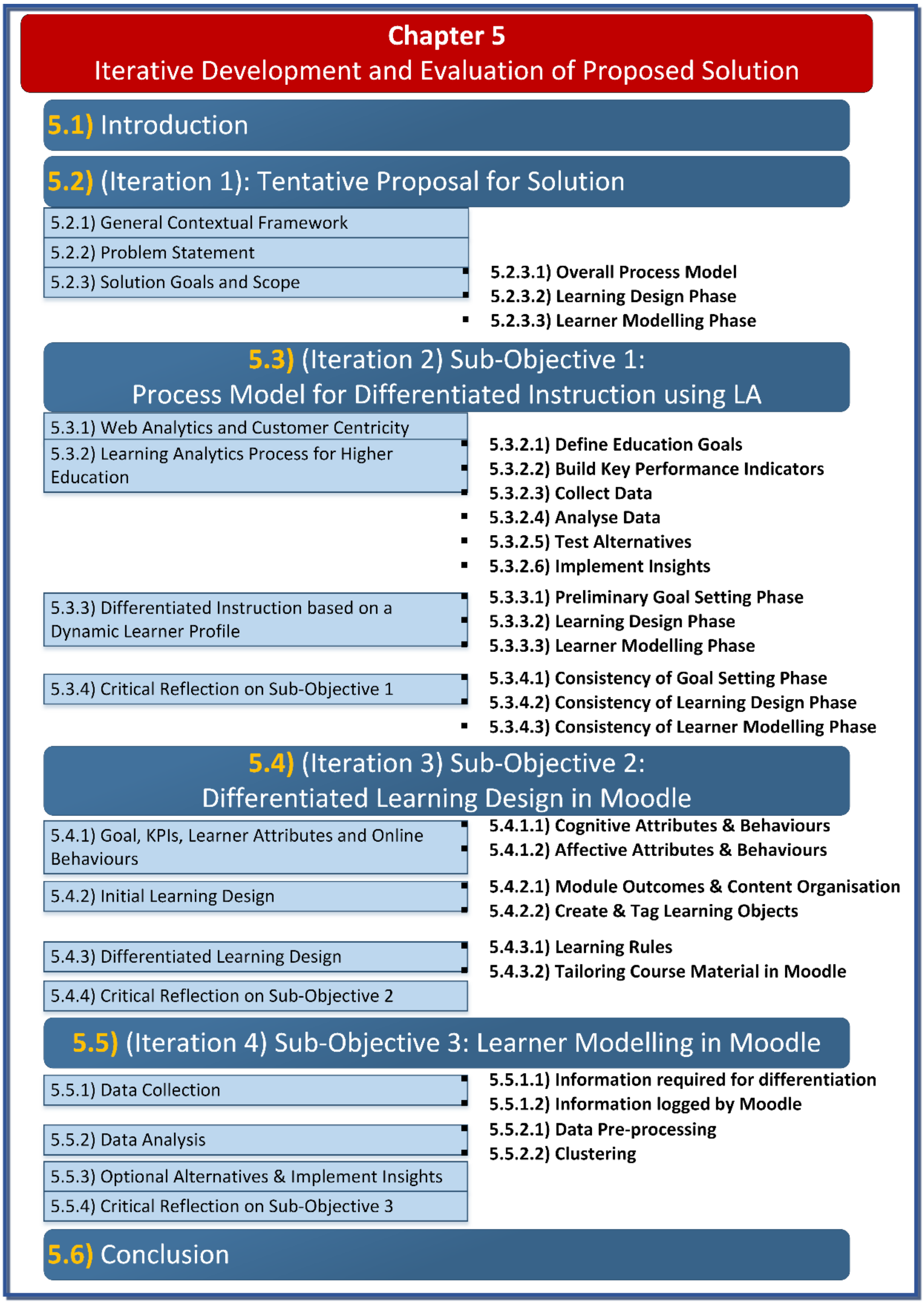
- **Predictive modelling** to model something that cannot be directly observed
- **Structure discovery** to find patterns in data that are not obvious
- **Relationship mining** to discover or confirm meaningful connections between variables that affect learning
- **Distillation and preparation of data** into meaningful information that teachers and learners can use to make informed decisions

One way of optimising the learning environment is the provision of differentiated instruction. To make suitable learning design choices, the learner profile must be validated by educational theories. Learning style theories, although contentious, represent one way of explaining differences in learner behaviour and their impact on pedagogy. Some learning style theories are grounded in a belief that learners' predilections are fixed, a claim that is not always scientifically supported. This thesis supports the notion that no single learning style theory exists that adequately represents complex learning processes. Instead of focusing on a single theorist, this thesis proposes a technique to select suitable learner attributes with pedagogic implications in a technology enhanced learning environment. To build a dynamic learner profile, these attributes must also be inferable by observing learners' online behaviour as recorded in log files. This dynamic learner profile requires an implicit modelling technique, which in turn requires a comprehensive ethical learning analytics code of practice.

This chapter identified shortcomings in existing learning analytics process models. Current models all reference data collection, data analysis and an appropriate intervention based on the outcome of the analysis phase. One deficiency in current models is a focus predominantly on the technical aspects of data collection and analysis, without specifying the need for an explicit goal setting phase before data is collected. Existing models also do not explicitly acknowledge new hypotheses that may arise after analysis is concluded. Finally, several models are silent on the need for a solid theoretical grounding rooted in educational theory and very few models mention a coordinated, comprehensive and integrated approach to ethics.

These deficiencies are addressed in Chapter 5, where a learning analytics model is iteratively developed and evaluated.

Chapter 5. Iterative Development and Evaluation of Proposed Solution



The aim of Chapter 5 is to show the iterative refinements of the proposed solution towards the problem identified in Chapter 1. The main points from Chapter 5 include:

- Elaboration on the process through which the model was iteratively designed and evaluated according to the DeRTEL methodology synthesised in Chapter 2.
- A description of the characteristics of a learning analytics model to assist with learner profile building and differentiated instruction in Moodle. The model incorporates the conceptual framework synthesised in Chapter 3 and Chapter 4 and attempts to solve the problem identified in Chapter 1.

Previous chapters elaborate on the existing knowledge base and provide the conceptual framework from which to argue steps of the proposed model. The sections in this chapter reference previous chapters to show the evolution of the process model for using learners' online behaviour to inform differentiated instructional design in a Learning Management System (Appendix D).

5.1 Introduction

Learning style theorists hypothesise that mental processes will be supported if the learning design is tailored to learners' cognitive needs. In addition, proponents of learning styles theorise that tailoring the learning design will positively influence learners' feelings about learning if their affect is considered. While studies on the effect of learning styles on learning remain inconclusive, failure of a particular learning style theory may simply be due to two issues:

(1) *Learning is a complex process that no single learning style theory can adequately capture*. This thesis proposes a technique for extracting attributes from learning style theories in acknowledgement of the complex nature of the learning process. While the extent of the impact of these attributes and related learning design choices is the subject of future work, the primary need that this thesis addresses is the practicality of implementing differentiated instruction in an existing Learning Management System based on a selection of learner attributes (Section 5.4).

(2) *There are flaws in the instrument to categorise learners into suitable categories.* This thesis proposes a learner modelling technique that seeks to use educational data mining techniques to infer learner attributes from metrics captured by online Learning Management Systems (Section 5.5).

This thesis, therefore, reports on establishing the overall process model (Sections 5.2 and 5.3) and technology used to instantiate parts of the process (Sections 5.4 and 5.5) and not on the impact of the process on learning. The proposed solution is developed for educators at any education institution that utilises a Learning Management System, and Moodle is representative of the general context. A contextual analysis of the institution would be conducted as part of future work. Only when the model is completely deployed and assessed in context at a specific higher education institution will the application domain be more rigorously examined.

Section 5.2 presents a tentative proposal for the solution as Iteration 1 of the study. Section 5.3 presents the global design of the solution (Iteration 2), by establishing a process for differentiated instruction based on a dynamic learner profile. Section 5.4 refines the learning design phase of the process model (Iteration 3) and Section 5.5 refines the learner modelling phase of the process model (Iteration 4).

5.2 Tentative Proposal for Solution (Iteration 1)

In this iteration, the tentative form and function of the proposed solution is drafted during the preparation cycle. In terms of the sub-objectives identified in Chapter 1, the proposed solution calls for three main components:

- Objective 1: A student-centric process model for differentiated instruction using learning analytics (described in Section 5.3 as the global design)
- Objective 2: A differentiated learning design phase (refined in Section 5.4)
- Objective 3: A learner modelling phase (refined in Section 5.5)

As discussed in the methodology chapter, the technical risk and efficacy evaluation strategy is deemed appropriate for this study (Section 2.3.3.2). Under this strategy, the emphasis is first on multiple artificial formative evaluations to evaluate subcomponents of the final solution. Once subcomponents are satisfactorily tested, the focus shifts to summative evaluations of the solution in the naturalistic environment. This thesis only reports on the iterative prototyping cycles and not on the assessment phase of the current study (Figure 2.7).

The research design requires an evaluation cycle for each iteration in the prototyping phase. The technical risk and efficacy evaluation strategy is selected, since the design risk is largely technical in nature (Table 2.3). Before the solution is deployed in the real environment (i.e. during the Assessment phase), parts of it must be rigorously tested to reflect on the expected practicality and utility of the solution. Any intervention that makes use of the proposed solution must therefore undergo careful formative evaluation of parts of the solution before a final impact analysis is conducted in practice.

In terms of the technical risk and efficacy evaluation strategy, all iterations must undergo formative assessment. For each objective, the preparation cycle gives a high-level overview of each component, while the critical reflections represent the evaluation cycle of each iteration. This is in line with Dewey's Model of Inquiry (Figure 2.2) as the epistemological basis for the current study. Furthermore, the evaluation of practicality and utility of the subcomponents of the proposed solution accentuates the pragmatic philosophical view underpinning the study. Similarly, evaluation of the utilitarian value of the subcomponents and the explicit emphasis on ethical reflection highlight the researcher's axiological considerations.

- The evaluation cycle of iteration 1 critically reflects on the relevance of the proposed solution (Section 5.2), by:
    - Describing the general context of the application domain
    - Deriving a problem statement from learning design and instructional technology
    - Proposing tentative goals and scope of a potential solution (Figure 5.1)
- The evaluation cycle of iteration 2 critically reflects on the consistency of each element of the global design of the proposed solution (Section 5.3.4)
- The evaluation cycle of iterations 3 and 4 critically reflects on the expected practicality and utility of the learning design phase and learner modelling phase respectively (Section 5.4.4 and Section 5.5)

The relevance and consistency are argued through citations to existing literature where parts of the proposed solutions were applied in different contexts. The expected practicality and effectiveness of the solution subcomponents are evaluated in a pilot study that instantiates the proposed learner modelling and learning design phases in the Moodle context. In this thesis, the learning design is not evaluated in terms of its impact on learning effectiveness, efficiency or learner satisfaction.

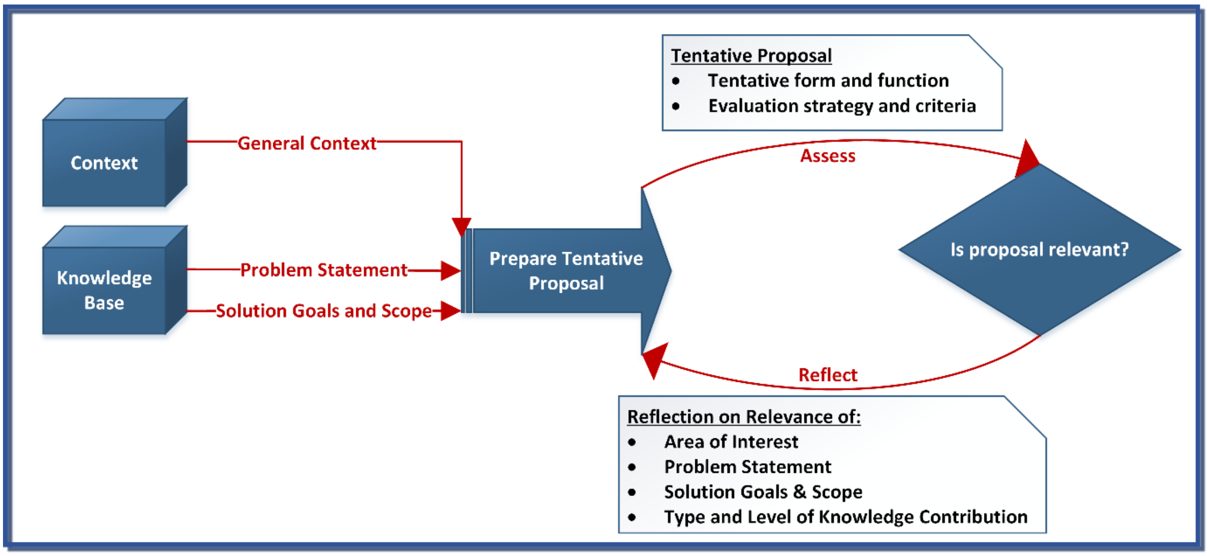


Figure 5.1 Tentative Proposal for Solution (Own Construction)

In the preliminary phase of this study, existing literature identifies a situation of concern from the problem domain and proposes a solution towards this situation of concern from the solution domain. General contextual analysis produces a contextual framework of the application domain (Section 5.2.1). A literature review of the problem domain results in a problem statement (Section 5.2.2) and literature review of the solution domain results in tentative solution goals and scope (Section 5.2.3).

5.2.1 General Contextual Framework

When the solution is deployed at an institution, a more rigorous contextual analysis is necessary to set the scope and parameters of the actual intervention. Upon deployment, it is necessary to determine if the staff and students are likely to accept the intervention. It is also necessary to examine the current infrastructure and determine whether management support can be relied upon. Since this thesis only reports on the iterative development of the procedural and technological aspects of the proposed intervention, specific institutional analysis is beyond the scope. This data driven solution is proposed for any education institution making use of a Learning Management System (LMS).

5.2.2 Problem Statement

The problem domain falls in the general area of learning design and instructional technology, with a focus on the provision of differentiated learning opportunities in an LMS. An LMS provides learning resources and activities through which learning opportunities are delivered (Section 3.4). Effective learning opportunities are generally developed through some procedural or conceptual instructional design model (Section 3.2). These models typically propose an evaluation phase during which the effectiveness and efficiency of the learning design is measured against given criteria. Feedback from these evaluations are used to improve the learning design in some way. Frequently, the learning design is assessed by focusing on the relevance and consistency of Learning Objects. When the focus is on the Learning Objects themselves and not on the way students interact with them, we are potentially missing the voice of the student when improving the learning design.

One way to learn more about the student is by using questionnaires. However, there are some drawbacks with the use of questionnaires, such as the static nature of survey results and the time needed to complete surveys (Section 4.3.1.1). There is a growing move towards using the actual learner behaviour to dynamically make inferences about each learner (Section 4.3.1.2). At institutions that use Moodle as their Learning Management System, the behaviours of learners are logged and available for educators to analyse (Section 3.4.2). But how do we collect and analyse the raw data of learner behaviours from Moodle log files into a format suitable to assist educators to tailor the learning design towards unique learner characteristics? This question leads to the problem statement:

There is limited prescriptive guidance on how to create a meaningful learner profile from Moodle logs that can inform differentiated learning design choices in Moodle, leading to inadequate instructional designs.

5.2.3 Solution Goals and Scope

The study into the stated problem is guided by the research question:

What are the steps of a comprehensive, learner-centric process model to enable differentiated instruction in Moodle based on a dynamic learner profile?

Learning analytics initiatives require a systematic process to guide effective educational interventions (Section 4.2.3). One type of intervention involves tailoring instruction based on a learner profile. In order for differentiated online instruction to be successful, it is necessary to build a dynamic learner profile of attributes with an influence on technology enhanced learning (Section 4.3). Educational data mining is increasingly used as part of these interventions (Section 4.2.1) to make sense of data generated by learners when completing online courses. In particular, educational data mining has been effectively applied in learner modelling to build this learner profile. A profile of learner attributes can be used to differentiate instruction towards the needs of individual learners. Therefore, three objectives have been identified for the proposed solution:

- A comprehensive, learner-centric process model to enable differentiated instruction based on a dynamic learner profile (Sections 5.2.3.1 and 5.3).
- The requirements for a Learning Design Phase (Section 5.2.3.2), instantiated in Moodle (Section 5.4).
- The requirements of a Learner Modelling Phase (Section 5.2.3.3), instantiated in Moodle (Section 5.5).

5.2.3.1 Overall Process Model

A generic goal of Learning Analytics Process models is the optimisation of online platforms used in Higher education. A Learning Management System (LMS) is an example of one such platform that facilitates online learning. There is an emerging trend towards educators embracing blended learning in their facilitation of modules. In South Africa, the Department of Higher education and Training (2013), in the White Paper for Post School education and Training, encourages the use of digital technology to optimise learner engagement. Since there is a positive relationship between student retention and their academic experience, which includes the use of technology (Carter and Yeo, 2016), higher education institutions should endeavour to understand the needs of their student body and optimise the online learning platforms accordingly. The online learning environment should be set up in such a way that it enables educators to build a dynamic profile of their learners, which in turn can be used to inform differentiated learning design choices based on individual profiles. The learner profile and the learning design should be backed by educational theories to ensure pedagogically sound interventions.

There is, therefore, a need for a systematic process that is student centred, data driven and backed by reputable educational theories. An established web analytics process model developed by Waisberg (2015) has been identified as a suitable model that can be applied to higher education. By applying concepts from the problem domain (Chapter 2) and solution domain (Chapter 3), the customer-centric web analytics process model can be adopted as a student-centric learning analytics model suitable for building a dynamic learner profile upon which differentiated learning design choices can be made.

5.2.3.2 A Learning Design Phase

In this phase, learning design choices must be based on attributes stored in each learner profile. This phase describes the different ways in which the Learning Objects are presented and/or sequenced according to the behaviours exhibited by each learner. Part of the learning design phase also involves triggering interventions based on attributes that may have a negative effect on learning. Section 5.3.3 describes the generic steps in the learning design phase, and Section 5.4 explores the provision of differentiated instruction in Moodle.

5.2.3.3 A Learner Modelling Phase

There is a need for a method to build a learner profile upon which differentiated instructional choices can be based. The learner profile should contain learner attributes satisfying the following characteristics (Section 4.3.2):

- The learner attributes influence learning (either positively or negatively) based on educational theories
- The learner attributes can be inferred from online behaviours
- The learner attributes have implications for online instructional design

For each selected learner attribute, associated behavioural metrics must be identified, collected and analysed in order to infer these attributes.

Section 5.3.3 describes the generic steps in the learner modelling phase and Section 5.5 explores the use of data mining techniques to perform learner modelling from Moodle log files.

5.3 Sub-objective 1: Process Model for Differentiated Instruction using Learning Analytics (Iteration 2)

The tentative proposal identified the need for a model for differentiated instruction based on a dynamic learner profile. Section 4.2.3 examined several learning analytics process models, all of which concur explicitly on at least three steps: data collection, data analysis and action based on the analysis. By focusing on the technical aspects of learning analytics there is a danger of interventions being technology-centric instead of student-centric. A customer-centric web analytics process is examined (Section 5.3.1) and its suitability for use in higher education argued (Section 5.3.2). A deficiency in existing learning analytics processes is the lack of educational theories backing pedagogical interventions (Section 4.2.3).

The provision of differentiated online instruction (Section 3.3.2) based on a collection of learner attributes from well-known learning style theories (Section 4.3.2) is proposed as a suitable goal for a learning analytics intervention. The steps are described in Section 5.3.3 and the consistency of the process steps is argued from relevant literature where the concepts have been applied in different contexts (Section 5.3.4).

5.3.1 Web Analytics and Customer Centricity

The success of e-commerce websites relies on understanding customer behaviours (Waisberg, 2015). Once the business has insight into its customers, it can tailor the website to improve their online experience. By analysing customers' engagement with the e-commerce website, interventions can be implemented to improve customer acquisition, development or retention (Bijmolt et al., 2010). The effective use of web analytics will ensure that the business can improve its financial situation, optimise website usability and maintain a competitive edge (Phippen, Sheppard and Furnell, 2004). Therefore, customer centricity is at the centre of improving the success of the e-commerce website (Bijmolt et al., 2010). Web metrics must be identified to measure customer engagement and to track and understand customer behaviour.

The advent of Web 2.0 has given rise to more complex customer interactions that are worthy of analysis. User interactions are moving beyond the e-commerce website and include user-generated content that has relevance to a business. In order to optimise their e-commerce sites, businesses must focus on website usability and interoperability with third party websites where their potential and current customers converge. This sophistication of customer interactions and the ubiquitous nature of the data their actions generate, highlights the need for a methodical process to collect and analyse data using relevant techniques. In response to this need, Waisberg (2015) proposes a web analytics process for commercial website optimisation. This process is the foundation for a learning analytics process model proposed in this thesis (Section 5.3.2 and Section 5.3.3).

Web analytics are used by companies to understand their customers' online behaviours in order to optimise their websites. Waisberg (2015) developed a process of six steps that commercial website designers can use to optimise e-commerce websites under their control.

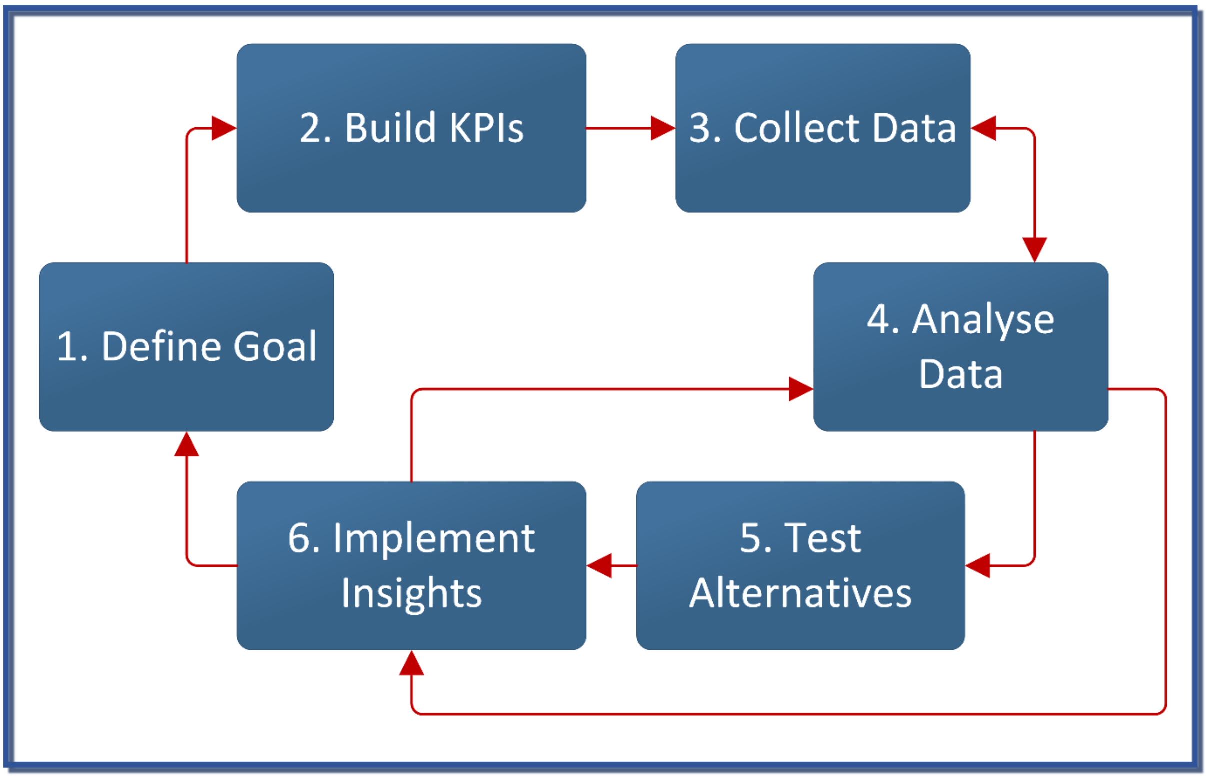


Figure 5.2 Web Analytics Process (adapted from Waisberg (2015))

(1) **Define Business Goals:** Website developers, in conjunction with business owners must define the primary business goal of the website. The goals of websites differ, with some aiming to increase page views in order to sell more advertising, while others may aim to decrease page views by providing quick answers to customer queries. Since business goals vary widely from one business website to the next, it is clear that the goal of the website must be expressed before any optimisation endeavour is to be undertaken.

(2) **Build Key Performance Indicators (KPI):** Once the business goal of the website is defined, relevant KPIs must be established. A precise KPI should enable decision makers to establish whether the business goal of the website is achieved. KPIs should be evaluated for simplicity and relevance and they must be timely and useful. In order to define KPIs to measure commercial website success, the following questions can be asked[2]:

(1) What is your desired outcome?
(2) Why does this outcome matter?
(3) How are you going to measure progress?
(4) How can you influence the outcome?
(5) Who is responsible for the business outcome?
(6) How will you know you have achieved your outcome?
(7) How often will you review progress towards the outcome?

The relevance of a particular KPI is typically evaluated through the use of the so-called SMART criteria:

- Is your objective **S**pecific?
- Can you **M**easure progress towards that goal?
- Is the goal realistically **A**ttainable?
- How **R**elevant is the goal to your organisation?
- What is the **T**imeframe for achieving this goal?

(3) **Collect Data:** Relevant metrics must be collected that can be used to measure KPIs. Identified metrics can be collected from a wide variety of different sources where appropriate and joined for further analysis. For example, data collection can be from web logs, through JavaScript tagging, through web beacons or with packet sniffing hardware. Google Analytics store metrics such as visits, bounce rate, page views, pages per visit, average time on site and percentage new visits, amongst others. To evaluate the data collection step, two questions are appropriate: (1) "Is the data accurate?" (2) "Am I collecting all the metrics that I need to understand customer behaviour?"

(4) **Analyse Data:** The data analysis step involves transforming the metrics from the data collection step into useful information about the customers. Appropriate data

[2] https://www.klipfolio.com/resources/articles/what-is-a-key-performance-indicator

analysis and visualisation techniques must be employed to discover meaningful patterns and trends over time. The data analysis step can also involve discovering clusters of customers with similar behaviours. The online environment can then be tailored towards the needs of these unique customer segments. After analysing the data, the process will branch into one of three possibilities:

- If important data is missing, go back to step 3 (Collect Data)
- If new hypotheses emerge, go to step 5 (Test Alternatives)
- If the analysis produces satisfactory results, go to step 6 (Implement Insights)

(5) **Test Alternatives:** The test alternative step is optional and necessary on condition that the data analysis step produces unexpected results that may require further exploration. During this step all new hypotheses are tested.

(6) **Implement Insights:** As soon as the "Analyse Data" or "Test Alternatives" steps reveal actionable insights into the customer's behaviour, the website optimisation can proceed. Stakeholder buy-in during this step is crucial. In order to achieve support from company executives, small incremental changes that produce fast but satisfactory results should be made to the e-commerce websites.

The next Section argues for the suitability of the Web Analytics Process to provide a holistic approach to student-centric interventions in online education platforms.

5.3.2 Learning Analytics Process for Higher education

In the same way that businesses can use web analytics to learn more about their customers, higher education institutions can learn more about their students through learning analytics. Learning analytics refers to the measurement, collection, analysis and reporting of data about learners and their contexts to optimise learning, and their learning environment (Siemens, 2013). Optimising the online environments may improve the acquisition of prospective students, and the development and retention of current students. Prospective students require ongoing support from the moment of application and registration. Students require holistic development throughout their studies until graduation. Beyond graduation, keeping contact with alumni provides a useful resource for the development of future graduates. With such numerous stakeholders using higher education online platforms with different goals, it is clear that a systematic and coordinated learning analytics effort is needed. A student-centric learning analytics process should provide the right intervention at the right time and an appropriate level of support. These interventions should be based on the diverse needs, backgrounds, skills, contexts and goals of individual learners.

The steps in the proposed learning analytics process uses the same keywords as that of the web analytics process. However, the application of the steps is discussed in the context of higher education online platforms and not to commercial websites.

5.3.2.1 Define education Goals

There are numerous types of online platforms in higher education. Some examples include:

- An external website that provides information and recommendations to prospective students
- Staff and student portals that provide relevant information to employees and registered students respectively
- Administrative systems for academic staff to maintain student marks
- Social sites to interact with alumni
- Learning Management Systems that deliver online facilitation of modules

Typically, someone will be accountable for the maintenance and optimisation of the online platforms that they control. It is the role of this decision maker to define the goal of the platform and to identify optimisation goals for improving the website. The following, non-exhaustive, list provides an idea of the kinds of goals that may be relevant to education (Chatti et al., 2012):

- Monitor and analyse the level of engagement of staff/current students/potential students/alumni on the various platforms
- Predict learner knowledge based on past and current activities and performance
- Predict future performance based on past and current activities and performance
- Identify the need for additional tutoring
- Provide academic staff with insight into learner interest or their learning contexts
- Provide an adaptive or adaptable learning environment
- Provide recommendations on qualifications
- Provide recommendations of the most appropriate learning path
- Advise current students and alumni of trends in the job market

5.3.2.2 Build Key Performance Indicators:

Decision makers who are responsible for initiating optimisation initiatives on their online platforms need relevant information for each project. KPIs must be set to measure whether the website goals are achieved. The following is a non-exhaustive list of metrics that form part of KPIs relevant to the context of education:

- Prospective: Number of requests for information on qualifications
- Alumni: Graduate placement figures, Number of graduates, Length of time before graduates find employment in their field of study
- Current students: Student marks and class averages, Duration and frequency of online learning activities, Student satisfaction

Table 5.1 illustrates the process of defining a KPI.

Table 5.1 Process for Establishing KPIs

Questions	Example Responses
(1) What is your desired outcome?	Goal: Improve the recommendations on qualifications given to prospective
(2) Why does this outcome matter?	Matching prospective students with appropriate qualifications would improve their chances of successfully graduating
(3) How are you going to measure progress?	KPI: Number of current students satisfied with their choice of qualification must be 80% or more
(4) How can you influence the outcome?	Use machine learning algorithms to base recommended qualifications on school results and aptitude tests
(5) Who is responsible for the business outcome?	Faculty Officers
(6) How will you know you’ve achieved your outcome?	Student satisfaction questionnaires
(7) How often will you review progress towards the outcome?	Annually

When evaluating the KPI ("Number of current students satisfied with their choice of qualification must be 80% or more") based on the SMART criteria, the following holds:

- *Specific*: The objective is to improve the number of students satisfied with their choice to over 80%
- *Measurable*: Progress towards the goal can be measured through questionnaires
- *Attainable*: By improving the method of making recommendations, the goal is realistically attainable
- *Relevant*: The goal of registering learners for appropriate courses is relevant to a higher education institution
- *Timeframe*: By measuring the goal annually, continual improvements can be made to the recommendation process

5.3.2.3 Collect Data

There are numerous data sources for metrics relevant to higher education. This may include:

- Staff/student portal analytics
- Learning Management System log files
- Academic Administration System databases
- Online surveys
- Social media interactions

The goal of the optimisation intervention and the KPIs defined will ultimately dictate the source of the data. Decision makers in charge of online platforms are encouraged to carefully brainstorm all potential metrics towards measuring the defined KPIs and to identify appropriate sources of these metrics.

5.3.2.4 Analyse Data

It will often be necessary to pre-process the data before analysis can start. After pre-processing, standard data analysis techniques may be applied. This may include, but is not limited to: (1) descriptive or inferential statistics; (2) information visualisation; (3) data mining; and (4) social network analysis (Chatti et al., 2012).

(1) Learning Management Systems, for example, collect usage metrics such as a timestamp of user actions. This can be converted to useful data like time spent online or frequency of assignment submissions, among others. Simple sorting

and filtering could also be enough to measure KPIs. When **descriptive statistics** are insufficient or inappropriate, **inferential statistics** may provide suitable answers.

(2) In some cases, numeric data may hide interesting trends that can be revealed through data **visualisation** techniques – charts and plots, maps or 3D visualisations.

(3) Beyond basic descriptive or inferential statistics, more advanced **data mining** techniques may be necessary for a deeper analysis. This may include techniques like prediction, feature engineering, item response theory, association rule mining, clustering or factor analysis (Section 4.2.1).

(4) Social Network Analysis quantifies relationships between individuals or organisations. This may be combined with visualisation techniques to reveal insightful patterns that may be required to measure goal attainment.

If the data analysis step revealed the required information about the user of the education website, the optimisation intervention can be implemented. If data is missing, it will be necessary to return to the data collection step. If new hypotheses emerged after original data analysis, and these need to be tested, the model proposes a fifth step to investigate alternatives.

5.3.2.5 Test Alternatives

In this step, it may be necessary to augment data from alternative sources or to use a different technique for data analysis. This is the step where a small tweak may be done, and a pilot test can be conducted to monitor the change in online behaviour of the website users.

5.3.2.6 Implement Insights

As soon as the behaviour of the website user is understood, the optimisation intervention can proceed. The following are the types of activities that can be implemented in online academic platforms:

- The learning environment can be optimised
- Students deemed at risk can receive proactive assistance
- New students can receive the necessary orientation and introduction to modules
- Career guidance can be provided
- Feedback can be provided on the learning process

- Learning objects can be presented and sequenced adaptively
- Recommendations to foster self-directed learning can be provided
- Self-reflection in students can be promoted

On completion of the final step, the loop closes, and a new project can be initiated.

5.3.3 Differentiated Instruction Based on a Dynamic Learner Profile

This Section illustrates how the Learning Analytics Process model can be adapted when the goal is to enable differentiated instruction (Figure 5.3). The model is derived by integrating the learning design and learner modelling phases from adaptive education systems, the abstracted learning design layers (Section 3.3.2.2), the steps from the learning analytics process (Section 5.3.2), and principles of an ethical learning analytics code of conduct (Section 4.2.2).

The aim of building a learner profile is shared by researchers who create automated adaptive education systems (AES). Two core phases of a typical AES are the learner modelling phase during which the learner profile is built, and an adaptation phase during which instruction is tailored towards unique learner profiles. Generally, in an AES the profile is built in real time and adaptation is provided automatically. In some systems, like an adaptable education system, the learning platform can be set up in such a way that the learner has control over some aspects of the learning design. In other cases, educators can be provided with information about their learners, which enables them to implement differentiated interventions when needed. To provide teachers with appropriate information that can be acted upon, raw data needs to be distilled into a format backed by educational theory. This information about the learner will form the content of a learner profile.

The content of a learner profile varies. One type of profile is based on learning styles (Section 4.3.2). Numerous learning style based adaptive education systems (LSAES) build a learner profile through a questionnaire associated with a particular learning style. Several problems have been identified with the use of questionnaires to build a learner profile (Section 4.3.1.1). To overcome these problems, an implicit learner modelling technique that enables dynamic profile building is proposed (Section 4.3.1.2). Implicit modelling requires careful consideration of the ethical issues arising from learning analytics initiatives (Section 4.2.2), leading to a potential need for negotiated learning modelling and an open learner modelling approach (Section 4.3.1.3).

Learning design involves the creation or discovery and reuse of suitable Learning Objects (Section 3.3.1). Learning design is accomplished in layers, each layer building on the previous one (Section 3.3.2.2). The following abstracted layers simplify the complex task of learning design:

- A domain layer where module outcomes are defined
- A goal and constraint layer where pedagogical rules are applied based on the learner knowledge and goals
- A learner model layer where content and techniques to build learner profiles are established
- A resource layer where Learning Objects are created or repurposed and tagged with appropriate educational metadata
- A course layer where Learning Objects are sequenced appropriately based on learners' unique characteristics
- A validation layer that bases all learning design choices on educational theory

In differentiated instruction, the updated learner profile is used to inform tailored learning design choices. Differentiation is enabled through tailoring Learning Objects in the course layer based on the rules defined in the goal and constraint layer (Section 3.3.2).

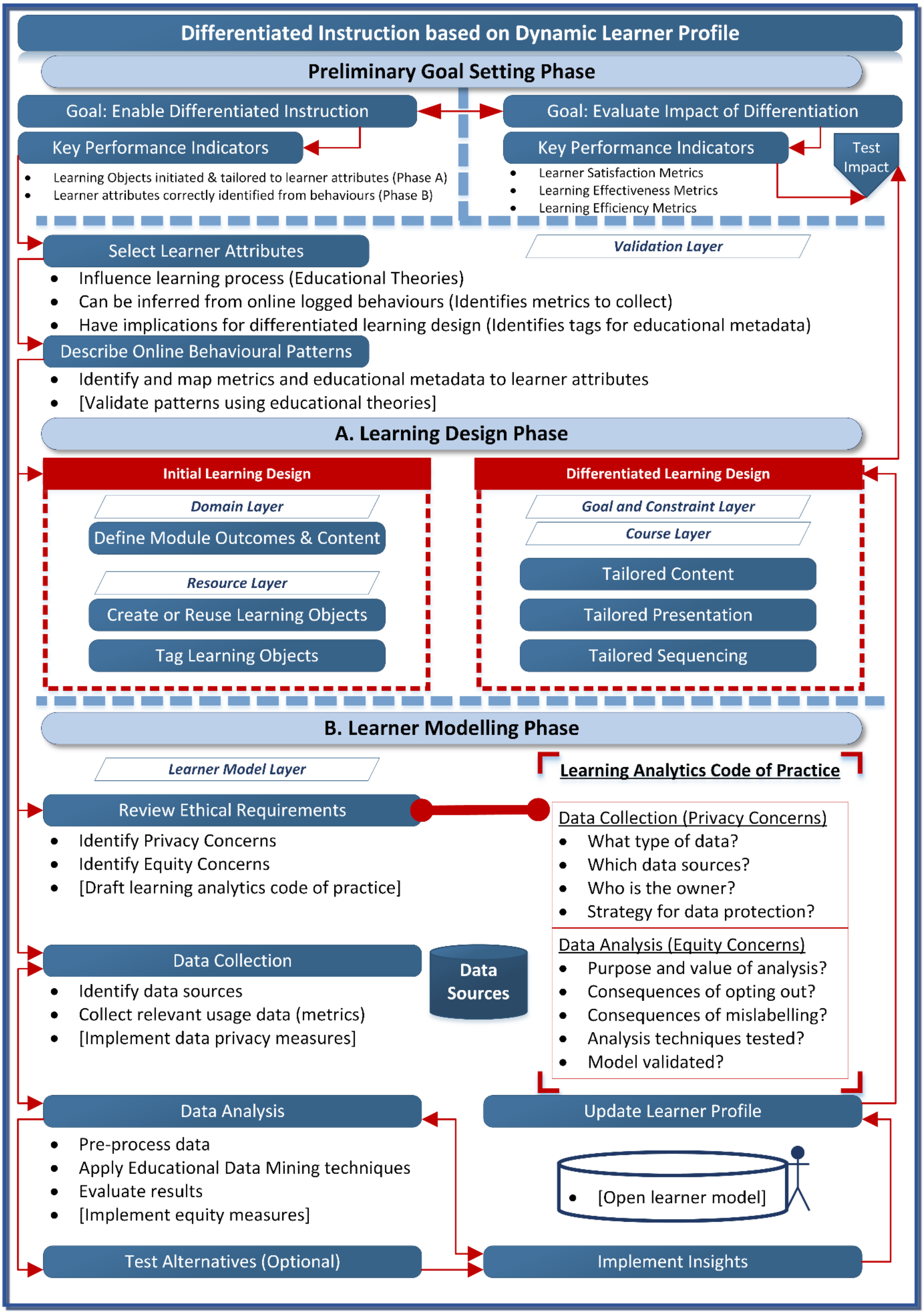


Figure 5.3 Differentiated Instruction Based on a Dynamic Learner Profile (Own Construction)

5.3.3.1 Preliminary Goal Setting Phase

One of the abstracted learning design layers is the *validation layer* that proposes any learning design choice should be backed by recognised educational theories. One such theory, or group of theories, is the identification of learning styles and the tailoring of instruction based on unique learner attributes associated with the learning style model. This thesis proposes a pragmatic approach of identifying relevant attributes from multiple learning style theories (Section 4.3.2.1).

Identify goal and set Key Performance Indicators: Enabling differentiated instruction is a goal compatible with the use of learning style theories to tailor instruction to the unique needs of learners. The generic goal proposed in the model (Figure 5.3) is therefore enabling and optimising differentiated learning design. Since differentiated instruction shares common phases of tailored learning design and learner modelling with adaptive education systems, two sub-goals are identified:

- Correctly identifying relevant learner attributes from learners' online behaviours
- Appropriately tailoring instruction based on the identified learner attributes

Select learner attributes and describe online behavioural patterns: One of the biggest challenges when integrating learning styles into adaptive learning systems is the selection of an appropriate learning style theory. Mounting criticism from some dissenting voices (Cook, 2012; Coffield et al., 2004; Kirschner, 2016) is pointing to theoretical incoherence, conceptual confusion, lack of scientific basis and seemingly never-ending overlapping characterisation of learner attributes. Further criticism is levelled at the questionnaires used to determine student attributes. This thesis proposes that instead of focusing on the model of one particular theorist, we focus instead on the student attributes defined in various learning style theories. By limiting the adaptive education system to only one learning style theory, we may be missing out on other attributes with an equally significant impact on teaching and learning. The following criteria should be applied to the selection of suitable attributes:

- The learner attributes must influence the learning process in some way, based on an educational theory
- The learner attributes must have implications for differentiated learning design
- It should be possible to infer the learner attributes from metrics that represent online logged behaviours

The focus on collecting and analysing patterns of students' online behaviours to build a learner profile dynamically is precisely in response to the criticism against the use of questionnaires to determine student attributes. When using implicit learner modelling techniques, relevant metrics must be identified that describe the online behaviour of the learner. These metrics must be mapped onto the chosen learner attributes validated by an existing educational theory.

5.3.3.2 Learning Design Phase

The learning design phase consists of two sub-sections, one performed before learner modelling (initial learning design) and one initiated in response to changes in the learner profile (differentiated learning design).

Initial Learning Design: During the initial learning design phase, the focus is on the domain layer and the resource layer.

For the **domain layer**, a theoretically sound online instructional design process should be followed to create a significant student-centric learning experience (Section 3.2). Module outcomes need to be defined and matched with suitable content. At this stage the content will be described and later instantiated when the focus shifts to the resource layer. A detailed description and application of online instructional design processes is beyond the scope of this thesis. The initial learning design can be represented in the form of a domain ontology.

The input for the **resource layer** comprises the learner attributes defined in the goal setting phase and the concepts represented in the domain ontology. The Learning Objects presented to the students in the online environment should be linked to the stated module outcomes, and be based on the pedagogic needs associated with the selected attributes. These Learning Objects must be tagged with educational metadata to record the teachers' pedagogic intention. The Institute of Electrical and Electronics Engineers Learning Object Metadata (IEEE LOM) standards (Section 3.3.1.3) provides a suitable vocabulary for educational metadata.

Differentiated Learning Design: While learners navigate the course material, the learner modelling phase continuously updates a learner profile. This profile provides the input into the differentiated learning design sub-Section. During differentiated learning design the focus is on the goal and constraint layer and the course layer.

Learning rules are created in the *goal and constraint layer*. Pre- and post-conditions based on the learner profile are overlaid on the domain ontology. These rules influence the sequencing, content and presentation of Learning Objects. Learning objects are differentiated based on pre-requisite knowledge, learner goals, and cognitive and affective needs contained within the learner profile.

Rules for differentiated learning design based on learner attributes can be represented using IF statements of the format in Equation 5-1 (Popescu, 2008):

Equation 5-1 Differentiated learning rules template

$IF\ Attribute\ THEN\ Action\ Object\ Value$, **where**

- Action = Sort | Dim | Hide | Highlight | Trigger | Show
- Object = Metadata tag of Learning Object | UI element
- Value = Value of Metadata tag

The metadata tags of Learning Objects and their associated values are linked to the fields and values from the educational category of the IEEE LOM standard (Section 3.3.1.3). The list of actions suggested above is not exhaustive. The mentioned actions are illustrative of the typical type of techniques used in adaptive education systems to tailor Learning Objects.

- “Sort” represents the sequencing of LOs/UI elements
- “Dim” represents greying out or disabling a LO/UI element (e.g. button/hyperlink)
- “Hide” represents the removal of a LO/UI element
- “Highlight” represents a recommendation of a particular LO/UI element
- “Trigger” represents an action such as the sending of an automated message
- “Show” represents displaying LO/UI element (e.g. table of contents/annotation)

A learning pathway based on learner knowledge and goals can be represented through directed acyclic graphs (DAG) (Karampiperis and Sampson, 2004). In a DAG, the vertices represent the concept to be learned and the edges between each vertex represent the relation between each concept. Typical relations can be classified as:

- Is part of / Has part
- References / Is referenced by
- Is based on / Is basis for
- Requires / Is required by

The learning rules designed in the goal and constraint layer are implemented in the **course layer**. The Learning Objects from the resource layer are tailored according to the rules defined in the goal and constraint layer. The Learning Objects can be differentiated on their sequence (Action: Sort), content (Actions: Dim, Hide, Highlight, Trigger, Show) or the presentation UI. The chosen educational theory will determine the form of the actions to be taken based on the learner attribute. Any tailored Learning Object must still guide the learners towards the same learning outcomes defined in the domain layer. Based on the prerequisites represented in a DAG, the learners' knowledge as stored in the learner profile can be used to recommend the next concept in the predesigned learning pathway. A dynamic learner profile is necessary to inform the differentiated learning design choices. The learner profile will be updated during the learner modelling phase, described next.

5.3.3.3 Learner Modelling Phase

The techniques recommended for the learner modelling phase proposed in this thesis are predominantly based on implicit and negotiated modelling (Sections 4.3.1.2 and 4.3.1.3), but it does not preclude explicit modelling techniques (Section 4.3.1.1) if and when required. The steps are based on the learning analytics process model derived in Section 5.3 and the learning analytics code of ethical practice described in Section 4.2.2. Incorporated into the learner modelling phase are activities and techniques associated with learning style based adaptive education systems (Section 4.3) and educational data mining (Section 4.2.1).

Review Ethical Requirements: Any learning analytics initiative must be conducted ethically (Section 4.2.2.1) and practitioners must carefully address privacy (Section 4.2.2.2) and equity (Section 4.2.2.3) concerns. To ensure buy-in from learners, their privacy must be guaranteed during data collection and they must be convinced that the benefits that will accrue from the data analysis outweigh potential risks. A learning analytics code of practice must be drafted and used to acquire informed consent from all participants whose data will be analysed and used for changes to the learning design. This code of practice must incorporate principles of ethical research, i.e. respect for persons, beneficence and justice.

Data Collection: During the data collection step, metrics identified during the goal setting phase must be collected. All potential data sources that may supply these metrics need to be identified. In implicit modelling (Section 4.3.1.2), these metrics represent learner cognitive and affective behaviours linked to learner attributes associated with educational theories (Section 4.3.2). In explicit modelling, data can be elicited directly from learners responding to questions (Section 4.3.1.1). During data collection all privacy measures as drafted in the learning analytics code of practice must be implemented.

Data Analysis: Learner attributes as identified during the goal setting phase are inferred during the data analysis step. For differentiated instruction, data can be analysed on demand as needed by the lecturer (Jugo, Kovačić and Tijan, 2015). The goal and the nature of the raw data collected in the previous step will determine the sequence of activities in the data analysis step. It may be possible, for example, to use simple inferential statistics if inferences and predictions are to be made on a small dataset. More complex goals and large datasets may require more advanced educational data mining techniques, such as described in Section 4.2.1 and summarised in Table 5.2.

Table 5.2 Educational Data Mining Techniques

Goal	Technique
Predictive Modelling	• Classification • Latent Knowledge Estimation • Regression
Structure Discovery	• Clustering • Factor Analysis • Social Network Analysis
Relationship Mining	• Association Rule Mining • Correlation Mining • Sequential Pattern Mining • Causal Data Mining
Distillation of data for human judgement	• Data Visualisation • Text Mining

Large datasets from disjoint sources may require pre-processing to prime data for analysis. Pre-processing can include data cleaning, integration, reduction or transformation. It is beyond the scope of this thesis to report on all possible pre-processing techniques, but the following serve as illustration to the potential strategies commonly applied to data mining (summarised in Figure 5.6):

- **Data cleaning** is responsible for removing inconsistencies and errors in the data. For example, there may be missing values, noisy, i.e. meaningless or unstructured data, outliers or inconsistent data.
- **Data integration** is responsible for consolidating data from multiple disjoint data sources. Learners frequently need to consult resources outside of the learning environment or perform offline activities; or biometric data needs to be integrated with online behavioural metrics to measure affect, for example. Metrics may, therefore, come from several sources and need to be combined in a sensible way.
- **Data reduction** focuses on deciding which data features to include or exclude for analysis. The aim of data reduction is to find a smaller dataset that can produce similar analytical results. Data reduction can be performed through several techniques such as:
    - Aggregation – combining two or more attributes
    - Sampling – selecting a subset from the population
    - Feature subset reduction – removing redundant or irrelevant features
- **Data transformation** converts data into a different format. Common techniques to transform data include:
    - Normalisation – scaling values into a pre-determined range
    - Smoothing – the removal of outliers
    - Aggregation – preparing data into a summarised format
    - Discretisation – mapping raw numeric data points onto interval or conceptual labels
    - Generalisation – substituting data points into hierarchical layers

When data is ready, analysis can proceed through a suitable educational data mining technique from Table 5.2 (summarised in Figure 5.6).

Data pre-processing and analysis is concluded by evaluating the results of the analysis. Evaluation methods (summarised in Figure 5.6) will depend on the data mining technique used and is necessary to measure the quality of the learner model that result from the data analysis.

When **predictive modelling techniques** are used, the performance of the classifier must be measured in terms of error rate. That is the proportion of the incorrect classifications made over the whole dataset. Common methods include cross-validation (hold-out test, k-fold cross-validation, random sub-sampling, leave-one-out method), confusion matrix and Receiver Operating Characteristic (ROC) curves.

Cross-validation involves splitting data one or more times into a training set and a test set. The training set is used to develop the model and the test set is used to estimate the risk of the algorithm.

- In the hold-out test, cross-validation is performed on a single split of the data.
- In the k-fold cross-validation is performed by splitting the dataset into k subsets and repeating the hold-out test k times. In each iteration, k subsets are the training set and k-1 subsets the test set.
- In the random sub-sampling method, the hold-out test is performed several times.
- In the leave-one-out method, k represents all data sets and a prediction is made for the one data point not in the training set.

A **confusion matrix** uses a contingency table to present the number of true and false positives and false and true negatives (Figure 5.4).

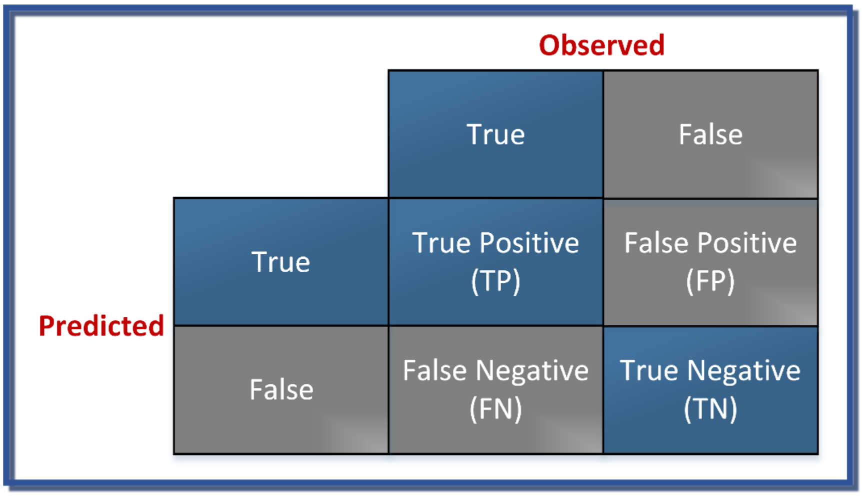


Figure 5.4 Confusion Matrix

Confusion matrices show number of hits, correct rejections, false alarms and misses.

- Hits: true positives (TP) where the true predicted classification matches the observed classification
- Correct rejections: true negatives (TN) where false predicted classifications match false observed values

- False alarms: false positives (FP) where a true prediction should have been false
- Misses: false negatives (FN) where a false prediction should have been true

A confusion matrix uses three metrics to measure performance of the predictive model: accuracy, precision and recall (Equation 5-2).

Equation 5-2 Metrics from Confusion matrix to Evaluate Predictive Models

$$Accuracy = \frac{TP + TN}{TP + TN + FP + FN}$$

$$Precision = \frac{TP}{TP + FP}$$

$$Recall = \frac{TP}{TP + FN}$$

Equation 5-3 Metrics from ROC Curves

$$TPR = \frac{TP}{TP + FN}$$

$$FPR = \frac{FP}{FP + TN}$$

ROC curves visually represent the information from a confusion matrix (Figure 5.5).

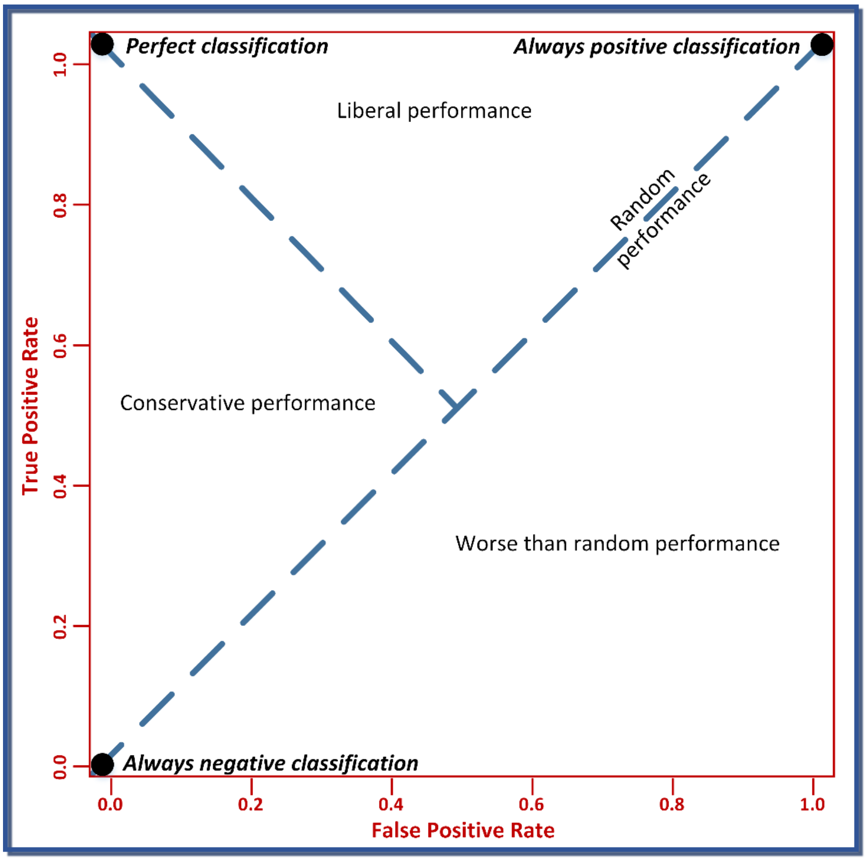


Figure 5.5 ROC Curve

The graph plots True Positive Rates (TPR) and False Positive Rates (FPR) on two-dimensional axes, where TPR and FPR are metrics as shown in Equation 5-3.

When **relationship mining techniques** are used, measures to evaluate the resulting rules include significance (through examining minimum thresholds on confidence and support) and appeal (through examining lift or conviction).

Measures of significance (Equation 5-4):

Equation 5-4 Measure of Relationship Rule Significance

$$Confidence: conf(X \Rightarrow Y) = \frac{supp(X \cup Y)}{supp(X)}$$

Where supp(X) is the proportion of transactions in the dataset which contains X.

Measures of appeal (Equation 5-5):

Equation 5-5 Measure of Relationship Rule Appeal

$$Lift: lift(X \Rightarrow Y) = \frac{supp(X \cup Y)}{supp(Y) \times supp(X)}$$

$$Conviction: conv(X \Rightarrow Y) = \frac{1 - supp(Y)}{1 - conf(X \Rightarrow Y)}$$

All equity measures as drafted in the learning analytics code of practice must be implemented during the data analysis step. Evaluating the performance of the model generated by the data mining technique will ensure the analysis results are validated. Validation of the model empowers teachers to take appropriate action and apply the intervention consistently to all participants.

All pre-processing, data mining and evaluation techniques described in this Section are summarised in Figure 5.6.

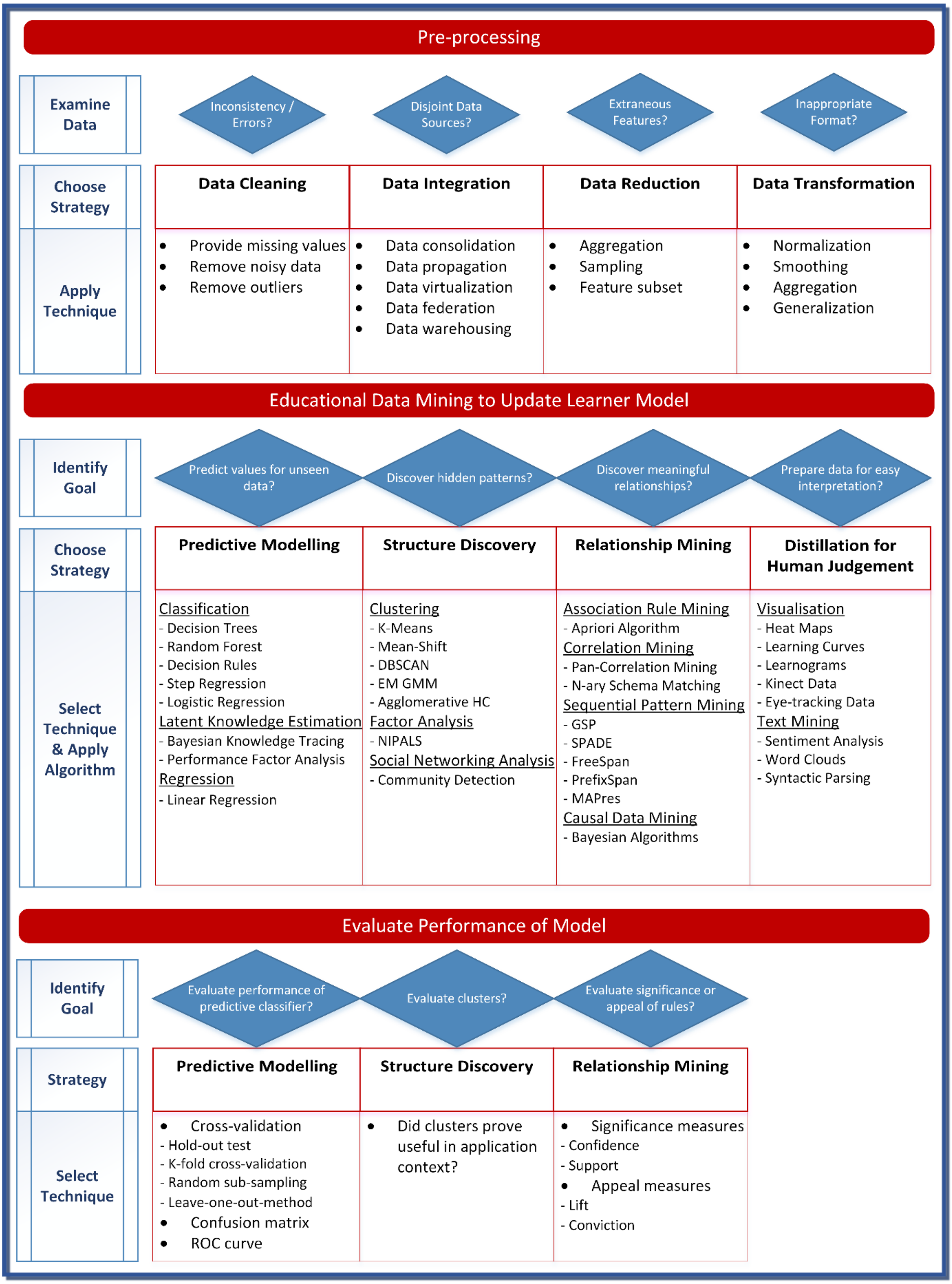


Figure 5.6 Overview of Data Analysis Techniques (Own Construction)

Test Alternatives and Implement Insights: The educational data mining step may reveal unexpected results that need further investigation. The proposed model allows an optional step to generate new hypotheses that may require:

- Exploration of different data sources
- Addition of new attributes/features
- Application of different educational data mining techniques, for example
    - Trying different algorithms
    - Tweaking clusters
    - Using the results of one analysis technique as input into another
- Applying negotiated learner modelling to seek the learners' approval of the conclusions made in the data analysis step
- Making a quick change to the learning design and conducting a small-scale pilot study to measure the effect of the change

Once satisfactory results are achieved, the necessary action can be taken ("Implement Insights"). This involves a two-part process:

- Updating the learner profile with inferred information
- Initiating the differentiated learning design in response to the changes in the learner profile (Section 5.3.3.2)

Evaluation of the impact of the differentiated learning design on learner satisfaction, learning effectiveness and efficiency closes the process model loop. This step is represented in the model as an off-page reference, since this step is yet to be modelled as part of future work.

5.3.4 Critical Reflection on Sub-Objective 1

In terms of the methodology developed in Section 2.3.3, the prototyping phase of design research for technology enhanced learning must argue the consistency of the global design of the proposed solution. Chapter 5 reports on the prototyping phase of the study that the thesis is based on, and hence the need to reflect on the proposed model to ensure the phases and steps described in Figure 5.3 are logically designed, viable and conform to the existing knowledge base. Each phase of the proposed model will be critically reflected on next.

5.3.4.1 Consistency of Goal Setting Phase in the Proposed Model

In the model proposed in Figure 5.3, the explicit inclusion of a goal setting phase at the start of any learning analytics initiative is consistent with both the Web Analytics Process (Section 5.3.1) and the need for a validation layer backed by educational theory (Section 3.3.2.2). In this thesis, the proposal for using multiple learning style theories to inform profile building and learning design stems from researchers and practitioners developing learning style based adaptive education systems (Section 4.3 and Section 3.3.2).

Table 5.3 Consistency of Key Performance Indicators

Questions	Answers
What is your desired outcome?	• Attribute X identified in Learner A • Learning design optimised for Learner with Attribute X
Why does this outcome matter?	Learning design choices are validated by a recognised educational theory
How are you going to measure progress?	Evaluating the effect of the learning design choice on: learner satisfaction, learning effectiveness or efficiency
How can you influence the outcome?	• For learner modelling: Suitable data sources must be identified and analysed using relevant techniques • For learning design: Learning objects must be tailored appropriately and provided on a suitable learning platform
Who is responsible for the business outcome?	Teachers, instructional designers
How will you know you've achieved your outcome?	If it can be proven that the learning design, that is based on a learner profile, had a positive effect on learning
How often will you review progress towards the outcome?	Choice will vary based on context

With the goal identified in this thesis as enabling differentiated instruction and optimising the learning design based on learner profiles, the Key Performance Indicators are linked to the two sub-objectives of the learner modelling phase and the learning design phase. These proposed KPIs are consistent with the key questions (Table 5.3) and SMART objectives (Table 5.4) described in Section 5.3.2.2.

Since the outcome of the learner modelling phase is a learner profile of attributes from selected learning style theories, a generic KPI for a successful learner modelling exercise can be "Attribute X is identified in Learner A". Similarly, the generic KPI to measure a successful learning design phase can be "Learning design is optimised for a learner with Attribute X". A model for evaluating optimal learning design is beyond the scope of this thesis, but the step is included as part of the goal setting phase of the proposed model (see the "Test Impact" shape on Figure 5.3). This impact study is necessary to measure progress in improving the learning design.

To conclude the learner modelling phase successfully, suitable data sources and metrics must be identified and analysed to build a learner profile. To conclude the learning design successfully, Learning Objects must be tailored and presented through the selected online learning platform that presents the learning design. If the learning design has a positive influence on learning and it can be shown through the impact study that this positive impact can be attributed to the learning design based on the learner profile, we can conclude that we have achieved the desired outcome. KPIs are evaluated with SMART criteria, i.e. they must be Specific, Measurable, Attainable, Relevant and Timely (Table 5.4).

For this thesis, the main goal is enabling differentiated instruction in a Learning Management System (LMS), specifically examined in Moodle, but the proposed model is generic enough to apply to any LMS. At this stage the assumption, backed by existing literature, is that differentiated instruction based on a dynamic learner profile will have a positive outcome on learning. This assertion must be validated through an impact study following steps that still need to be modelled as part of future work.

Table 5.4 Evaluating KPIs Based on SMART Criteria

Criteria	
Is your objective Specific?	Enable differentiated instruction in a Learning Management System
Can you Measure progress towards that goal?	Impact evaluation
Is the goal realistically Attainable?	See Section 5.4 and Section 5.5
How Relevant is the goal to your organisation?	Academic institutions are responsible for creating optimal learning environments
What is the Timeframe for achieving this goal?	Learning environments should be continuously evaluated and improved

The mapping of behavioural patterns onto learner attributes is supported by pedagogical practices in technology enhanced learning (Figure 4.8), and successfully implemented in learning style based adaptive education systems (Section 4.3.2.2).

5.3.4.2 Consistency of the Learning Design Phase in the Proposed Model

The learning design phase in the proposed model is split into two sub-sections to reflect the initial learning design before learner modelling and the updated learning design in response to changes in the learner profile.

The term "differentiated learning design" is used to distinguish the tailoring of the learning environment from other types (Table 3.5). For example, the learning design in this model will not be automatically tailored in real time, as is the case in adaptive and personalised learning systems. Instead, the lecturer will be able to create a predesigned learning pathway, group learners with similar characteristics together and recommend Learning Objects tailored towards these groups. In addition, UI elements such as special hyperlinks, table of content, sequential navigation buttons, annotations or any relevant element as dictated by selected learner attributes can be tailored for different groups. Differentiation can be done at any time as decided by the lecturer. The model, though, can be extended in future to include automation. The decision not to incorporate automation yet is in response to reservations regarding the use of learning style theories (Section 4.3.2.1). Several studies have been inconclusive about the impact of the provision of instruction based on learning styles. By making recommendations on suitable Learning Objects or the sequence in which they should be attempted, learners will feel more in control of their own learning. In the proposed model, learners should still have a choice of following the recommendations or not. This type of learner control is consistent with a negotiated learner model (Section 4.3.1.3) and with the ethical principle of respect for persons as prescribed in the Belmont report (Section 4.2.2.1). Furthermore, any future experimental studies to evaluate the differentiated learning design will be augmented with data showing whether learners followed or ignore the recommendations. Over time, such evidence will strengthen the confidence that appropriate learning design choices are being made based on the learner profile and automation can be gradually introduced.

In differentiated learning design, the lecturer sets the same learning outcome and pace for all learners, in contrast to individualised/self-directed learning (Table 3.5). Even though it is not explicitly mentioned in the proposed model, the learning environment may include elements that can be manipulated by learners themselves, as is the case in adaptable systems.

Incorporating the domain layer, resource layer, goal and constraint layer and course layer is consistent with the abstracted steps to simplify learning design (Figure 3.17). The domain and resource layers provide the base design relevant to all learners, while the differentiation is designed in the goal and constraint layer and implemented in the course layer.

This type of layered design allows lecturers and instructional designers to focus on relevant activities at different times during the learning design process. It should be noted that the goal and constraint layer, responsible for designing learning rules, is conceptually placed inside the differentiated design layer. However, temporally, the design activities are performed prior to the learner modelling phase. The course layer implements the rules devised in the goal and constraint layer in response to changes in the learner profile.

5.3.4.3 Consistency of the Learner Modelling Phase in the Proposed Model

The entire learner modelling phase is consistent with the aim of the learner model layer as described in the layered learning design model (Figure 3.17), i.e. building and updating a learner profile.

The learner modelling phase integrates steps of a learning analytics process model (Section 5.3.2), derived from a web analytics process model (Section 5.3.1), with an ethical learning analytics code of conduct (Section 4.2.2). The learner modelling phase includes steps evident in existing learning analytics process models (Section 4.2.3), i.e. “Data Collection”, “Data Analysis” and “Implement Insights” where action is taken based on the results of the data analysis. The optional “Test Alternatives” step is included from the web analytics process model.

The addition of the “Review Ethical Requirements” step is consistent with the principles for ethical research prescribed in the Belmont report (NCPHS, 1979), i.e. respect for persons, beneficence and justice. The ethical learning analytics code of practice is designed in the “Review Ethical Requirements Step” and implemented during the rest of the learner modelling phase. Even though the implementation of data privacy measures is explicitly mentioned under the “Data Collection” step, it does not mean that this is the only step where privacy should be maintained. Naturally, privacy should be maintained throughout the entire initiative. Placing the issue of data privacy under “Data Collection” is in acknowledgement of the need for data protection predominantly at the source of the data. Similarly, explicitly mentioning implementation of equity measures under “Data Analysis” does not preclude equity measures from being implemented when differentiated instruction is implemented. The implementation of the measures designed in the learning analytics code of conduct is one of the decisions to be made when ethical requirements are reviewed. The questions presented in Figure 5.3 do not form an exhaustive list, but are representative of some of the main privacy (Section 4.2.2.2) and equity (Section 4.2.2.3) concerns raised by learners during learning analytics initiatives. More questions are likely to arise when a contextual analysis is performed of the site where the learning analytics initiative will be implemented. The “Open Learner Model” is recommended from the concept of negotiated learner modelling (Section 4.3.1.3) and addresses equity concerns the learners may have (Section 4.2.2.3).

In this study, the next two iterations involve refining the steps proposed in the learning design phase (Section 5.4) and learner modelling phase (Section 5.5) to enable differentiated learning design in a Learning Management System.

5.4 Sub-Objective 2: Differentiated Learning Design in Moodle (Iteration 3)

With the global design of the proposed model prepared and evaluated for consistency against the existing knowledge base (Section 5.3), the focus of the study shifts to the two sub-components. This Section refines and evaluates the learning design phase for expected practicality and effectiveness. To instantiate the learning design phase in Moodle, the following questions are explored:

- What are the typical behavioural patterns associated with learner attributes selected from learning style theories? (Section 5.4.1)
- How should the initial learning design be implemented? (Section 5.4.2)
- How can the desired differentiation be achieved in a Learning Management System? (Section 5.4.3)

Moodle was selected as the Learning Management System based on its widespread adoption, 75,830 sites over 232 countries (Moodle, 2018), as well as the fact that it is the platform of choice at the Nelson Mandela University where the learning design is instantiated for evaluation. Through Moodle, educators and online learning designers have access to a wide range of customisable resources and activities with which to design an online course. Moodle, however, is not naturally developed to be used as an adaptive education system. Consequently, there is a need to verify whether Moodle has the potential to provide differentiated instruction (Section 5.4) based on changes to attributes stored in a learner profile (Section 5.5).

5.4.1 Goal, KPIs, Learner Attributes and Online Behaviours

With the pedagogic goal defined as adapting learning design towards the needs of learners based on their profile, suitable attributes must be selected in the preliminary goal setting phase of the proposed model (Figure 5.3). These attributes must be validated from educational theory for implications in an online learning environment.

Section 4.3.2.1 established potential attributes from existing learning style theories. These attributes have been selected and proposed based on their hypothesised implications on learners' cognition or affect. For each attribute, behavioural patterns in online learning can potentially be identified. Sections 5.4.1.1 and 5.4.1.2 present these attributes and hypothesises matching behavioural patterns that learners are likely to exhibit when navigating online learning material. A technique to confirm these hypotheses is illustrated in Section 5.5.

5.4.1.1 Learner Cognitive Attributes and Behaviours with Potential Implications for Online Learning

Table 4.6 lists potential attributes motivated by various learning style theories (Section 4.3.2.1). Examination of the definition of these attributes in the original theory reveals potential behaviours in the cognitive domain relevant in the online learning context. These attributes are summarised in Table 5.5 and Table 5.6 and grouped together where behaviours overlap. Behaviours are broken into their constituent parts, i.e. the metric extracted from a Learning Management System log and a description of the relevant learning resource or UI element. The vocabulary used to describe the learning resource stem from IEEE LOM (Section 3.3.1.3). At this stage the behaviours are hypotheses which will need further investigation during the data analysis step.

Perceptual Modality: This dimension refers to the preferred media type through which learners best extract information.

- **Visual** learners prefer to learn from Learning Objects that are predominantly graphical in nature, e.g. still images or dynamic animations or videos.
- **Verbal/Aural** learners prefer to learn from Learning Objects that include audio narration, such as podcasts or narrated presentations.
- Learners classified as **Read/Write** prefer to learn from predominantly textual Learning Objects.
- **Kinaesthetic/Tactile** learners prefer learning through whole body movement or through manipulating objects. This can be enabled online using virtual or augmented reality.

Table 5.5 Cognitive Attributes and Associated Online Learning Behaviours (media types)

<table>
<tr><th rowspan="2">Cognitive Attributes</th><th colspan="2">Learning Behaviours regarding media types</th></tr>
<tr><th>Learning Resource\UI Element</th><th>Metric from log</th></tr>
<tr><td>Visual</td><td>• LO Media Type: Graphic / Image / Video / Animation</td><td rowspan="4">• High number of hits
• Longer duration per access
• Consistently first access to
• Better performance on assessments related to</td></tr>
<tr><td>Verbal/Aural</td><td>• LO Media Type: Audio / Podcast / Narrated Presentation</td></tr>
<tr><td>Read/Write</td><td>• LO Media Type: Written / Textual</td></tr>
<tr><td>Kinaesthetic / Tactile</td><td>• LO Media Type: Augmented Reality</td></tr>
</table>

Information Processing, Organisation and Reasoning: This dimension refers to the mental processes and affordances through which learners best make sense of new information.

- **Abstract/Deductive/Reflective** learners prefer to learn theory (e.g. facts, details, definitions) before viewing examples or applying the theoretical concepts. They explore a topic from a wide variety of sources/opinions and critically think about the theoretical concepts.
- **Concrete/Inductive/Active** learners prefer to first actively work through examples or try things out before learning about the underlying theories.
- **Serial/Linear/Field-Dependent/Synthesis** learners prefer to navigate content using previous and next buttons, breadcrumbs or in the order presented. They are distracted if additional material to related content is interspersed with core content and need more guidance and feedback while navigating through the course material. These learners need a broad overview of a topic before going into detail.
- **Holistic/Alternating/Field-Independent/Analysis** learners prefer the navigational flexibility provided by a hyperlinked table of content or a clickable concept map. They find it easy to switch between different tasks and frequently explore related content while mastering core content. These learners need to master a single concept in-depth before moving to the next one.

Table 5.6 Cognitive Attributes and Associated Online Learning Behaviours (mental strategies)

Cognitive Attributes	Learning Behaviours regarding mental strategies	
	Learning Resource\UI Element	**Metric from log**
Abstract / Deductive / Reflective	• LO Fundamental (Definition, Fact, Law, Theory)	• High number of hits • Longer duration per access • Consistently first access to LO • Better performance on assessments related to LO
Concrete / Inductive / Active	• LO Auxiliary-Illustration (Example, Counter Example, Case Study) • LO Auxiliary-Interactivity (Exercise, Exploration, Simulation)	
Serial / Linear / Field Dependent / Synthesis	• UI Element: Prev/Next • UI Element: Annotation • LO Auxiliary-Explanation (Hints)	• High number of hits • Consistently first access to LO
Holistic / Alternating / Field Independent / Analysis	• UI Element: Index • LO Auxiliary-Explanation (Additional Info)	

While behaviours in the cognitive domain reflect mental processes to acquire and make sense of new information, affective behaviours reflect feelings and emotions that dictate how learners will react to learning tasks.

5.4.1.2 Learner Affective Attributes and Behaviours with Potential Implications for Online Learning

Affective behaviours defined by learning style theories that may potentially be inferred in the online learning environment are summarised in Table 5.7. This thesis proposes techniques to infer implicitly whether learners are intrinsically or extrinsically motivated by analysing engagement metrics. Techniques to infer learners' preferences for social interaction are also proposed, since social aspects have an influence on motivation and engagement.

Motivation: This dimension dictates why learners behave in a particular way when they engage with the course material. Learners can be intrinsically or extrinsically motivated at any time during their studies. Motivation is not necessarily a fixed personality trait but may fluctuate depending on external factors such as their affinity towards the module. While a deeper interest cannot be forced, strategies may be employed to cultivate intrinsic motivation by harnessing extrinsic stimuli. While external rewards may motivate some learners, it may distract others. It is therefore necessary to continuously determine learners' current motivation to best support their progress.

- **Intrinsically** motivated learners learn for the sake of the experience and mental challenges inherent in learning. Intrinsic motivation will be evident if the learner enjoys a subject and has an innate desire to learn more.
- **Extrinsically** motivated learners are motivated by external rewards, e.g. good grades, well-paid jobs, recognition, etc.

Engagement: Motivation will influence how engaged learners are with the course material. Engagement, and therefore motivation, manifests through several attributes:

- **Deep** engagement can be distinguished in intrinsically motivated individuals who spend more time on a range of learning activities and frequent access of additional material. They tend to actively participate in online discussions and will likely have a higher performance in assessments.
- **Surface** engagement can be distinguished in extrinsically motivated individuals who will try to get through the course material as quickly as possible. Quizzes are completed by trial and error and there is a heavy reliance on practice tests to prepare for summative assessments. Performance in assessments is likely to be low to medium.
- **Strategic** engagement can be observed in extrinsically motivated individuals who seeks recognition from teachers. Frequent online interaction with the teacher will be evident. Access to formative assessments and course material will increase closer to deadlines for summative assessments. Links to additional information will largely be ignored, since they are typically not included in module outcomes and will not be assessed. Performance in assessments is likely to be medium to high.

- **Resistance** or very little to no engagement can be distinguished in learners who have no interest in the module. There is extremely low study time or access to Learning Objects, and low participation in activities. Performance in assessments will suffer as a result and is often below criteria for passing the module.
- **Meticulous** approach to submissions is evident when quizzes or assignments are submitted with careful revisions, correlated with attention to detail and high performance. These learners may take longer to complete activities as a result.
- **Careless** approach to submissions is evident when quizzes or assignments are submitted with careless errors or omissions, correlated with low performance. The time to complete activities will likely range from slow to fast.
- **High Persistence** can be detected when learners spend a longer duration engaged in learning activities, frequently returning to course material to master the content. High persistent learners also seek out feedback and resubmit quizzes or assignments to improve.
- **Low Persistence** can be detected when learners spend less time studying and fewer returns to course material already accessed. If hints are provided in quizzes, for example, there will be frequent access to this facility. Assignments or quizzes will be submitted with unanswered questions.

Table 5.7 Affective Attributes and Associated Online Learning Behaviours Linked to Engagement

<table>
<tr><th rowspan="2">Engagement Attributes</th><th colspan="2">Learning Behaviours regarding engagement</th></tr>
<tr><th>Learning Resource\Performance</th><th>Metric from log</th></tr>
<tr><td>Deep</td><td>• Entire course (High)
• LO Auxiliary-Explanation (Additional Info)
• Forum / Chat (Peers and Teachers)
• Medium to High performance</td><td rowspan="4">• Total time studying
• Total number of hits
o Per LO
o All LOs
• Time of first access
• Performance in assessments
o No. Errors
o No. Omissions
o No. Revisions
o Score</td></tr>
<tr><td>Surface</td><td>• Entire course (Low to Med)
• Low to Medium performance</td></tr>
<tr><td>Strategic</td><td>• Entire course (Med to High)
• Forum / Chat (Mostly Teachers)
• Medium to High performance</td></tr>
<tr><td>Resistant</td><td>• Entire course (Low)
• Low performance</td></tr>
<tr><td>Meticulous</td><td rowspan="2">• Learning activities
o Quiz
o Assignment</td><td>• Low No. Errors
• Low No. Omissions
• High No. Revisions
• Medium to High Performance</td></tr>
<tr><td>Careless</td><td>• High No. Errors
• High No. Omissions
• Low No. Revisions
• Low to Medium Performance</td></tr>
<tr><td>High Persistence</td><td rowspan="2">• Entire Course
• Learning Activities
o Quiz\Assignment
• LO Auxiliary-Explanation (Hints)
• LO Auxiliary-Explanation (Feedback)</td><td>• High Duration\Hits
• High No. Revisions</td></tr>
<tr><td>Low Persistence</td><td>• Low Duration\Hits
• Low No. Revisions
• High No. Hints</td></tr>
</table>

Social: Learners' motivation and engagement with the course material will be affected by their preferences regarding interaction with peers and teachers while learning. It is therefore necessary to understand learners' predilection towards social interaction.

Learners with high preference for **Individual** work will tend to avoid group work if given a choice, while those who prefer working in **Groups** will seek out opportunities to share the workload. A preference for working as individuals or groups does not necessarily correlate with a person being an **Introvert** or **Extrovert**. Introversion/Extraversion will be evident in their approach to online discussions in forums or chat with peers or teachers. Furthermore, learners can be classified as **Competitive** if they tend to compare their progress with others and if they are hesitant to offer their assistance to peers. By contrast, **Collaborative** learners will be motivated through being able to offer assistance if the opportunity presents itself.

Table 5.8 Affective Attributes and Associated Online Social Interaction Behaviours

<table>
<tr><th rowspan="2">Social Attributes</th><th colspan="2">Learning Behaviours regarding social interaction</th></tr>
<tr><th>Learning Resource</th><th>Metric from log</th></tr>
<tr><td>Individual</td><td>• Individual Assignment</td><td rowspan="6">• Content of forums / chat
• No of messages
• No of assignment type (if choice given between individual/group)
• Choice of activity type</td></tr>
<tr><td>Group</td><td>• Group Assignment</td></tr>
<tr><td>Introvert</td><td>• No online communication or Forum (only if anonymous)</td></tr>
<tr><td>Extrovert</td><td>• Forum or Chat</td></tr>
<tr><td>Competitive</td><td>• Gamified activities</td></tr>
<tr><td>Collaborative</td><td>• Peer help facility</td></tr>
</table>

The behavioural patterns identified in this Section forms the basis for any learning design choices made during the initial and differentiated learning design (described in Section 5.4). The goal of the learner modelling phase (Section 5.5) is to infer relevant attributes from these behavioural patterns recorded in LMS log files.

5.4.2 Initial Learning Design

The initial learning design phase is responsible for defining the module outcomes and establishing Learning Objects to deliver the course content. The two learning design layers associated with the initial learning design are, therefore, the domain layer for defining the outcomes and the resource layer focusing on the Learning Objects.

5.4.2.1 Module Outcomes and Content Organisation

To achieve the level of differentiated instruction as described in Section 5.4.3, it is necessary for the initial design to establish Learning Objects at the lowest possible granularity (Section 3.3.1.2). In addition, these Learning Objects must be tagged with educational metadata using vocabulary from metadata schemas such as IEEE LOM (Section 3.3.1.3). The descriptors used to tag the Learning Objects must reflect the type of learning resource, since this is one part of the online learning behaviour that will be logged by a Learning Management System.

The following module organisation is prescribed for the process proposed in this thesis (Figure 5.7): Module > Unit of Learning > Learning Object > Information Object > Media Element. Below, the role of each of these building blocks is briefly defined in the educational context, and its instantiation in Moodle introduced.

Module represents the highest level of abstraction and the coarsest granularity. This level represents the building blocks of an academic programme, with credits from a collection of modules leading to the awarding of a qualification. In Moodle, the main area where learners access their learning material is called a *course*. Moodle Version 3.3 allows for courses to be published in several formats[3]:

- Weekly – Courses with a clear start and end date and used to set up sections by week to guide all learners through the course material at the same pace.
- Topic – The default format to create a course around specific objectives. In this layout, learners see all topics on and have to scroll to the relevant topic.
- Social – A course structure built around a single main forum.
- Single activity – A format that enables a teacher to upload a Sharable Content Reference Model (SCORM) package to the course as a self-contained bundle of content and interactive JavaScript activities.

[3] https://docs.moodle.org/33/en/Course_formats

Moodle is open source and has an active community of developers that contributes new features. Some of the course formats contributed by these developers include:

- Buttons format to provide a JavaScript menu to access different sections.
- Collapsed topics format to decrease the amount of scrolling.
- Daily format to organise topics by day instead of week.
- Grid format that uses a grid of icons and clicking on an icon opens only the topic represented by the icon.

A **Unit of Learning** represents a single lesson, topic or chapter, each one linked to the module's learning outcomes. A learning outcome is a statement that describes skills, competencies or knowledge that learners must demonstrate on completion of a course. The focus of a learning outcome is higher order thinking skills and behavioural changes. Learning outcomes can be constructed using measurable verbs from Bloom's Taxonomy. In a Moodle course, each unit of learning has its own Section when the topics format is used. Section headings can be modified to fit the topic it presents.

A Unit of Learning is constructed through a collection of **Learning Objects**. Each Learning Object represents a single concept and is linked to a specific Learning Objective. A Learning Objective is a statement that describes content that will be covered in a module. Learning objects, aimed towards specific learning objectives, enable learners to achieve the module's stated learning outcomes. The structure of a Learning Object as defined by Cisco Systems (Figure 3.15) is ideally suited to the provision of differentiated instruction (Section 5.4.3). From the Learning Object level, the instantiation in Moodle becomes flexible. Moodle provides a collection of learning resources and activities that can (Section 3.4) represent a single Learning Object. The same learning resources and activities can also represent information objects or media elements that can be combined to form Learning Objects. For example, a file (one of the Moodle resources) can be a Learning Object, an information object or a media element depending on the contextual use and content of the file. It is therefore necessary to tag the resource or activity with educational metadata to identify the objects pedagogic role and to make the content of the file known to the learner. Moodle Activities can be used as Learning Objects for formative or summative assessment to reinforce and measure learning.

Information Objects represent the building blocks of a Learning Object. An overview and a summary are two specialised information objects to bookend a Learning Object. An Overview Information Object should introduce the Learning Object, explain its relevance, present the learning objectives, mention any prerequisites and provide an outline of the rest of the Learning Object. A Summary Information Object should review the material, introduce the next step and provide links to additional material if desired. In addition to the overview and summary, five general types of information objects convey the content of the module:

- A concept in the form of a definition or example
- A fact providing relevant background information about the concept
- A procedure describing sequential steps to perform a task
- A process describing the flow of events in a system
- A principle describing best practices and guidelines

Media Elements represent the lowest atomic component of an Information Object. In Moodle a media element can be represented by a file, a page, a label or even a URL, depending on its content. The content of a media element can, for example, be audio, video, text or graphic.

When planning the required course content, knowledge can be represented in a domain ontology. An ontology formally and explicitly describes concepts in a domain of discourse. Several tools and methodologies exist for developing ontologies (Jones, Bench-Capon and Visser, 1998). The process model proposed in this thesis is non-prescriptive regarding the ontology formalism, since at this stage the processes are not automated. To build more automated intelligence into the learning design and learner modelling, as is the case in adaptive education systems, more formal knowledge representation is needed. For differentiated instruction, the ontology could simply be a representation of the domain vocabulary, a taxonomy of essential concepts and their relations in the form of hierarchies and constraints (Gašević, Djurić, and Devedžić, 2006).

Regardless of the way knowledge is represented, the output from this step will feed into the process of developing or selecting/reusing suitable Learning Objects to cover each Learning Objective.

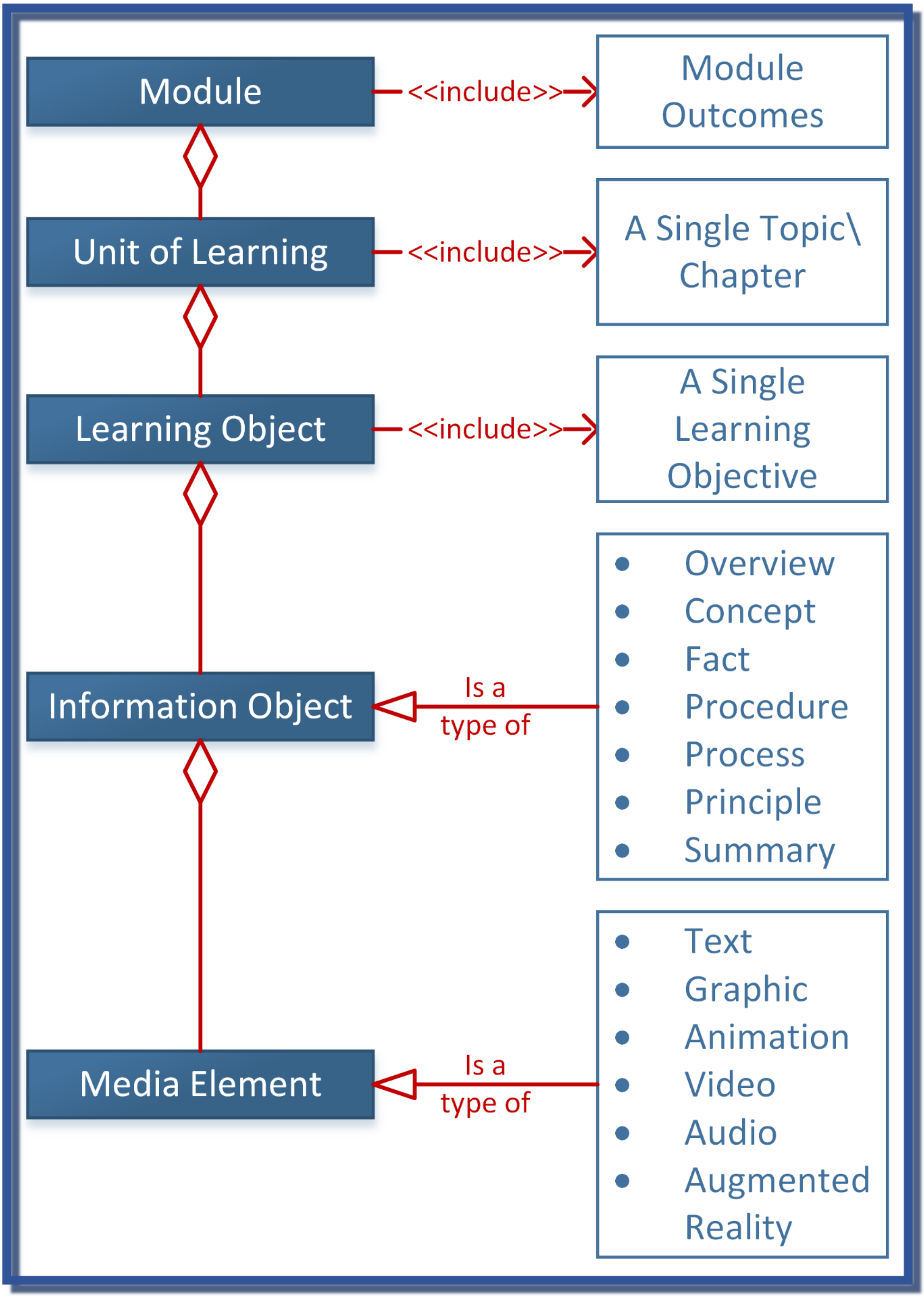


Figure 5.7 Module Organisation (Own Construction)

5.4.2.2 Create and Tag Learning Objects

Best Practice Guidelines developed by Cisco Systems (Section 3.3.1.2) recommend Learning Objects should be constructed by aggregating no less than five and not more than nine Information Objects. All content delivered through the Information Objects are grouped together to guide learners toward a single, stated Learning Objective.

A standard installation of Moodle provides at least seven different types of resources (Table 5.9). Theoretically, any resource added to a Moodle course can qualify as a Learning Object. However, each resource has implications for the proposed model in terms of granularity of the Learning Object and the learner interaction data it records. Data considerations are addressed in Section 5.5.

Table 5.9 Moodle Resource Recommendations

Resources	Recommended Usage (in terms of granularity)
Book	Learning Object
File	Information Object or Media Element
Folder	N/A (UI element to group objects together)
IMS Content Package	Learning Object
Labels	Media Element or Information Object
Page	Information Object or Media Element
URL	Information Object or Media Element

Moodle Activities are used to reinforce learning. The standard activities can be categorised as formative or summative assessment (Table 5.10), as tools for collaboration (Table 5.11) or communication (Table 5.12). While grades can be attached to collaboration and communication activities, their primary focus is not assessment. Since Moodle is open-source, a community of developers continually makes new activities available. The proposed model recommends that any newly submitted activity should be categorised similarly as presented here, i.e. the pedagogical intention of the learning activity must be clearly defined for differentiation.

Table 5.10 Assessment Activities in Moodle

Activities	Recommended Usage
Quiz	Formative/Summative Assessment (e.g. multiple-choice, matching, short-answer questions)
Assignments	Formative/Summative Assessment (e.g. digital submission of files)

Table 5.11 Collaboration Activities in Moodle

Activities	Recommended Usage
Database	Peer Collaboration (e.g. collaboratively adding to course content)
Glossary	Peer Collaboration (e.g. collaboratively adding terminology)
Wiki	Peer Collaboration (e.g. collaboratively adding to course content)
Workshop	Peer Collaboration (e.g. peer assessment)
Choice	Teacher-Student Collaboration (e.g. quick opinion polls)
Feedback	Teacher-Student Collaboration (e.g. custom survey instrument)
Survey	Teacher-Student Collaboration (e.g. verified survey instrument)

Table 5.12 Communication Activities in Moodle

Activities	Recommended Usage
Forum	Asynchronous Communication (e.g. ongoing discussion between students and teachers)
Chat	Real-Time Communication (e.g. chat between student-student or student-teacher)

The Lesson, LTI External and SCORM activities (Table 5.13) are typically used for self-contained courses/Learning Objects. LTI External and SCORM would usually be a fully-fledged, self-contained course with resources and activities pre-packaged. Self-contained courses can be developed using freeware or commercial third-party authoring tools and uploaded to Moodle via the SCORM activity[4]. A detailed analysis of available authoring tools is beyond the scope of this thesis and remains future work. For the proposed model to be effectively instantiated in a Learning Management System, such an analysis of third-party authoring tools should determine the extent to which Learning Objects can be differentiated and whether the learner interaction data is readily available for analysis.

Table 5.13 Moodle Activities as Self-contained Courses

Activities	Recommended Usage
Lesson	Self-contained Course (e.g. delivering resources and activities with rudimentary branching)
(LTI) External	Self-contained Course (e.g. course previously published elsewhere)
SCORM	Self-contained Course (e.g. interactive course created in 3rd party authoring tool)

[4] https://docs.moodle.org/35/en/Creating_SCORM_Content#Creating_SCORM_Content

Once the Learning Objects are created and added to Moodle, they must be tagged with educational metadata. This metadata that describes the learning resources and activities, is necessary for the learner modelling and differentiation learning design phases. The description of the learning resource is one part of the metrics analysed during learner modelling to infer learner behaviours (Section 5.4.1). The learners must also be informed of the type of resource to enable them to select the most appropriate one for their needs. It should be noted that metadata in this instance is not necessarily machine-readable. The focus in differentiated instruction is on making information available for human (teacher and learner) decision-making. In Moodle, metadata for human consumption can be added in at least three different ways:

- Through establishing a convention for naming the resource/activity, where pre- or post-fixes can be used to convey relevant information quickly
- Through completing the “Description” field, available when adding a new resource/activity
- Through completing the “Tag” field, available when adding a new resource/activity

When the Tag field is used, learners can quickly filter resources and activities based on the tag provided by the lecturer. The lecturer must activate the Tag block and change the context to the module. This will display a word cloud showing all tags used to describe the content in a course. Learners that are more interested in examples of concepts rather than definitions, for example, can click on the “Concept_Example” tag to see a list of all resources and activities tagged with this keyword. The Description field can be used to give more details about the content of the resource or activity. Descriptions, if provided, can be displayed on the page before clicking the link, or in the window that opens after the link to the resource/activity is clicked.

5.4.3 Differentiated Learning Design

The Differentiated Learning Design phase focuses on designing the learning rules in the Goal and Constraint Layer (before learner modelling, Section 5.4.3.1) and implementing the learning rules in the Course Layer (after learner modelling, Section 5.4.3.2).

5.4.3.1 Learning Rules

Rules for differentiated learning design based on learner attributes can be represented using IF statements of the format in Equation 5-6:

Equation 5-6 Differentiated Learning Rules for Proposed Model

$\boldsymbol{IF\ Attribute\ THEN\ Action: Object\ \{Value\}}$, **where**

- Action = Sort | Dim | Hide | Highlight | Trigger | Show
- Object = MediaType | InformationType | ActivityType | UI element
- Value = Value of metadata tag for Object

Using the template proposed in Equation 5-6, the suggestions in Table 5.14 and Table 5.15 can be combined with the actions (Section 5.3.3.2) to plan differentiation.

Table 5.14 Potential Dichotomous Attributes for Learning Rules

Cognitive Domain		Affective Domain	
Modality	**Mental Strategies**	**Engagement**	**Social**
• Visual • Verbal • RW • Tactile	• Abstract versus Concrete • Deductive versus Inductive • Reflective versus Active • Serial versus Holistic • Linear versus Alternating • FD versus FI • Synthesis versus Analysis	• Deep versus Surface • Strategic versus Resistant • Meticulous versus Careless • Persistent versus Irresolute	• Individual versus Group • Introvert versus Extrovert • Competitive versus Collaborative

Table 5.14 shows dichotomous attributes derived from learning style theories (see progression from Table 4.5, Table 4.6, Table 4.7, Table 5.5, Table 5.6, Table 5.7 and Table 5.8). When setting up the condition in the learning rule (i.e. the attribute), a

learner will fall either in one or other pole of each dichotomy, e.g. abstract or concrete, strategic or resistant, etc.

Table 5.15 Potential Object-Value Pairs for Learning Rules

Object	Value
MediaType	• Text • Graphic • Animation • Video • Audio • AR/VR
InformationType	• Overview • Overview_Scenario • Concept_Definition • Concept_Example • Fact • Procedure • Process • Principle • Summary • Summary_Additional
ActivityType	• PeerCollaboration • TeacherCollaboration • FormAssessment • SummAssessment • Forum • Chat
UIElement	• Index • NavButtons • Annotation

Table 5.15 shows a list of potential metadata tags for objects (regarding media types, learning resource and activity types or UI element descriptors) and their associated values. The consequent in the learning rule, the action statement (e.g. sort, dim, hide, highlight, trigger or show), is followed by an object-value pair that identifies the media type, or information type or activity type or UI element on which the action should be performed.

- Media types represent the main multimedia elements and are associated with the attributes related to modality.
- Information types represent the values associated with expositive information objects and are linked to mental strategies.
- Activity types represent the values associated with active information objects and are linked to engagement and social attributes.
- UI elements are special affordances required for learners with certain attributes, e.g. Field-dependent learners and their need for sequential navigation buttons or Field-independent learners flourishing with a hypermedia index.

Below are a few sample scenarios to illustrate setting up learning rules.

Scenario 1: Table 5.5 describes a visual learner as someone with a preference for extracting information from Learning Objects in the form of still images, animations or video. Should the teacher decide to recommend all graphics, videos or animations the learning rule can be:

- IF Visual THEN
  - Highlight: MediaType {Graphic | Animation | Video}

    (where "|" is used to show "or")

If the decision is made to first present all graphics, animation and videos before any textual descriptions or audio narration (e.g. podcasts), and grey out text and audio, the rule will be:

- IF Visual THEN
  - Sort: MediaType {Graphic | Animation | Video; Text | Audio}
  - Dim: MediaType {Text | Audio}

    where the semi-colon (;) is used to set the order of the objects inside { }.

Scenario 2: Table 5.6 describes an "Abstract" learner as someone who has a higher preference for theoretical concepts, while a "Concrete" learner shows a predilection towards active experimentation. Suppose the teacher wants to set up rules for accommodating Abstract and Concrete learners, a rule could be:

- IF Abstract THEN
  - Sort: InformationType {Concept_Definition; Concept_Example}
  - Highlight: InformationType {Concept_Definition}
- IF Concrete THEN
  - Show: InformationType {Overview_Scenario} //Optional object
  - Sort: InformationType {Concept_Example; Concept_Definition}
  - Highlight: InformationType {Concept_Example}

Scenario 3: Table 5.7 shows a "Resistant" learner is one who shows little interest in the course material. Suppose the lecturer needs to perform an intervention of some kind, the rule could be:

- IF Resistant THEN
  - Trigger: Intervention //Where the form of the intervention would be defined on a case-by-case basis

The Section that follows explores how differentiation can be practically achieved in Moodle.

5.4.3.2 Tailoring Course Material in Moodle

To achieve differentiated instruction learners are grouped together based on their attributes and tailored course material presented to one group will be slightly altered from the material presented to another (Section 3.3.2.1). The Moodle Groups, Groupings and Cohort settings can be utilised to enable differentiated instruction.

The Groups[5] mode allows a teacher to cluster learners together so that content can be filtered based on their relevance to the group members. The default setting is No Groups, but if groups are required, two different modes are available:

- Separate Groups allow members to only see content relevant to them

[5] https://docs.moodle.org/35/en/Groups

- Visible Groups allow members to only interact with group members, but they can view the content of other groups

Groups can be set on a course level, that filters down to all activities or is set on activity-by-activity level. The Groupings[6] setting allows a teacher to create a collection of groups. So, while a group is a collection of learners, a grouping is a collection of groups. Members are added manually to groups through list boxes. In large classes, this is not practical, so a new plugin or work-around is needed. In Moodle version 3.5, Cohorts[7] are groups available site-wide and enable quicker enrolment through a spreadsheet. Since cohorts are available to all courses in a site, learners only need to be added once to a specific cohort grouping (e.g. all learners exhibiting behaviours of “abstract” learners in one cohort and all learners exhibiting behaviours of “concrete” learners in another). If “Cohort sync” is enabled site-wide, teachers in different courses who wish to tailor material for different cohorts can easily take advantage of these different learner categories.

If learners have not yet been uploaded to the system, the CSV file to bulk upload users must have the structure:

```
username, password, firstname, lastname, email, cohort1
```

If learners are already on the system, but not yet in a cohort, only the username and cohort1 fields are necessary.

If a teacher wants to use the learner modelling phase (Section 5.5) to categorise learners as either concrete or abstract, a spreadsheet must be created with the attribute (“Concrete” or “Abstract”) recorded in the last column under the “cohort1” heading. This spreadsheet must be furnished to the site administrator for updating cohort lists. With learners assigned to different categories they can be given restricted access to different topics, resources or activities created during the initial learning design phase (Section 5.4.2).

To summarise, the sequence of steps in Moodle to enable the grouping of learners based on attributes is as follows:

[6]https://docs.moodle.org/35/en/Groupings
[7]https://docs.moodle.org/35/en/Cohorts

(1) Cohorts, linked to relevant learner attributes, added by site administrator and Cohort sync enabled
(2) If data is available, learners assigned in bulk to cohorts through CSV file
(3) Teacher creates Groups in their course that are related to learner attributes
(4) Teacher adds Cohorts to relevant groups
(5) If necessary, Groups further added to Groupings
(6) For each topic, resource or activity configure the Restrict access setting to only those applicable groups

Through the Cohort, Group and Grouping settings, it is technically possible to split and filter content to different sets of learners in Moodle. Whether the provision of differentiated instruction had a positive impact on learning is beyond the scope of this thesis.

5.4.4 Critical Reflection on Sub-Objective 2

Sub-Objective 1 establishes the global design of the model for differentiated instruction based on a dynamic learner profile (Section 5.3). Sub-Objective 2 (Section 5.4) focuses on enabling differentiated learning design in a Learning Management System, specifically instantiated and verified in Moodle. In terms of the methodology developed in Section 2.3.3, Sub-Objective 2, therefore, deals with the refinement of a sub-component of the proposed solution. The evaluation cycle of this iteration primarily reflects on the expected practicality and effectiveness of the sub-component (Section 2.3.3.2). Consequently, this Section reflects on whether Moodle shows potential for enabling differentiated learning design (expected utility) and how usable the tools are to tailor course material (expected practicality).

Section 5.3.2.2 proposes a generic KPI for evaluating learning design as “Learning design is optimised for a learner with Attribute X”. The underpinning assumption, that learning design suitable for one group of learners may frustrate another, means that it is necessary to identify dichotomous attributes towards which learning design can be differentiated. In Table 5.14, several potential attributes have been suggested based on their hypothesised cognitive and affective implications.

In the cognitive domain, the focus is on attributes associated with:

- Perceptual modality, i.e. the format in which the learner extracts information the best
- Mental strategies employed by learners to make sense of new information

In the affective domain, the focus is on attitudes, feelings and emotions associated with learning. In particular, the affective domain classifies attributes associated with engagement levels and preferred social interactions that affect learners' motivation.

The attributes identified in this thesis do not make an exhaustive list of all characteristics with an influence on learning. The primary objective of the proposed model is to enable discovery of cognitive and affective attributes with an impact on learning. The initial list of proposed attributes represents a starting point for future investigation into attributes that can be inferred by analysing learners' online behaviours. Once learner attributes have been identified, impact studies can commence to determine the impact, if any, that the differentiated learning design has on learning effectiveness and efficiency and learner satisfaction. While the nature of the impact studies is beyond the scope of this thesis, the model proposed to identify relevant attributes from learner behaviours is a vital first step. Section 5.4.1 illustrates a technique to use the definition of suitable attributes from learning style theories to describe expected online learning behaviours that will likely be exhibited by a learner who possesses the stated attribute. For differentiated learning design to be practical and effective in a Learning Management System, the behaviour must be described using two components relevant to the online learning environment (as illustrated in Table 5.5, Table 5.6, Table 5.7 and Table 5.8):

(1) The learning resource/activity accessed by the learner and/or the affordances built into the UI of the system
(2) The temporal, navigation and performance metrics recorded by the Learning Management System

The first component has implications for the initial and differentiated learning design choices. Identifying the types of learning resources, activities or UI elements preferred by learners dictates the type of Learning Objects and UI affordances that should be included in the course content.

To evaluate the practicality of the Learning Management System and the effectiveness of the proposed techniques to enable differentiated instruction, the following must be determined in terms of learning design:

- Does the Learning Management System have the necessary support in terms of the types of learning resources and activities to provide the required Learning Objects and UI affordances?
- Does the Learning Management System provide the ability to describe the Learning Object with educational metadata?
- Does the Learning Management System allow the formation of different groups of learners?
- At what level of granularity can Learning Objects best be differentiated? Can the Learning Object be realised in the Learning Management System at an appropriate granularity?
- What actions can be performed to tailor Learning Objects for different learners? Can these actions be performed in the Learning Management System or authoring tool used to construct Learning Objects?

Section 5.4.2.1 proposes the following levels to organise learning content in a Learning Management System, to enable differentiated instruction while achieving the same module outcomes (illustrated conceptually in Figure 5.8):

- At the highest level, define a Module and associated learning outcomes
- Modules consist of several units of learning, each one linked to a single topic
- Units of learning are constructed by aggregating several Learning Objects, each one linked to a specified Learning Objective
- A Learning Object comprises a collection of information objects that provide an overview and summary, and convey content in the form of a concept, fact, procedure, process or principle
- The content in an information object can be presented through any of the six types of multimedia elements (text, graphic, animation, video, audio or augmented reality).

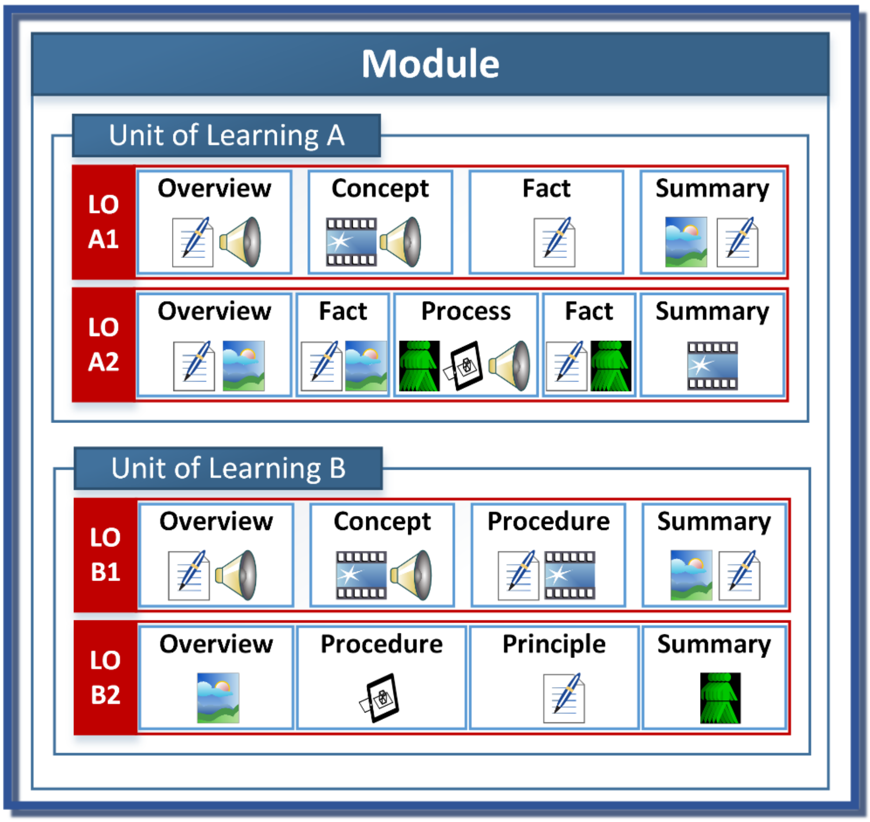


Figure 5.8 Course Organisation Example (Own Construction)

It is proposed in this thesis that differentiation is best achieved if information objects are tailored based on attributes associated with mental strategies, and media elements are tailored based on attributes associated with perceptual modality. Attributes associated with the social dimension dictates how a learner will respond to activities involving collaboration, and attributes associated with engagement should trigger appropriate interventions by the teacher if required.

Moodle has a rich toolset of customisable learning resources and activities, provided in a default installation and developed by the open source community. This Section (5.4) illustrated and confirmed that Moodle has the potential to enable differentiated learning design through the Cohort, Groups and Groupings functionality.

The next step is to further refine the requirements for the learner modelling phase in a Learning Management System. Like the learning design phase, the investigation will instantiate the learner modelling phase in Moodle in order to evaluate expected practicality and effectiveness.

5.5 Sub-Objective 3: Learner Modelling in Moodle (Iteration 4)

The learner modelling phase uses several analysis techniques, including educational data mining to build a learner profile. This thesis adopts an implicit method for modelling that builds the profile by analysing how learners navigate through the online course material. This Section reports on a pilot study involving an instantiation of the learner modelling phase of the proposed model in the context of a Learning Management System (Appendix B, Appendix C). Moodle is selected for this study based on its widespread adoption, but the principles derived from this Section should equally apply to most Learning Management Systems with minor adjustments.

The requirements imposed by the proposed model on data collection from a Learning Management System are explored in Section 5.5.1. Analysis techniques for finding relevant patterns in this data is proposed and briefly described in Section 5.5.2. Section 5.5.3 illustrates alternative scenarios after initial analysis and the form of the learner profile proposed for this thesis. Section 5.5 critically reflects on the expected practicality and effectiveness of maintaining a learner profile from Moodle data.

5.5.1 Data Collection

Section 5.4.3.2 proposes a technique for filtering content based on different groups of learners in order to satisfy the goals of differentiated instruction. The aim of the data collection and analysis phase is to discover meaningful patterns in the learning activity logged by the system. These patterns are derived from learning style theories that hypothesise behavioural differences in learners (Section 5.4.1). Section 5.5.1.1 describes the requirements for data extracted from Learning Management System log files, and Sections 5.5.1.2 explores whether the data provided by Moodle databases satisfies these requirements.

5.5.1.1 Information required for differentiation

Section 5.4.1 categorises several online learning behaviours as described in a selection of learning style theories. One grouping of behaviours is related to cognition, whereby learners reveal preferences regarding the media elements through which they best extract information and the mental strategies used to process new information. Another grouping of behaviours reveals affective states, specifically how engaged learners are with the course material and their preference regarding interaction with others in the learning process.

Educational Metadata: The pedagogic intent of Learning Objects is identified through metadata tags assigned to resources and activities uploaded to the Learning Management System. Per illustration, the following are examples of metadata tags that could be assigned to resources and activities:

- **MediaType:** -Text, -Graphic, -Animation, -Video, -Audio, -AR/VR
- **InformationType:** -Overview, -Overview_Scenario, -Concept_Definition, -Concept_Example, -Fact, -Procedure, -Process, -Principle, -Summary, -Summary_Additional
- **ActivityType:** -PeerCollaboration, -TeacherCollaboration, -FormAssessment, -SummAssessment, -Forum, -Chat
- **UIElement:** -Index, -NavButtons, -Annotation

These **object**-**value** pairs must be identifiable in the log file produced by the Learning Management System (Section 5.4.1).

Logged Metrics: Online behaviour is revealed by examining metrics about time, sequence of activities and performance. In particular, based on Table 5.5, Table 5.6, Table 5.7 and Table 5.8, the following metrics are required:

- Number of hits on:
  - A specific type of resource or activity
  - All resources and activities
- Duration spent on:
  - A specific type of resource or activity
  - All resources and activities
- The type of resource/activity consistently accessed before another type
- Grades in assessments in relation to the identified behaviours
- Number of errors made in assessments
- Number of omissions in submitted assessments, e.g. abandoned quiz questions
- Number of revisions made to assessments before submitting
- Content of forums and chat activities
- Number of messages in forums and chat activities
- Participation in activities involving Peer Collaboration
- Participation in activities involving Teacher Collaboration

The educational metadata and metrics listed above must be provided by the log files produced by the Learning Management System.

5.5.1.2 Information Logged by Moodle

Moodle produces several reports describing learner activity (Section 3.4.2).

Table 5.16 Fields from a Moodle Log File

Field Name	Description
Time	Timestamp showing date (mm/dd/yy) and time (mm:ss) of initialising the event
User full name	Student number, name and surname of the user that performed the action
Affected user	Student number, name and surname of the user on who the action is performed, where applicable
Event context	The type of resource and activity (Section 3.4.1), followed by the name given to the object by the lecturer
Component	Either System for course level events or the name of the type of resource/activity
Event name	Short description of the action performed
Description	Description of the event containing the user id, event id, and course id
Origin	• Web Interface (usually resulting from student/teacher), or • Command Line Interface (usually resulting from site admin activities)
IP address	IP address of the device from where course material is accessed

A default installation includes the following under Module Administration > Reports:

- Competency breakdown – a report that describes the level of proficiency of a learner in relation to the module outcomes

- Dates – a tool with which due dates can be set for all activities
- Logs – a log of activities filtered by student, day, activity or action
- Live logs – real time updates of all student activities
- Activity report – the number of views on each activity and resource
- Course participation – a list of students, the actions they performed and the number of times they performed the action on a specific activity
- Activity completion – a report showing the progress of the learners through the course by listing the completion status of resources and activities where activity completion status is enabled
- Statistics – a report of the amount of activity on the course

Of these reports, the logs, activity completion and grader report warrant a closer inspection to extract or calculate the metrics required for differentiation (Section 5.5.1.1). All logs can be downloaded in Microsoft (MS) Excel format. The fields that can be viewed from a standard Moodle Log are shown in Table 5.16.

The fields that can be viewed from an Activity Completion Report are shown in Table 5.17.

Table 5.17 Fields from a Moodle Activity Completion Report

Field Name	Description
First Name / Surname	Student ID and Full Name of student
ID number	Unique student number
Email address	Student email address
Resources and Activities	Columnar list of all resources and activities added to the course, showing for each student: • Completion Status – Completed or Not completed • Date Completed

The fields that can be viewed from a Grader Report are shown in Table 5.18.

Table 5.18 Fields from a Moodle Grader Report

Field Name	Description
Surname	Student surname
First name	Student first name
ID number	Unique Student number
Department	Faculty/Department where student is registered
Email address	Student email address
Activities	Columnar list of all activities added to the course and the grade students achieved in each activity
Module total	Average mark obtained in all activities

Based on the fields available, only the grades in assessments are usable in their current form for analysis. Most of the metrics need to be transformed from the available logs. The required transformations are described in Section 5.5.2.1. It is also clear that the educational metadata is not readily available in existing reports. This points to a need for pre-processing to transform the data into an appropriate format and for metadata to be included in the log file.

The Moodle Quiz activity provides several reports that can be used to improve quiz questions and to discern differences in learner behaviour. The details from a Quiz report are more complete than the actions stored in a general Moodle log. The general log reports on the date and time that the quiz was viewed, started, submitted or abandoned, and reviewed. The Quiz Grade report shows each attempt made, the start and end time, the time taken, the grade for each question and the total quiz grade. Through this information, one can extract:

- Quiz grades
- Number of quizzes started but not submitted
- Number of errors made in assessments

- Number of omissions in submitted assessments
- Number of times the quiz is re-attempted

The number of revisions made to a quiz before it is submitted is currently not being tracked in Moodle. To view the actual answers given by the learners, the Quiz Responses report must be consulted. Another useful Quiz report provided in Moodle shows complete statistics. The statistics provided in this report can be used to determine whether the test questions are within acceptability standards based on classical test theory[8].

The standard Moodle log reports the number of times the forum or chat activities were viewed, and the number of messages sent. Should the need arise to analyse the content, this can be extracted from the actual objects.

5.5.2 Data Analysis

The aim of the data analysis is to infer the learner attributes identified during the preliminary goal setting phase (Section 5.3.3.3). These attributes are inferred by analysing the behaviour of learners logged by the Learning Management System. These logs record several events from hundreds of learners, producing large datasets. Consequently, educational data mining techniques are necessary to distil usable information from the raw data (Section 4.2.1). Section 5.5.1 revealed that the data logged by Moodle require pre-processing (Section 5.5.2.1) to get the correct mix of metrics into an appropriate format to be usable as input for the educational data mining algorithms. For the objectives of this study, learners must be clustered into dichotomous groups exhibiting similar online behaviour; hence, the use of clustering techniques to find these patterns (Section 5.5.2.2 and Appendix B).

5.5.2.1 Data Pre-processing

The analysis phase of this study focuses on modelling techniques to analyse data from log files downloaded from a standard Moodle installation. These logs are produced by integrating data from various Moodle databases. For the initial analysis, no further integration is done during pre-processing.

[8] https://docs.moodle.org/dev/Quiz_report_statistics

When comparing the metrics required by the proposed model (Section 5.5.1.1) and the data provided by Moodle (Section 5.5.1.2) it is evident that data reduction, transformation and cleaning is needed. The following techniques need to be applied to Moodle logs:

- Sampling – removal of irrelevant records
- Feature subset reduction – removal of irrelevant fields
- Aggregation – transforming data into a summarised format
- Discretisation – replacing numeric values with interval/conceptual labels
- Smoothing – removing outliers

Sampling: The Moodle log files include actions performed by students, lecturers, tutors and the site administrator. All actions performed by users with a non-student role must first be removed from the log. While the roles are not provided in the standard Moodle log, this information is easily obtained from the Moodle course. Depending on the goal of the analysis, certain records can further be removed. For example, if the preliminary goal setting phase decided to tailor instruction along the abstract-concrete dichotomy, only actions on resources and activities tagged as being suitable for abstract or concrete learners need to be kept in. Or in another scenario, the teacher may wish to discover groups based on their preference for one of the multimedia elements and therefore would only need actions on resources tagged with one of the MediaType keywords. To enable filtering based on metadata tags, a new field is needed in the standard Moodle log that shows the keywords used to describe the resource or activity. A workaround could be to include the keyword in the name of the Learning Object and extract it through a Microsoft (MS) Excel text function. However, the drawback of this approach is that multiple keywords are used to describe the Learning Object.

Feature subset reduction: The Moodle reports all have fields that are not necessary as input into the educational data mining algorithm or the resulting action after analysis.

- The standard Moodle log can be reduced to:

```
Log(Time, FullName, EventContext, Component, Description)
```

  - The Time field should be separated into two columns: Date and Time
  - The student number can be extracted from FullName

  - The name of the Learning Object can be extracted from EventContext
  - The type of resource/activity can be extracted from Component
  - System generated UserID can be extracted from Description, and used as a privacy measure instead of the student number and name
- The Activity Completion report can be reduced to:

```
ActivityCompletion(ID, CompletionStatus, CompletionDate)
```

- The Grader report can be reduced to:

```
GraderReport(ID, Grade)
```

Feature construction and aggregation: Temporal metrics are not readily available in the original dataset and need to be constructed and aggregated from the timestamp and metadata tag of the Learning Objects. The metadata tag is currently not displayed in a standard Moodle log. Since the educational metadata plays such a large role in the implicit learner modelling process, this highlights the need for modification of the existing Moodle source code.

The time field displays the date and time of access to a particular resource or activity. This timestamp can be used as a basis to calculate session duration and to determine the sequence in which Learning Objects were accessed. For certain resources and activities, the time spent on each Learning Object can also be estimated from the timestamp. The quiz activity records the exact duration in the Quiz Grade report. The File resource, if it is downloadable, is an example of a Learning Object for which it is impossible to gauge an exact duration. Furthermore, certain assumptions need to be made when calculating duration spent on a resource. One such assumption is that the learner was actually perusing the Learning Object before clicking on the next one. The following illustrates, at a high level of abstraction, how to construct and aggregate new metrics needed for data mining:

- Number of hits on a specific type of resource or activity
  - Specify the MediaType/InformationType/ActivityType required
  - Define “access” for each resource/activity type, e.g. access can be defined as “Viewed” only or as a combination of “Viewed” and “Updated”
  - For each learner, count the number of times the selected objects were accessed

- Number of resources and activities used
    - Add up the number of hits on all Learning Objects of any type
- Duration spent on a specific type of learning resource or activity
    - Specify the MediaType/InformationType/ActivityType required
    - For each learner, calculate the duration spent on the selected objects that were accessed (Caveat: Duration does not work for all types of learning resources. For example, if the learner downloads a file it may appear in the log as a short stay. However, more time will be spent offline studying from the resource.)
- Duration spent on all resources and activities
    - Add up the total time spent in each online session (Caveat: Additional data may be needed to verify that learners spent the time studying and not on offline tasks while being logged in to Moodle.)

Discretisation: Numerical data like grades is often more easily interpreted and understood when represented as a range. Typical labels used to represent a learner's grade are:

- FAIL: If Grade <50%
- PASS: If Grade >= 50% AND Grade < 75%
- DISTINCTION: If Grade >=75%

Another example where discretisation can simplify interpretation is if temporal metrics are converted to labels instead of recording hours, minutes and seconds. Typical labels used to represent both the number of hits or the duration spent on Learning Objects are (amounts used as illustration):

- LOW: Less than five minutes total spent on resource
- MEDIUM: Between five and ten minutes total spent on resource
- HIGH: More than ten minutes total spent on resource

To define these ranges, the actual values for duration and number of hits on a specific Learning Object should be estimated for each resource and activity. The metric to keep track of the order in which Learning Objects of a given type are accessed, can also be discretised based on the time of access. Since the attributes that need to be inferred are dichotomous, it is often the case that their behaviour will also be a choice between one or another type of Learning Object. For example, the values for each record could then be:

- 0s: where the time of first access is frequently before the alternative
- 1s: where the time of first access is frequently after the alternative

Table 5.19 Discretisation of Sequence

UserID	Concept_Definition	Concept_Example
16179	0	1
48083	1	0
57411	0	1

In Table 5.19, the learners with UserID 16179 and 57411 frequently accessed definitions before examples, while the learner with UserID 48083 frequently accessed examples before definitions. The process to determine preferred sequence can be achieved through the following steps:

- Firstly, in the initial learning design an information object must be tagged as a Concept_Definition and another as a Concept_Example
- For each concept, the time of first access to the definition and related example must be determined and values of 0 and 1 assigned as appropriate
- For each learner, determine the total frequency of 0's and 1's over all concepts and aggregate to a single 0 and 1 (where 1 represents the max access frequency).

Smoothing: When analysing Moodle logs, certain events represent a temporary stop on the way to their target destination. For example, if a learner closes a popup window, the return to the main course page will be recorded as an event that the module is accessed. Seen in isolation to other events it may, therefore, seem like multiple accesses to the module in rapid succession. Within a single session, these events represent outliers and must be removed from the log. Another example outlier would be a learner opening a resource (perhaps in error or simply to view the description) and immediately closing it. These entries should also be removed prior to running the data mining algorithm.

The pre-processing of the Moodle reports will involve creating new spreadsheets with all required metrics needed as input for the data mining algorithm. If a tool such as WEKA (Frank, Hall and Witten, 2016) is used, these spreadsheets must be converted to text files with ARFF (Attribute-Relation File Format) format. WEKA can be used to convert spreadsheet files with the .CSV extension into the ARFF format.

5.5.2.2 Clustering

Clustering is a technique applied to structure discovery problems (Section 4.2.1.2). Patterns in datapoints emerge when clustering algorithms such as K-Means are invoked on a large dataset. For this study, clustering is used to group together learners exhibiting similar behavioural patterns. This is done to identify dichotomous attributes upon which differentiated learning design choices can be based. Once learners are grouped based on their shared attributes, Learning Objects that suit their preferences can be presented to them.

Recall that the preliminary goal setting phase of the proposed model (Figure 5.3) requires the selection of a clear goal. It is proposed that the goal should be specific down to a single set of dichotomous attributes. These attributes have hypothesised online learning behaviours and a first step in the testing of these hypotheses is to partition learners based on their behaviours. In the process modelled in this study, the behaviours are described using metrics extracted from Moodle log files and stored as feature vectors in an ARFF file. These feature vectors are used as input into the chosen clustering algorithm.

In some clustering algorithms, the number of clusters can be specified. Since we are trying to find two groups labelled at opposite poles of a dichotomy, it is proposed that the “number of clusters” parameter is set at two. Alternatively, three clusters can also be tested. Using three clusters, learners can be partitioned into groups exhibiting strong, weak or balanced preferences towards a specific pole of a dichotomy. For example, with three clusters a learner may show strong or weak tendencies towards favouring “Abstract” Learning Objects, or they may exhibit a balance between “Abstract” and “Concrete” attributes. The resulting centroids reported in the output of the clustering algorithm can act as threshold values to classify and identify learners according to the relevant attributes.

Clustering algorithms provided in WEKA are:

- Canopy
- Cobweb
- EM
- FarthestFirst
- FilteredClusterer
- HierarchicalClusterer
- MakeDensityBasedClusterer
- SimpleKMeans

While it is beyond the scope of this thesis to do a detailed comparison of the different algorithms, the wide choice emphasises the need for the "Optional Alternatives" step in the proposed model (Figure 5.3).

5.5.3 Optional Alternatives and Implement Insights

The model proposed in this thesis (Figure 5.3) recommends an "Optional Alternatives" step before invoking the differentiated instruction step. The following are illustrative of the types of choices that could potentially impact the outcome of the data analysis phase that may warrant further investigation:

- In terms of the Clustering using the WEKA workbench, a different algorithm may be selected
- The number of clusters can be changed from two to three
- The selected feature vectors can be expanded by adding more metrics from the Moodle reports
- Association rule mining can be used on the output of the clustering in order to discover the correlation between the final marks and learner behaviour

The "Implement Insights" step of the proposed model culminates in invoking the differentiated learning design. The learner profile must be updated as an intermediate step post analysis. In Moodle, this means that a .csv file must be created with the ID of all learners enrolled in the Module and their discovered attributes. This .csv file can be used to create Cohorts (Section 5.4.3.2). Teachers can add these cohorts to the Groups and Groupings they created for their modules.

5.5.4 Critical Reflection on Sub-objective 3

Sub-Objective 1 establishes the global design of the model for differentiated instruction based on a dynamic learner profile (Section 5.3) and Sub-objective 2 deals with the instantiation of the learning design in a Learning Management System. Sub-objective 3 (Section 5.5) focuses on the expected instantiation of the learner modelling phase, specifically building a learner profile from data provided by Learning Management System reports. In terms of the methodology developed in Section 2.3.3, Sub-objective 3, therefore, deals with the refinement of a sub-component of the proposed solution. As is the case with Sub-objective 2, the evaluation cycle of this iteration primarily reflects on the expected practicality and effectiveness of the sub-component (Section 2.3.3.2). Consequently, this Section reflects on whether Moodle shows potential for building a learner profile (i.e. expected utility) and how usable the available tools are for building a learner profile using the implicit modelling technique proposed in this study (i.e. expected practicality).

The success of the learner modelling phase relies heavily on the appropriate selection of learner attributes and the fact that the mapping between these attributes and their associated online behaviours are backed by a valid educational theory. Based on these hypothesised behaviours, teachers/instructional designers must implement an initial learning design satisfying all pedagogical needs as suggested by the chosen theory.

At the stage of the initial design (Section 5.4.2) the material is presented in a one-size-fits-all approach and it is up to the learners to select appropriate Learning Objects based on their preferences. To empower learners to make the best possible choice of Learning Objects, information about the resource or activity must be shared with the learners. In Moodle, the Learning Management System evaluated in this study, at least three ways have been identified to guide learners towards appropriate Learning Objects (Section 5.4.2.2). When a naming convention is used to describe the type of resource/activity, the learner must be oriented towards this naming convention. The name is recorded in a standard Moodle log file, so the educational metadata can be extracted from the relevant field. When the “Description” field is used to describe the resource/activity, more information can be made available to the learner. This information is visible to the learner either on the main page or on the landing page after opening the resource/activity, or in Moodle 3.5 the description appears in a tooltip when pointing to the object. If the description is too long, though, the tooltip is not as effective. The description is not displayed in the report, so this information is not readily available for learner modelling. If the "Tag" field is used to describe a new resource/activity, this description is also not in the report. More tags can be used to describe the resource/activity than through the naming convention. Learners can also use a Tag block in Moodle to quickly filter all relevant resources and activities of the type they prefer. However, the information stored in the tags are not available in the Moodle reports.

Since Moodle is open source, this information can be added to the reports through joining and querying relevant tables. Hence, while it is possible to guide learners to appropriate Learning Objects, some alterations are needed to make the process more user friendly and practical. One set of object-value pairs that is not currently logged is the usage of an index or navigation buttons if the Moodle Book resource is used. This information is a necessary feature when trying to discern, for example, Field Independent or Field Dependent learners. Further investigation is needed into click-level data that can be logged by Moodle to record this type of UI level interaction beyond access to resources and activities, as is currently the case.

Evaluation of the following reports showed that it is possible to extract the necessary temporal, navigation and performance metrics needed as input into educational data mining algorithms:

- Standard Moodle log report
- Grader report
- Activity Completion report
- Quiz (Grade report, Responses report, Statistics report)

However, the considerable effort required to get the raw data into an appropriate format makes any attempt at manual pre-processing unrealistic. Once the data is in an appropriate format (e.g. the ARFF file format used by WEKA), the actual process of applying the data mining algorithm is quick. The difficulty with the analysis is that teachers need to be competent in interpreting the output of the educational data mining technique for it to be of any value. If appropriate clusters have been identified, it is possible to add the identified attributes as a new label to a class list in .csv format that can be used to add learners automatically to Cohorts that Moodle can use to create the groups required for the differentiated learning design phase (Section 5.4.3.2).

5.6 Conclusion

The process model described in this thesis (Figure 5.3) acknowledges that learning analytics should not focus on technological issues alone but should also address questions of a pedagogical and ethical nature. Current learning analytics process models suffer from a myopic view and often ignore or under-report pedagogical validation and ethical oversight.

The model developed through this study proposes that educational theory should inform all learning design choices and guide all learning analytics intervention. Pedagogical alignment is achieved through an explicit preliminary goal setting phase that sets the goal and key performance indicators towards measuring goal achievement. For this study, the general goal is set as the differentiation of the learning design in a Learning Management System based on a dynamic learner profile. More specifically, the learning design choices are based on dichotomous learner attributes extracted from influential learning style theories.

This thesis does not take a stance on the validity of the learning style theories. Theories and associated attributes used as examples in this thesis are chosen for the following reasons:

- The theories have been identified by others as influential in the field of technology enhanced learning (Popescu et al., 2007b; Labib, Canós and Penadés, 2017)
- The theories describe dichotomous attributes that manifest through different online learning behaviours, making Clustering a suitable technique to discern these hypothesised differences
- Content, presentation or navigation of the learning material can be differentiated to accommodate these hypothesised differences in behaviour

The learner modelling phase exploits behavioural similarities exhibited by groups of learners sharing the same attributes and differences between groups of learners with attributes from an opposing pole along the same dichotomy.

Evaluation of the expected practicality and effectiveness of a Learning Management System that can be used to instantiate the proposed model reveals that little usable insight can be drawn directly from unprocessed Moodle reports. It is impossible to discern groups of learners exhibiting similar behaviours. Pre-processing is necessary to process raw data into a format suitable for educational data mining.

Pre-processing manually is an extremely time consuming and labour-intensive process, rendering the manual method impractical. Tools are available to facilitate the pre-processing of log file data, however, none of them meet the requirements imposed by the proposed model. For example, no tool was found to link attribute selection directly to educational theory. Ideally, the selection of a particular learner attribute should prompt the initial suggestion of relevant metrics based on hypothesised behaviours. It should also be possible for the teacher to add additional features as alternatives to the suggested ones.

The success of the learner modelling phase requires specific actions during the initial learning design. In particular, since the process of describing online behaviours refers to specific types of information objects and media elements, all resources and activities added to Moodle must be tagged with relevant educational metadata. While Moodle does provide the means to describe the Learning Objects in a way for learners to make an informed choice, this metadata is not recorded in current Moodle reports. These limitations highlight the need for Moodle to be extended with the necessary functionality to accommodate all requirements imposed by the proposed model. The open source nature of Moodle makes these extensions feasible.

Chapter 6. Discussion and Conclusions

Chapter 6
Discussion and Conclusion

6.1) Introduction

6.2) Summary of Research Process

- 6.2.1) Preliminary Phase
- 6.2.2) Prototyping Phase
- 6.2.3) Assessment Phase

6.3) Discussion and Critical Reflection

- 6.3.1) The Type and Level of Knowledge Contribution
- 6.3.2) Theoretical Contribution
- 6.3.3) Practical Contribution
- 6.3.4) Limitations of the Research

6.4) Recommendations

- 6.4.1) Recommendations for Policy and Practice
- 6.4.2) Recommendations for Refinement of Solution
- 6.4.3) Recommendations for Further Research

6.5) Epilogue

The aim of Chapter 6 is to highlight the main contributions of this thesis, examine the broader implications of the findings and recommend a way forward.

6.1 Introduction

This thesis reports on a study prompted by a desire to tailor instructional design presented in a Learning Management System according to learners' unique needs. Differentiated instruction is frequently accomplished through proprietary software based around a particular learning style theory. However, higher education institutions are increasingly abandoning proprietary learning environments and adopting Learning Management Systems to create an online learning environment. Learning Management Systems, though, are mostly suited to a one-size-fits-all approach to instructional design. The proposed solution is underpinned by a belief that the repeated online behavioural patterns exhibited by learners interacting with course material can be exploited successfully to discern cognitive strategies and affective states. In turn, this knowledge can inform differentiated instructional design choices that facilitate the cognitive and affective development of learners. This thesis, then, describes the iterative development and evaluation of a process model to enable differentiated instruction in a Learning Management System based on a dynamic learner profile built through learning analytics. Chapter 6 presents:

- Section 6.2: A narrative discussion of the model described in Chapter 5, related to the research question and objectives established in Chapter 1 and its iterative development and evaluation using the methodology described in Chapter 2.
- Section 6.3: Critical reflection on the contributions and limitations of the thesis.
    - Theoretical contribution: A summary of the main findings are presented in Chapter 3 and Chapter 4, resulting in an emerging model for differentiated instruction based on a dynamic learner profile (Section 5.3).
    - Practical contribution: Recommendations for instantiating the learning design and learner modelling phases in a Learning Management System in general and specifically an evaluation of the suitability of Moodle for supporting the process (Sections 5.4 and 5.5).
    - Limitations: A discussion of the methodological restrictions that constrained the current study.
- Section 6.4: Recommendations for policy and practice, refinement of the proposed solution and further research.

6.2 Summary of Research Process

In searching for a research design to guide this study, it was decided to synthesise a methodology that incorporates elements of Design Science Research (DSR), with a focus on information systems, and Design Based Research (DBR), with a focus on optimising educational processes. DSR and DBR are two closely related research methodologies with an iterative design focus. The synthesis of DSR and DBR resulted in Design Research in Technology Enhanced Learning (DeRTEL) (Section 2.3.3). DeRTEL is applicable to this study, since the general research area is Technology Enhanced Learning, and the proposed solution is a data driven approach to support differentiated instructional design choices in a Learning Management System. Ontologically, DeRTEL is conceptualised as a methodology rooted in pragmatism. Epistemologically, the prototyping and assessment phases proposed in DeRTEL follow Dewey's model of enquiry, i.e. suggesting a possible solution to an identified problem and reflecting on the logical design and practical implications of the solution. Axiologically, the values of utility and ethics underpins the development of the solution proposed in this thesis.

The steps proposed in the DeRTEL methodology are used to frame the narrative discussion of the research process followed in this study (Sections 6.2.1, 6.2.2 and 6.2.3).

6.2.1 Preliminary Phase

In the preliminary phase of the research design, the general context, problem domain and solution domains are described. The problem identified and addressed in this study is:

There is limited prescriptive guidance on how to create a meaningful learner profile from Moodle logs that can inform differentiated learning design choices in Moodle, leading to inadequate instructional designs.

The nature of the problem addressed in this study is situated in the domain of online instructional design. Consequently, literature is reviewed on the form and function of procedural and conceptual instructional design models (Section 3.2). The following are chosen as example of instructional design based on their widespread use:

Procedural Models

- ADDIE
- Dick and Carey
- Gerlach-Ely
- ASSURE
- Backward Design
- Successive Approximation Model (SAM)

Conceptual Models

- Kemp's Model
- 3P Model
- Conceptual Framework of High-Quality Online Learning Environments

The procedural models emphasise that instructional design should be based on learner needs and advocate iterative refinement based on continual evaluation of the learning environment. These refinements are often driven by generic heuristic learning design guidelines suited to a stereotypical learner or based solely on analysis of learner performance metrics. Analysis of learner behaviours in the online learning environment is not yet as widely used to inform the optimisation of learning design. Online learning design (Section 3.3) is increasingly being instantiated in Learning Management Systems that provide an array of tools to deliver and manage instruction (Section 3.4). Learner Management Systems keep track of all learner behaviours while they interact with the learning material. These large datasets provide a prime opportunity to analyse learners' online behaviours, which in turn can inform changes to be made to the instructional design. The field of Learning Analytics uses educational data mining techniques to find meaning in large datasets (Section 4.2). Learning Analytics can, therefore, be used to build a dynamic learner profile from these large datasets (Section 4.3). The literature review of the problem and solution domains informed the objectives and scope of the proposed solution:

- Objective 1: Develop and evaluate a comprehensive, learner-centric process model to enable differentiated instruction based on a dynamic learner profile (Sections 5.2.3.1 and 5.3).
- Objective 2: Establish requirements for a Learning Design Phase (Section 5.2.3.2), instantiated in Moodle (Section 5.4).
- Objective 3: Establish requirements for a Learner Modelling Phase (Section 5.2.3.3), instantiated in Moodle (Section 5.5).

6.2.2 Prototyping Phase

The prototyping phase for DeRTEL involves the development and evaluation of a proposed solution through several iterations.

Iteration 1: The first iteration prepares a tentative proposal and evaluates the proposal for relevance. The form of the proposed solution is expressed through the three objectives identified for this study. The relevance of these objectives is argued in:

- Section 5.2.3.1, concluding that there is a need for a process model for differentiated instruction based on a dynamic profile
- Sections 5.2.3.2 and 5.2.3.3, concluding that there is a need for the learning design and learner modelling to be further refined for use in a Learning Management System

From iteration 1, the study concludes that the proposed model addresses a real need and the form of the tentative solution is based on current literature.

Iteration 2: The second iteration is responsible for preparing the global design of the solution and evaluating the design for consistency/construct validity.

Objective 1: Develop and evaluate a comprehensive, learner-centric process model to enable differentiated instruction based on a dynamic learner profile

In the context of this thesis, the proposed model (Objective 1) describes how differentiated instruction can be provided based on a learner profile, dynamically maintained through implicit learner modelling and guided by an ethical learning analytics code of practice (Section 5.3). The model is derived by integrating steps of a tried-and-tested web analytics process with pedagogical considerations, ethical reflection and a layered abstraction of online learning design. The thesis recommends optimisation of the learning environment should be conducted in three phases: a preliminary goal setting phase, a learning design phase (initial and differentiated learning design) and a learner modelling phase (Figure 6.1).

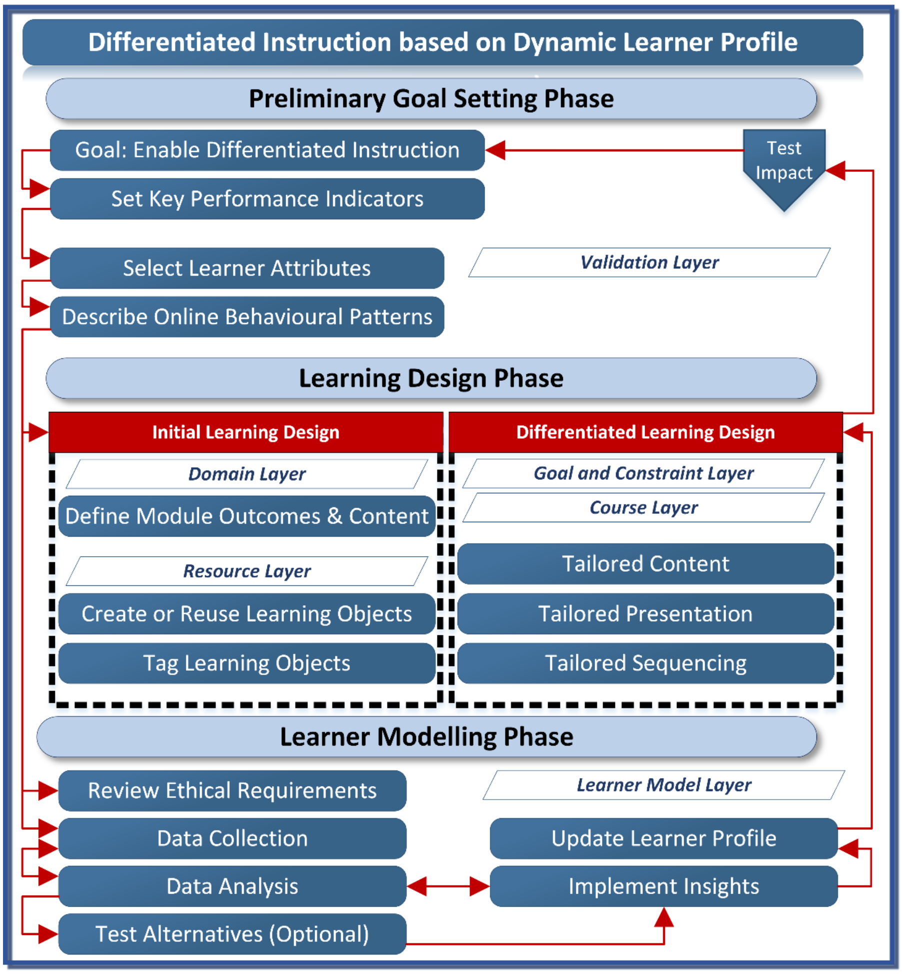


Figure 6.1 Differentiated Instruction based on a Dynamic Learner Profile (Own Construction)

This thesis argues that pedagogical alignment between learning design and learner modelling can be realised if the process is initiated with a preliminary goal setting phase rooted in educational theory (Figure 6.2). It is recommended that all learning design and learner modelling activities should be informed by a clear educational goal and the setting of KPIs to measure whether the goal was achieved. Furthermore, to illustrate the steps in the goal setting phase the thesis presents the identification of learner attributes with an impact on online learning. These learner attributes are assembled from dominant learning style theories hypothesising that learners with dichotomous attributes will behave differently when interacting with learning material. The learning design should incorporate affordances to accommodate these behavioural differences and the learner modelling should examine the different behaviours to infer relevant learner attributes. This thesis is based on a belief that learner behaviours are not fixed but will likely change over time and within different contexts. The learner profile should, therefore, be a dynamic entity that continuously evolves.

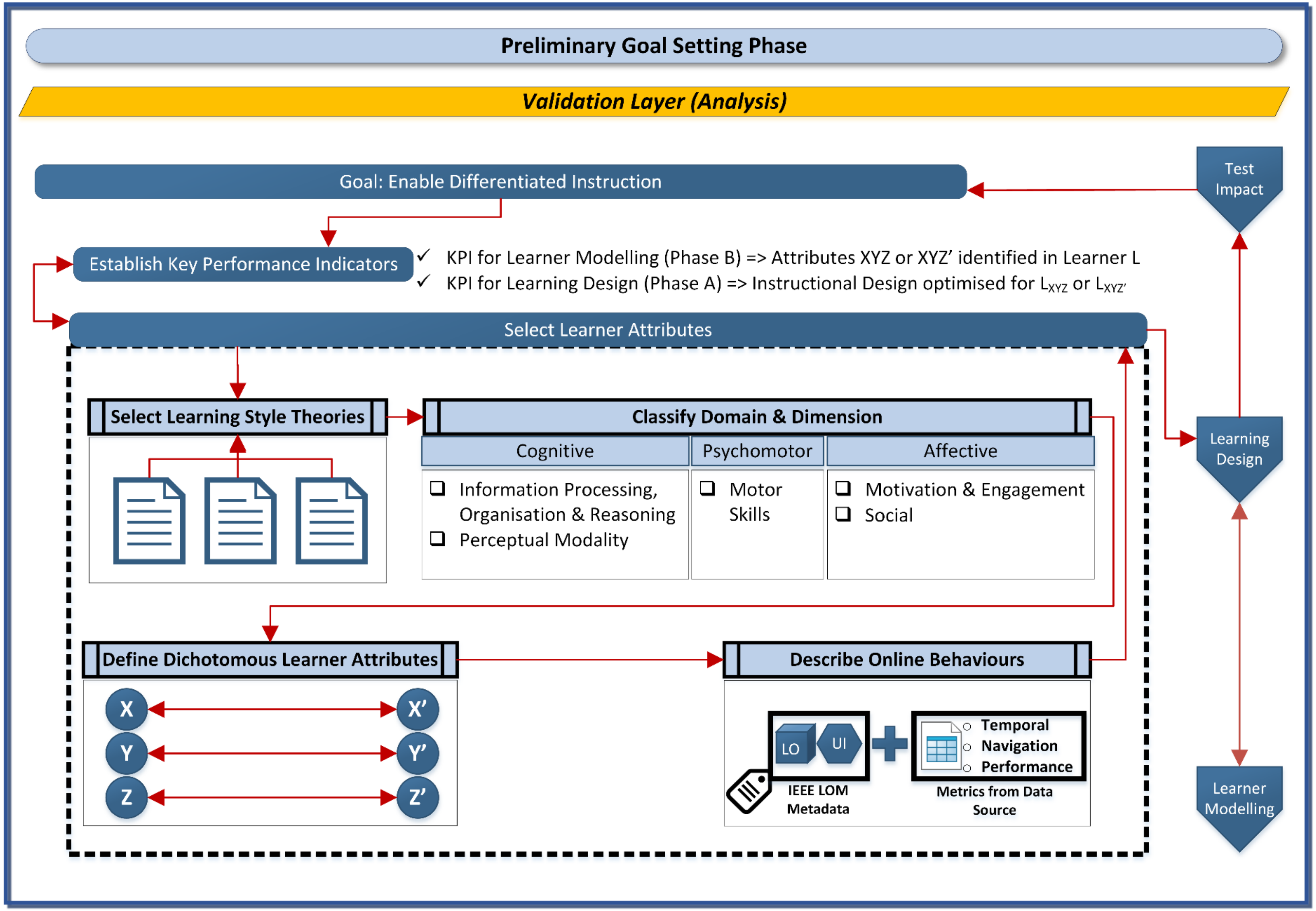


Figure 6.2 Preliminary Goal Phase

It is proposed that the learning design should be implemented in two parts, the initial learning design before learner modelling and differentiated learning design based on updates to the learner profile (Figure 6.3). The initial learning design is informed by hypothesised differences in a stereotypical learner, while the differentiated design tailors the course material based on the actual behaviours and choices made by learners interacting with the course material.

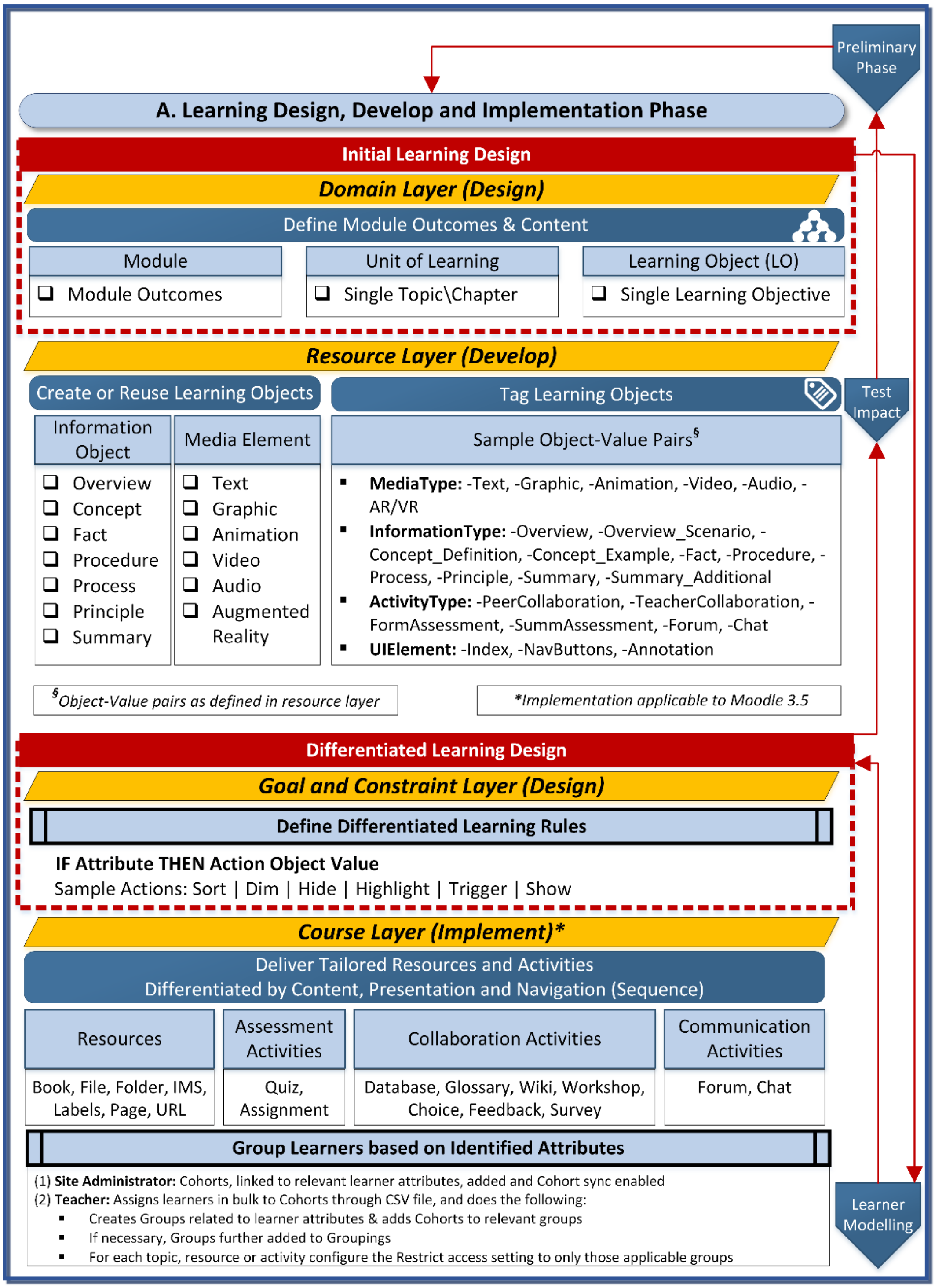


Figure 6.3 Learning Design Phase

Finally, the thesis proposes that learner modelling (Figure 6.4) should always be accompanied by ethical reflection, especially on privacy and equity issues as they relate to learning analytics. Data should be collected from any relevant source, internal or external to the LMS, and analysed using appropriate analysis techniques as dictated by the overall aim set during the preliminary goal setting phase. Data analysis involve pre-processing, applying educational data mining and evaluating the results of the analysis. If alternative hypotheses arise after data analysis, these could be further explored before the learner profile is updated. The learner profile consists of static data, provided by the learner, and dynamic data resulting from inferences made during data analysis. The profile should be open for inspection by the learner before it is used for tailoring of the learning design.

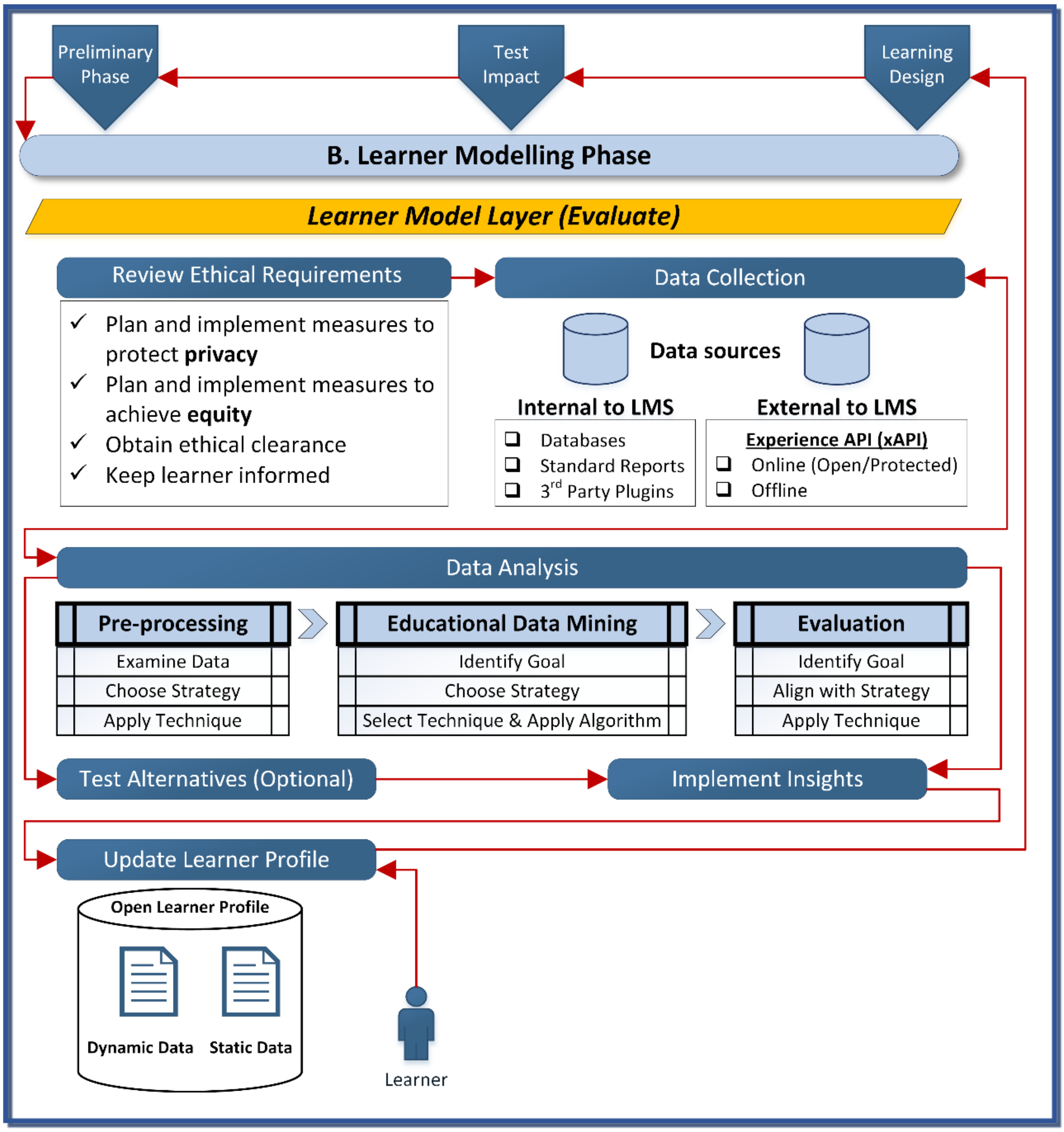


Figure 6.4 Learner Modelling Phase

From iteration 2, the study concludes that the three phases proposed in the model (Appendix D) are viable and logically designed when compared to existing solutions and best practices (Section 5.3.4).

Iteration 3: The third iteration is responsible for refining the proposed model's learning design phase aimed at instantiation in a Learning Management System (Objective 2, Section 5.4).

Objective 2: Establish requirements for a Learning Design Phase, instantiated in Moodle

The evaluation of iteration 3 focuses on the expected utility and practicality of the initial and differentiated learning design in Moodle. The requirements are derived by exploring the following questions:

- What are the typical behavioural patterns associated with learner attributes selected from learning style theories? (Section 5.4.1)
- How can the initial learning design be implemented in a Learning Management System? (Section 5.4.2)
- How can the desired differentiation be achieved in a Learning Management System? (Section 5.4.3)

From iteration 3, the study concludes that dichotomous attributes with hypothesised differences in behaviour can be extracted from learning style theories (Section 5.4.4). These behaviours reveal learners' cognitive strategies and affective states. To simultaneously cater for and discover the existence of these behavioural differences in a Learning Management System, it is necessary to describe the behaviours in two parts:

- The learning resource or activity preferentially accessed by the learner and/or the affordances built into the UI of the system
- The temporal, navigation and performance metrics recorded by the Learning Management System

Iteration 3 of the study further concludes that Moodle can accommodate the proposed course organisational structure, i.e. Module > Unit of Learning > Learning Object > Information Object > Media Element. Moodle also provides the necessary mechanisms to describe the pedagogic intent behind Learning Objects, to divide learners into groups based on their attributes and to filter or recommend learning material based on the unique requirements of each group. It can therefore be concluded that Moodle has the potential to achieve the goals of the learning design phase, and the mechanisms provided in Moodle have the potential to be usable by teachers with little prior Moodle experience. One aspect of the learning design phase, the creation of Cohorts, requires assistance from the site administrator. This aspect may have implications for the practicality of the grouping of learners.

Iteration 4: The fourth iteration is responsible for refining the proposed model's learner modelling phase aimed at instantiation in a Learning Management System (Objective 3, Section 5.5).

Objective 3: Establish requirements for a Learner Modelling Phase, instantiated in Moodle

The evaluation of iteration 4 focuses on the expected utility and practicality of data collection and analysis in Moodle and other tools. Learner modelling requires the identification of metadata along with temporal, navigation or performance metrics from available data sources. Moodle stores all learner interactions in several tables. Extracting information from these tables requires database administration skills most teachers do not have. For the learner modelling phase to be practical for a standard teacher, easier access to this data is necessary. Moodle provides several reports that can be downloaded in spreadsheet format which most teachers are familiar with. While previous studies have shown it is possible to pre-process the data in the spreadsheets through macros and Microsoft (MS) Excel functions, it places an additional burden on teachers. This study experimented with a few standard data pre-processing techniques and the manual method was found to be technically possible, but highly impractical. Once the data is pre-processed, the actual analysis through data mining techniques is quick. Output is produced within seconds after the process is initiated. The difficulty lies in the pre-processing phase and in extracting pedagogic meaning from the output. The standard teacher would, for example, need to know about educational data mining techniques such as clustering if they wanted to partition learners into appropriate groups. This iteration concludes that, while the expected utility is evident, the laborious nature of manual pre-processing and the required knowledge of educational data mining tools and techniques renders learner modelling in Moodle impractical at this stage. Alterations to the Moodle source code are necessary to improve the practicality of the learner modelling technique proposed in the model contributed in this thesis.

6.2.3 Assessment Phase

The assessment phase is beyond the scope of this study. Before the proposed solution is deployed at a specific higher education institution, a site inspection is necessary to draft a deployment plan. Continual impact studies are necessary to evaluate the effects of the differentiated instructional design on learning effectiveness, efficiency and learner satisfaction. Output from these impact studies would close the cycle of the proposed model and initiate further adjustments to the learning design.

6.3 Discussion and Critical Reflection

A model is a simplified representation used to explain the procedures in a real-world system or event, or a process of steps to simplify a complex task. The primary artefact produced in this study is a model that describes the process of enabling differentiated instruction in a Learning Management System by using learner interaction data recorded by the system.

6.3.1 The Type and Level of Knowledge Contribution

The process model was derived from an established web analytics process model used in business. As such it is a level 2 contribution of the exaptation type, since the learning analytics model extends a known solution for a business problem to a problem in online instructional design. The learner modelling and differentiated learning design phases are also both classified as a level 2 contribution of the exaptation type. The structure of the learner profile and techniques to construct the learner profile are known solutions in learner modelling, but they are rarely applied to an open source Learning Management System. Similarly, the provision of differentiated learning design, while a known solution in proprietary online education platforms, is not frequently done in a Learning Management System that was not developed for tailored instruction.

6.3.2 Theoretical Contribution

The theoretical contribution of this study is linked to Objective 1 (Figure 6.1) and is accepted for publication in the South African Computer Journal (Leppan, Van Niekerk and Botha, 2018). The proposed model is an abstracted representation of a process to optimise online learning design through the provision of differentiated instruction that is informed by dynamic learner modelling.

The narrow focus on the technical issues of data collection and analysis identified in existing learning analytics process models may result in an oversight of privacy and equity issues inherent in learning analytics-based interventions. Additionally, without an explicit goal for learning analytics, there is a perceived lack of pedagogical reflection driving the process. The process model proposed in this thesis addresses these shortcomings in existing learning analytics process models. The proposed model emerged by incorporating steps of a tried and tested web analytics process with educational theory, an ethical code practice for learning analytics and a layered abstraction of online learning design.

6.3.3 Practical Contribution

This study produced two practical contributions directly related to the proposed model and one practical contribution in research design.

The first practical contribution is directly linked to Objective 2. The second objective is an instantiation of the learning design phase of the proposed model in a Learning Management System. This instantiation enabled the evaluation of the suitability of Moodle to provide the initial and differentiated learning design as required by the proposed model. A similar method can be applied to the evaluation of other Learning Management Systems with regard to the instructional design requirements imposed by the proposed model.

The second practical contribution is directly linked to Objective 3. The third objective is an instantiation of the learner modelling phase of the proposed model in a Learning Management System. This instantiation enabled the evaluation of the suitability of Moodle to generate a learner profile in the format required by the proposed model. A similar method can be applied to the evaluation of other Learning Management Systems with regards to the data collection and analysis imposed by the proposed model.

The research design synthesised for this study (DeRTEL) is the third practical contribution. DeRTEL can be used to iteratively develop, evaluate and report on an artefact to be used for Technology Enhanced Learning. This research design was used to plan, execute and document the current study that this thesis reports on.

6.3.4 Limitations of the Research

The experimental nature of the learning design and learner modelling process, coupled with time cost involved to test the model on real data, meant the model could not be fully tested in an actual classroom setting. Nevertheless, since the main risk investigated in the current iteration of the model is technical in nature, the technical risk and efficacy evaluation strategy is used to at least test the two sub-processes in two separate instances. This evaluation on the learning design phase and learner modelling phase reveals that while Moodle provides the necessary utility, it lacks in practicality, especially in the learner modelling phase.

The lack of a suitable tool for the pre-processing of data extracted from Moodle means that the educational data mining techniques could not be used on real data. In addition, the tagging of Moodle with the necessary educational metadata, while suitable for learners, is inadequate for the purposes of data extraction. This highlights the need for alterations to the existing Moodle code base to accommodate the requirements imposed by the proposed model.

While plugins have been developed for analysis of data logged by Moodle, it is beyond the scope to do a detailed analysis of these plugins in the current investigation. Since this study focuses on the establishment of the overall process model and the evaluation into the suitability of a standard Moodle installation to instantiate the process, a thorough examination of existing third-party plugins is not conducted.

The attributes and associated behaviours cited in this study were extracted from descriptions of existing learning style theories. At this stage the behavioural differences are hypothesised and would require implementation of the model in a real classroom to confirm or reject the existence of these hypothesised differences in behaviour.

6.4 Recommendations

From this study flow recommendations for policy and practice (Section 6.4.1), recommendations for refining the proposed model (Section 6.4.2) and recommendations for further research related to the proposed model (Section 6.4.3).

6.4.1 Recommendations for Policy and Practice

The model that emerged from the current investigation is still in its infancy. As such wide-scale adoption at higher education institutions is not recommended. More investigation is needed on a small scale by teachers who like to push the envelope. Those with beliefs similar to the ones that underpin this study, i.e. that instructional design should be informed by learners' actual behaviours, are encouraged to try the proposed model in their online classes and to share their experiences. Over time this will help grow the model into a nascent theory.

A policy recommendation that emerged from the current investigation is the need for the adoption of an institution-wide ethical code of practice regarding learning analytics. The proposed model relies heavily on all stakeholders buying in to the notion of harnessing learners' online data to optimise the learning environment. This acceptance will be enhanced if the institution clarifies issues of privacy and equity, and if a learning analytics culture is already embedded at the institution.

One aspect to consider by practitioners is the notion of equity as it relates to this type of intervention. Traditional learning environments are tailored towards the "middle of the pack". If we are to differentiate instruction to cater for, perhaps, the gifted learner, would we not be increasing the gap further? Conversely, would we not do a disservice to the gifted learner by not tailoring the instruction to provide them with a bigger challenge than the average learner?

6.4.2 Recommendations for Refinement of Solution

Since the model has not yet been fully tested in an authentic environment, there is scope for new insights to be worked back into the phases. The sample questions relating to privacy and equity concerns proposed in the model presented in Figure 5.3 are merely a starting point. Further investigation into the ethical requirements are needed for a more comprehensive treatment of ethical considerations. The Open Learner Model concept also needs further refinement to propose technical ways in which negotiated modelling can be realised.

The model as it is presented in Appendix D includes elements of the instantiation of the process in Moodle. Further experimentation with machine learning algorithms is necessary to add more depth to the analysis step. In addition, the model in Appendix D only reports on the requirements for instantiating the process in Moodle. Different

Learning Management Systems can be examined, and the process can be further adapted to be more widely applicable to any Learning Management System.

6.4.3 Recommendations for Further Research

It is highly recommended that plugins are developed for Moodle to simplify the pre-processing and subsequent analysis through data mining methods. Such tools should be developed to be as user-friendly to teachers as possible. They should be able to perform basic data mining tasks, while at the same time guiding the teacher to select appropriate attributes and features as input for the analysis. In addition, the output should be presented in such a way that it is usable enough for teachers, without knowledge of data mining techniques, to make informed choices. The proposed study should develop requirements that can be used to evaluate existing third-party tools that also pre-process and analyse Moodle data. This can include investigating existing dashboards that use sophisticated analysis techniques to distil data for teachers to learn more about their learners.

The current study only investigated data logged and extracted from Moodle. Learning is not confined to the Learning Management System alone but can happen at any time on any online or offline platform. Consider, for example, a teacher that provides a link to an externally-hosted lesson. Moodle will only record that this link has been accessed at a specific time and not the actions of the learner in this lesson. Or a learner may be required to read a Section in a textbook and subsequently complete an offline activity in response. No data will be logged by any system in such cases of offline learning. To collect and analyse data from online and offline learning experiences, further research is necessary in specifications such as Experience API (xAPI).

The “Test Impact” step is represented in the model (Figure 6.1) as an off-page reference. This shape denotes that the process is yet to be modelled in a follow-up study. This type of impact study would require a full deployment of the model in an online class with real users. The impact study would require investigation into techniques for measuring learning effectiveness (e.g. using Kirkpatrick’s model of evaluation, Section 3.2.2.5), learning efficiency and learner satisfaction based on the learning design choices made to accommodate the learners’ profiles.

6.5 Epilogue

The online learning environment in general and Learning Management Systems in particular are nowhere near the levels of personalisation found on websites like Amazon and Netflix. This thesis is a small step towards achieving similar kinds of recommendations by harnessing learner data generated in a widely-used Learning Management System.

Through this research, a multifaceted and multidisciplinary design-focused research agenda emerged that can be pursued to explore the relationships between various learner attributes and instructional design choices. By applying the contributions of this thesis, educators can identify learner attributes with an effect on online learning. With this ability, they can measure the effect of their instructional design choices against identified learner attributes. Continuously updating learner profiles will enable educators to refine their online instruction to continue meeting the learning needs of their cohort. The ability to build a profile from learner behaviours further enables educators to identify habits of successful students and recommend keys to success.

BIBLIOGRAPHY

Al-Azawei, A., & Badii, A. (2014). State of The Art of Learning Styles-Based Adaptive Educational Hypermedia Systems (Ls-Baehss). *International Journal of Computer Science and Information Technology*, *6*(3), 01-19. https://doi.org/10.5121/ijcsit.2014.6301

Aljafari, R., & Khazanchi, D. (2013). On the Veridicality of Claims in Design Science Research. In *International Conference on System Sciences* (pp. 3747–3756). IEEE. https://doi.org/10.1109/HICSS.2013.427

Allen, M. (2012). *Leaving ADDIE for SAM: An agile model for developing the best learning experiences*. American Society for Training and Development.

Allinson, C., & Hayes, J. (2012). Cognitive Style Index. Pearson Education.

Allinson, C. W., & Hayes, J. (1996). The cognitive style index: A measure of intuition-analysis for organizational research. *Journal of Management Studies*, *33*(1), 119–135.

Anderson, L. W., & Krathwohl, D. R. (2001). *A taxonomy for learning, teaching, and assessing: A revision of Bloom's taxonomy of educational objectives, abridged edition*. (L. W. Anderson, D. R. Krathwohl, P. W. Airasian, K. A. Cruikshank, R. E. Mayer, P. R. Pintrich, J. Raths, & M. C. Wittrock, Eds.). White Plains, NY: Longman.

Anderson, T., & Shattuck, J. (2012). Design-Based Research: A Decade of Progress in Education Research? *Educational Researcher*, *41*(1), 16–25. https://doi.org/10.3102/0013189X11428813

Apter, M. J., & Desselles, M. (2018). Zigzag and zeitgeist: motivation in organizations. *Development and Learning in Organizations*. https://doi.org/https://doi.org/10.1108/DLO-03-2018-0040

Apter, M. J., Mallows, R., & Williams, S. (1998). The development of the motivational style profile. *Personality and Individual Differences*, *24*(1), 7–18.

Atif, Y. (2010). An architectural specification for a system to adapt to learning patterns. *Education and Information Technologies*, *16*(3), 259–279. https://doi.org/10.1007/s10639-010-9125-9

Bailey, M., Dittrich, D., Kenneally, E., & Maughan, D. (2012). Applying Ethical Principles to Information and Communication Technology Research: A Companion to the Department of Homeland Security Menlo Report. US Department of Homeland Security.

Baker, R. S. J. ., & Yacef, K. (2009). The State of Educational Data Mining in 2009: A Review and Future Visions. *Journal of Educational Data Mining*, *1*(1), 3–16.

Baker, R. S. J. D. (2010). Data mining for education. In B. McGaw, P. Peterson, & E. Baker (Eds.), *International Encyclopedia of Education* (3rd ed.). Oxford: Elsevier. https://doi.org/10.4018/978-1-59140-557-3

Baker, R. S. J. D., Buckingham Shum, S., Duval, E., Stamper, J., & Wiley, D. (2012). Educational Data Mining Meets Learning Analytics. *Proceedings of the Second International Conference on Learning Analytics and Knowledge - LAK '12*, (May), 20. https://doi.org/10.1145/2330601.2330613

Barab, S., & Squire, K. (2004). Design-Based Research: Putting a Stake in the Ground. *Journal of the Learning Sciences*, *13*(1), 1–14.

Baskerville, R., Gergor, S., Baiyere, A., Hevner, A., & Rossi, M. (2018). Design Science Research Contributions: Finding a Balance between Artifact and Theory. *Journal of the Association for Information Systems*, *19*(5), 358–376. https://doi.org/10.17705/1jais.00495

Battou, A., Mezouary, A. El, Cherkaoui, C., & Mammass, D. (2011). Towards an Adaptive Learning System Based on a New Learning Object Granularity Approach. *International Journal of Advanced Computer Science and Applications*, *2*(9), 8–14. https://doi.org/10.14569/IJACSA.2011.020902

Bayasut, B., Pramudya, G., & Basiron, H. (2013). The Review of Adaptive Educational Hypermedia System Based on Learning Style. In *Colloquium on Active Learning UTeM* (pp. 1–5).

Bichelmeyer, B. (2005). "The ADDIE Model" – A Metaphor for the Lack of Clarity in the field of IDT. *Educational Media and Technology Yearbook*, *31*, 1–7.

Bichsel, J. (2012). Analytics in Higher Education: Benefits, Barriers, Progress, and Recommendations. *EDUCAUSE: Center for Applied Research*, 1–31. https://doi.org/10.13140/RG.2.1.1064.6244

Biggs, J. B. (2011). *Teaching for quality learning at university: What the student does*. McGraw-Hill Education (UK).

Bijmolt, T. H., Leeflang, P. S., Block, F., Eisenbeiss, M., Hardie, B. G., Lemmens, A., & Saffert, P. (2010). Analytics for Customer Engagement. *Journal of Service Research*, *13*(3), 341–356. https://doi.org/10.1177/1094670510375603

Bogarín, A., Cerezo, R., & Romero, C. (2018). A survey on educational process mining. *Wiley Interdisciplinary Reviews: Data Mining and Knowledge Discovery*, *8*(1). https://doi.org/10.1002/widm.1230

Brusilovsky, P. (1996). Methods and techniques of adaptive hypermedia. *User Modeling and User-Adapted Interaction*, *6*(2–3), 87–129.

Brusilovsky, P. (2012). Adaptive Hypermedia for Education and Training. *Adaptive Technologies for Training and Education*, *46*, 46–68. https://doi.org/10.1007/BF00143964

Brusilovsky, P., & Millán, E. (2007). User Models for Adaptive Hypermedia and Adaptive Educational Systems. In P. Brusilovsky, A. Kobsa, & W. Nejdl (Eds.), *The adaptive web*. Springer Berlin Heidelberg. https://doi.org/10.1007/978-3-540-72079-9_1

Brusilovsky, P., Wade, V., & Conlan, O. (2007). From Learning Objects to Adaptive Content Services for E-Learning. In C. Pahl (Ed.), *Architecture Solutions for E-Learning Systems* (pp. 243–261). IGI Global.

Bull, S. (2016). Negotiated learner modelling to maintain today's learner models. *Research and Practice in Technology Enhanced Learning*, *11*(1), 10. https://doi.org/10.1186/s41039-016-0035-3

Bull, S., & Kay, J. (2007). Student Models that Invite the Learner In: The SMILI:() Open Learner Modelling Framework. *International Journal of Artificial Intelligence in Education*, *17*, 89–120. Retrieved from http://iospress.metapress.com/index/D1363L5752666874.pdf

Bull, S., & Kay, J. (2010). Open learner models. In R. Nkambou, J. Bourdeau, & R. Mizoguchi (Eds.), *Studies in Computational Intelligence* (Vol. 308). Springer Berlin Heidelberg. https://doi.org/10.1007/978-3-642-14363-2_15

Carter, S., & Yeo, A. C.-M. (2016). Students-as-customers' satisfaction, predictive retention with marketing implications. *International Journal of Educational Management*, *30*(5), 635–652. https://doi.org/10.1108/17465681011017255

Carver, C., Howard, R., & Lane, W. (1999). Enhancing student learning through hypermedia courseware and incorporation of student learning styles. *IEEE Transactions on Education*, *42*(1), 33–38. https://doi.org/10.1109/13.746332

Chatti, M. A., Dyckhoff, A. L., Schroeder, U., & Thüs, H. (2012). A Reference Model for Learning Analytics. *International Journal of Technology Enhanced Learning*, *4*(5/6), 318–331. https://doi.org/10.1504/ijtel.2012.051815

Chembrakuzhi, M., & Haneefa, M. (2014). Metadata Standards for Open Educational Resources. In *59th ILA International Conference on Managing Libraries in the Changing Information World* (pp. 107–110).

Çinici, M. A., & Altun, A. (2018). Reusable content matters: a learning object authoring tool for smart learning environments. *Smart Learning Environments*, *5*(10). https://doi.org/10.1186/s40561-018-0060-3

CISCO Systems. (2001). Reusable Learning Object Strategy, (November), 48.

Clow, D. (2012). The learning analytics cycle. In *Proceedings of the 2nd International Conference on Learning Analytics and Knowledge* (pp. 134–138). https://doi.org/10.1145/2330601.2330636

Coffield, F., Moseley, D., Hall, E., & Ecclestone, K. (2004). *Learning styles and pedagogy in post-16 learning: A systematic and critical review*. Learning and Skills Research Centre London.

Conijn, R., Snijders, C., Kleingeld, A., & Matzat, U. (2017). Predicting student performance from LMS data. *IEEE Transactions on Learning Technologies*, *10*(1), 17–29. https://doi.org/10.1109/TLT.2016.2616312

Conole, G. (2014). The 7Cs of Learning Design – A new approach to rethinking design practice. In *9th International Conference on Networked Learning* (pp. 502–509).

Cook, D. A. (2012). Revisiting cognitive and learning styles in computer-assisted instruction: not so useful after all. *Academic Medicine*, *87*(6), 778–784. https://doi.org/10.1097/ACM.0b013e3182541286

Cuevas, J. (2015). Is learning styles-based instruction effective? A comprehensive analysis of recent research on learning styles. *Theory and Research in Education*, *13*(3), 308–333. https://doi.org/10.1177/1477878515606621

D'Mello, S., & Graesser, A. (2009). Automatic detection of learner's affect from gross body language. *Applied Artificial Intelligence*, *23*(2), 123–150. https://doi.org/10.1080/08839510802631745

De Arriba-Pérez, F., Santos-Gago, J. M., & Caeiro-Rodríguez, M. (2016). Analytics of biometric data from wearable devices to support teaching and learning activities. *Journal of Information Systems Engineering & Management*, *1*(1), 41–54. https://doi.org/10.20897/lectito.201608

De Bra, P., Aerts, A., Berden, B., De Lange, B., Rousseau, B., Santic, T., Smits, D., & Stash, N. (2003). AHA! The adaptive hypermedia architecture. In *Proceedings of the Fourteenth ACM Conference on Hypertext and Hypermedia* (pp. 81–84). https://doi.org/10.1145/900065.900068

De Rossi, M., & Trevisan, O. (2018). Technological pedagogical content knowledge in the literature: how TPCK is defined and implemented in initial teacher education. *Italian Journal of Educational Technology*, *26*(1), 7–23. https://doi.org/10.17471/2499-4324/988

Department of Higher Education and Training. (2013). *White Paper for Post-School Education and Training*. https://doi.org/10.5663/aps.v1i1.10138

Design-Based Research Collective. (2003). Design-Based Research: An Emerging Paradigm for Educational Inquiry. *American Educational Research Association*, *32*(1), 5–8.

Despotović-Zrakić, M., Marković, A., Bogdanović, Z., Barać, D., & Krčo, S. (2012). Providing Adaptivity in Moodle LMS Courses Adaptive e-learning systems. *Educational Technology & Society*, *15*, 326–338.

Dewey, J. (1997). *How we think*. *The Problem of Training Thought*. Dover Publications. https://doi.org/10.1037/10903-000

Dick, W., Carey, L., & Carey, J. O. (2014). *The Systematic Design of Instruction*. (J. Johnston, Ed.) (8th ed.). Pearson.

Dousay, T., & Logan, R. (2010). Analyzing and Evaluating the Phases of ADDIE. In *Design & Development Research Conference* (pp. 32–43). https://doi.org/10.13140/2.1.1715.5206

Drachsler, H., & Greller, W. (2012). The pulse of learning analytics understandings and expectations from the stakeholders. In *Proceedings of the 2nd International Conference on Learning Analytics and Knowledge* (pp. 120–129). https://doi.org/10.1145/2330601.2330634

Dunn, R., & Dunn, K. (2014). *Learning styles: Online learning style assessments and community*. Retrieved from http://www.learningstyles.net/ (Accessed: December 4, 2018)

Dunn, R. S., Dunn, K. J., & Price, G. E. (1981). *Learning style inventory*. Price Systems New York, Lawrence, KS.

Entwistle, N., Hanley, M., & Hounsell, D. (1979). Identifying distinctive approaches to studying. *Higher Education*, *8*(4), 365–380. https://doi.org/10.1007/BF01680525

Entwistle, N., & Ramsden, P. (2015). *Understanding student learning (Routledge revivals)*. Routledge.

Falvo, B. D. A., & Johnson, B. F. (2007). The Use of Learning Management Systems in the United States. *TechTrends*, *51*(2), 40–45. https://doi.org/10.1007/s11528-007-0025-9

Frank, E., Hall, M. A., & Witten, I. H. (2016). The WEKA Workbench. *Morgan Kaufmann, Fourth Edition*, 553–571.

Freeth, D., & Reeves, S. (2004). Learning to work together: Using the presage, process, product (3P) model to highlight decisions and possibilities. *Journal of Interprofessional Care*, *18*(1), 43–56. https://doi.org/10.1080/13561820310001608221

Friesen, N., Fisher, S., & Roberts, A. (2003). *CanCore Guidelines Version 2.0: Educational Category*.

Friesen, N., Fisher, S., & Roberts, A. (2004). *CanCore Guidelines for the implementation of LOM*.

Gagné, R. M. (1965). *Conditions of learning*. New York: Holt, Rinehart and Winston.

Gašević, D., Djurić, D., & Devedžić, V. (2006). *Model driven architecture and ontology development*. *Model Driven Architecture and Ontology Development*. https://doi.org/10.1007/3-540-32182-9

Gerlach, V. S., & Ely, D. P. (1971). *Teaching and media: a systematic approach*. Prentice-Hall.

Goswami, S., Chakraborty, S., Ghosh, S., Chakrabarti, A., & Chakraborty, B. (2018). A review on application of data mining techniques to combat natural disasters. *Ain Shams Engineering Journal*, *9*(3), 365–378.

Grabowski, S. (2003). *Teaching and Media: A Systematic Approach*. (Allyn and Bacon, Ed.) (Second). Pearson Education.

Graf, S. (2009). Advanced adaptivity in learning management systems by considering learning styles. *Proceedings of the 2009 IEEE/WIC/ACM International Conference*, *3*, 235–238.

Graf, S., & Kinshuk, K. (2007). Providing adaptive courses in learning management systems with respect to learning styles. In *World Conference on E-Learning in Corporate, Government, Healthcare and Higher Education* (pp. 2576–2583).

Graf, S., Kinshuk, K. B., Khan, F. A., Maguire, P., Mahmoud, A., Rambharose, T., Shtern, V., Tortorella, R., & Zhang, Q. (2012). Adaptivity and Personalization in Learning Systems based on Students' Characteristics and Context. In *The 1st international symposium on smart learning environment* (pp. 33–36).

Graf, S., Kinshuk, K., & Liu, T.-C. (2008). Identifying Learning Styles in Learning Management Systems by Using Indications from Students' Behaviour. In *2008 Eighth IEEE International Conference on Advanced Learning Technologies* (pp. 482–486). https://doi.org/10.1109/ICALT.2008.84

Graf, S., & Kinshuk, T. (2008). Analysing the Behaviour of Students in Learning Management Systems with Respect to Learning Styles. In M. Wallace, M. Angelides, & P. Mylonas (Eds.), *Advances in Semantic Media Adaptation and Personalization*. Springer Berlin Heidelberg. https://doi.org/10.1007/978-3-540-76361_3

Gravett, S., & Geyser, H. (2009). *Teaching and learning in higher education*. (L. Martini, Ed.) (First). Pretoria: Van Schaik.

Gregor, S., & Hevner, A. R. (2013). Positioning and Presenting Design Science Research for Maximum Impact. *MIS Quarterly*, *37*(2), 337–355. https://doi.org/10.25300/misq/2013/37.2.01

Gregor, S., & Jones, D. (2007). The Anatomy of a Design Theory. *Journal of the Association for Information Systems*, *8*(5), 312–335.

Gregorc, A. (2006). The mind styles model: theory, principles, and practice. *AFG, Columbia*.

Gregorc, A. F. (1982). The mind styles model: theory, principles, and applications. *Maynard, MA: Gabriel Systems*.

Griffiths, D. (2017). The Use of Models in Learning Design and Learning Analytics. *Interaction Design and Architecture(s) Journal*, *33*, 113–133.

Griffiths, D., Drachsler, H., Kickmeier-Rust, M., Steiner, C., Hoel, T., & Greller, W. (2016). *Is Privacy a Show-stopper for Learning Analytics? A Review of Current Issues and Solutions*. Retrieved from http://www.laceproject.eu/learning-analytics-review/files/2016/04/LACE-review-6_privacy-show-stopper.pdf (Accessed: October 23, 2018)

Gürer, M. D. (2013). *Utilization of Learning Objects in Social Studies Lesson: Achievement, Attitude and Engagement*. Middle East Technical University.

Gustafson, K. L., & Branch, R. M. (2002). What is Instructional Design? In *Trends and issues in instructional design and technology* (pp. 17–25).

Halim, N. D. A., Ali, M. B., & Yahaya, N. (2011). Personalized Learning Environment: Accommodating Individual Differences in Online Learning. In *2011 International Conference on Social Science and Humanity* (Vol. 5, pp. 398–400).

Hall, W., & Keynes, M. (2011). The Net Generation and Digital Natives Implications for Higher Education. *Higher Education Academy*, (June), 1–56.

Heinich, R., Molenda, M., Russell, J., & Smaldino, S. (2002). *Instructional media and technologies for learning. (7th ed.)*. *Pearson Education*.

Herrington, J., McKenney, S., Reeves, T., &, & Olive, R. (2007). Design-based research and doctoral students: Guidelines for preparing a dissertation proposal. In *World Conference on Educational Multimedia, Hypermedia and Telecommunications* (pp. 4089–4097).

Herrmann, N. (1991). The creative brain. *The Journal of Creative Behavior*, *25*(4), 275–295.

Herrmann, N., & Herrmann-Nehdi, A. (2015). *The Whole Brain Business Book: Unlocking the Power of Whole Brain Thinking in Organizations, Teams, and Individuals*. McGraw-Hill Education.

Hevner, A., & Chatterjee, S. (2010). Design Research in Information Systems. In *Integrated Series in Information Systems* (Vol. 22, pp. 9–23). https://doi.org/10.1007/978-1-4419-5653-8

Hevner, A., March, S., Park, J., & Ram, S. (2004). Design Science in Information Systems Research. *MIS Quarterly*, *28*(1), 75–105. https://doi.org/https://doi.org/10.2307/25148625

Hodgins, W. (2006). The Future of Learning objects. *Educational Technology*, *46*(1), 49–54.

Hoel, T., & Chen, W. (2016). Privacy-Driven Design of Learning Analytics Applications: Exploring the Design Space of Solutions for Data Sharing and Interoperability. *Journal of Learning Analytics*, *3*(1), 139–158. https://doi.org/10.18608/jla.2016.31.9

Honey, P., & Mumford, A. (1992). The manual of learning styles.

Honey, P., & Mumford, A. (2006). *Learning styles questionnaire: 80-item version*. Maidenhead.

Hundhausen, C. D., Olivares, D. M., & Carter, A. S. (2017). IDE-Based Learning Analytics for Computing Education: A Process Model, Critical Review and Research Agenda. *ACM Transactions on Computing Education*, *17*(3), 1–26. https://doi.org/10.1145/3105759

IEEE 1484.12.1. (2002a). *IEEE Standard for learning object metadata*. *IEEE Standard* (Vol. 1484).

IEEE 1484.12.1. (2002b). IEEE Standard for Learning Object Metadata. *Learning Technology Standards Committee of the IEEE*, *1484.12.1*-(July), 1–44. https://doi.org/10.1109/IEEESTD.2002.94128

Iivari, J. (2007). A Paradigmatic Analysis of Information Systems As a Design Science. *Scandanavian Journal of Information Systems*, *19*(2), 39–64. https://doi.org/10.1.1.218.2636

IMS Global Learning Consortium. (2006). *IMS Meta-data Best Practice Guide for Learning Object Metadata*.

Jones, D., Bench-Capon, T., & Visser, P. (1998). Methodologies For Ontology Development. In *IT & KNOWS Conference of the 15th IFIP World Computer Congress*. Budapest: Chapman-Hall.

Jones, D., & Gregor, S. (2006). The formulation of an Information Systems Design Theory for e-learning. In *First International Conference on Design Science Research in Information Systems and Technology* (pp. 356–373).

Jones, D., Gregor, S., & Lynch, T. (2003). An Information Systems Design Theory for Web-based Education. In *IASTED International Symposium on Web-based Education* (pp. 1–6).

Jugo, I., Kovačić, B., & Tijan, E. (2015). Cluster analysis of student activity in a web-based intelligent tutoring system. *Pomorstvo: Scientific Journal of Maritime Research*, *29*(1), 75–83.

Karampiperis, P., & Sampson, D. (2004). Adaptive instructional planning using ontologies. In *Proceedings - IEEE International Conference on Advanced Learning Technologies, ICALT 2004* (pp. 126–130). https://doi.org/10.1109/ICALT.2004.1357388

Kemp, J. E., Ross, S. M., & Morrison, G. R. (1998). *Designing effective instruction*. (D. A. Stollenwerk, Ed.). Prentice Hall.

Kirkpatrick, D. (1996). Revisiting Kirkpatrick's four-level model. *Training and Development*, *50*(1), 54–59.

Kirschner, P. A. (2016). Stop propagating the learning styles myth. *Computers & Education*, *106*, 166–171. https://doi.org/10.1016/j.compedu.2016.12.006

Kolb, A. Y., & Kolb, D. A. (2005). The Kolb learning style inventory-version 3.1 2005 technical specifications. *Boston, MA: Hay Resource Direct*, *200*, 1–72. https://doi.org/10.1016/S0260-6917(95)80103-0

Konstantinidis, A., & Grafton, C. (2013). Using Excel Macros to Analyse Moodle Logs. In *2nd Moodle Research Conference* (pp. 33–39). Tunisia.

Kop, R., & Fournier, H. (2010). New Dimensions to Self-Directed Learning in an Open Networked Learning Environment. *International Journal of Self-Directed Learning*, *7*(2), 1–18.

Kramer, B. J. (2005). Reusable Learning Objects: Let's give it another trial. In *Forschungsberichte des Fachbereichs Elektrotechnik*. Hagen.

Krathwohl, D. R. (2002). A Revision of Bloom's Taxonomy: An Overview. *Theory Into Practice*, *41*(4), 212–218. https://doi.org/10.1207/s15430421tip4104_2

Krathwohl, D. R., Bloom, B. S., & Masia, B. B. (1964). *Taxonomy of educational objectives, Book II: Affective Domain* (Vol. 1). New York: David MacKay Company, Inc.

Kuhn, T. (1977). *Second Thoughts on Paradigms*. *The Structure of Scientific Revolutions*. University of Chicago Press. https://doi.org/10.1119/1.11776

Labib, A. E., Canós, J. H., & Penadés, M. C. (2017). On the way to learning style models integration: a Learner's Characteristics Ontology. *Computers in Human Behavior*, *73*, 433–445. https://doi.org/10.1016/j.chb.2017.03.054

Lee, J., & Jang, S. (2014). A methodological framework for instructional design model development: Critical dimensions and synthesized procedures. *Educational Technology Research and Development*, *62*(6), 743–765. https://doi.org/10.1007/s11423-014-9352-7

Leppan, R. G., Van Niekerk, J. F., & Botha, R. A. (2018). Process Model for Differentiated Instruction using Learning Analytics. *South African Computer Journal*, *30*(2), 17–43. https://doi.org/10.18489/sacj.v30i2.481

Linden, T., & Lederman, R. (2015). Using Building Blocks to Construct Effective Learning Objects. In *PACIS 2015 Proceedings* (p. 248).

Lu, H. (2012). Learning styles and acceptance of e–learning management systems: an extension of behaviour intention model. *International Journal of Mobile Learning and Organisation*, *6*(3), 246–259. https://doi.org/10.1504/IJMLO.2012.050044

Luna, J. M., Castro, C., & Romero, C. (2017). MDM tool: A data mining framework integrated into Moodle. *Computer Applications in Engineering Education*, *25*(1), 90–102. https://doi.org/10.1002/cae.21782

Lwoga, E. (2012). Making learning and Web 2.0 technologies work for higher learning institutions in Africa. *Campus-Wide Information Systems*, *29*(2), 90–107. https://doi.org/10.1108/10650741211212359

Manning, S., Stanford, B., & Reeves, S. (2010). Valuing the advanced learner: Differentiating up. *The Clearing House*, *83*(4), 145–149. https://doi.org/10.1080/00098651003774851

Mcafee, A., & Brynjolfsson, E. (2012). Spotlight on Big Data: The Management Revolution. *Harvard Business Review*, (October), 1–9.

Mehlenbacher, B. (2012). Everyday Learning: Instruction and Technology Designs. In N. M. Seel (Ed.), *Encyclopedia of the Sciences of Learning* (pp. 1181–1185). Boston, MA: Springer US. https://doi.org/10.1007/978-1-4419-1428-6_1769

Melia, M., & Pahl, C. (2007). An Information Architecture for Courseware Validation. In *First International Workshop on Learning Object Discovery & Exchange (LODE'07)*.

Melia, M., & Pahl, C. (2009). Constraint-Based Validation of Adaptive e-Learning Courseware. *IEEE Transactions on Learning Technologies*, *2*(1), 37–49. https://doi.org/10.1109/TLT.2009.7

Merceron, A., Blikstein, P., & Siemens, G. (2015). Learning Analytics: From Big Data to Meaningful Data. *Journal of Learning Analytics*, *2*(3), 4–8. https://doi.org/10.18608/jla.2015.23.2

Minielli, M., & Ferris, S. (2005). Electronic courseware in higher education. *First Monday*, *10*(9). Retrieved from http://www.firstmonday.dk/ojs/index.php/fm/article/view/1279/1199 (Accessed: December 1, 2018)

Mishra, D., Gunasekaran, A., Papadopoulos, T., & Childe, S. J. (2018). Big Data and supply chain management: a review and bibliometric analysis. *Annals of Operations Research*, *270*(1–2), 313–336.

Mishra, P., & Koehler, M. J. (2006). Technological Pedagogical Content Knowledge: A Framework for Teacher Knowledge. *Teachers College Record*, *108*(6), 1017–1054. https://doi.org/10.1111/j.1467-9620.2006.00684.x

Molenda, M. (2003). In Search of the Elusive ADDIE Model. *Perfromance Improvement*, *42*(5), 34–36. https://doi.org/10.1002/pfi.4930420508

Moodle. (2018). Moodle Statistics. Retrieved December 1, 2018, from https://moodle.net/stats/

Moodle Docs. (2018a). Moodle Activities. Retrieved December 1, 2018, from https://docs.moodle.org/35/en/Activities

Moodle Docs. (2018b). Moodle Resources. Retrieved December 1, 2018, from https://docs.moodle.org/35/en/Resources

Mor, Y., Craft, B., & Hernández-Leo, D. (2013). Editorial: The art and science of learning design. *Research in Learning Technology*, *21*. https://doi.org/10.3402/rlt.v21i0.22513

Mor, Y., Ferguson, R., & Wasson, B. (2015). Editorial: Learning design, teacher inquiry into student learning and learning analytics: A call for action. *British Journal of Educational Technology*, *46*(2), 221–229. https://doi.org/10.1111/bjet.12273

Morgan, D. L. (2014). Pragmatism as a Paradigm for Social Research. *Qualitative Inquiry*, *20*(8), 1045–1053. https://doi.org/10.1177/1077800413513733

Morrison, G. R., Ross, S. M., Kemp, J. E., & Kalman, H. (2010). *Designing effective instruction*. John Wiley & Sons.

Mudrick, N. V, Azevedo, R., & Taub, M. (2018). Integrating metacognitive judgments and eye movements using sequential pattern mining to understand processes underlying multimedia learning. *Computers in Human Behavior*.

Mulwa, C., Lawless, S., Sharp, M., Arnedillo-Sanchez, I., & Wade, V. (2010). Adaptive Educational Hypermedia Systems in Technology Enhanced Learning: A Literature Review. In *2010 ACM Conference on Information Technology Education* (pp. 73–84). ACM.

Myers, I. B., Kirby, L. K., & Myers, K. D. (2011). *Introduction to type : a guide to understanding your results on the MBTI instrument*. Parkville, Vic.: CPP Asia Pacific Pty Ltd.

Myers, I. B., McCaulley, M. H., & Most, R. (1985). *Manual: A guide to the development and use of the Myers-Briggs Type Indicator* (Vol. 1985). Consulting Psychologists Press Palo Alto, CA.

NCPHS. (1979). The Belmont Report: Ethical Principles and Guidelines for the Protection of Human Subjects of Research. *OPPR Reports*, *44*(76), 1–5. https://doi.org/10.1002/9780471462422.eoct093

Nicolae, M. (2014). Teachers' Beliefs as the Differentiated Instruction Starting Point: Research Basis. *Procedia - Social and Behavioral Sciences*, *128*, 426–431. https://doi.org/10.1016/j.sbspro.2014.03.182

Nithya, P., Umamaheswari, B., & Umadevi, A. (2016). A Survey on Educational Data Mining in Field of Education. *International Journal of Advanced Research in Computer Engineering & Technology (IJARCET)*, *5*(1), 69–78.

Özyurt, Ö., & Özyurt, H. (2015). Learning style based individualized adaptive e-learning environments: Content analysis of the articles published from 2005 to 2014. *Computers in Human Behavior*, *52*, 349–358. https://doi.org/10.1016/j.chb.2015.06.020

Palavitsinis, N., Manouselis, N., & Sanchez-Alonso, S. (2014). Metadata quality in learning object repositories: a case study. *The Electronic Library*, *32*(1), 62–82. https://doi.org/10.1108/EL-12-2011-0175

Papamitsiou, Z., & Economides, A. A. (2014). Learning Analytics and Educational Data Mining in Practice: A Systematic Literature Review of Empirical Evidence. *Educational Technology & Society*, *17*(4), 49–64.

Papanikolaou, K. A., Grigoriadou, M., Kornilakis, H., & Magoulas, G. D. (2003). Personalizing the Interaction in a Web-based Educational Hypermedia System: the case of INSPIRE. *User Modeling and User-Adapted Interaction*, *13*(3), 213–267. https://doi.org/10.1023/A:1024746731130

Pask, G. (1976). Styles and strategies of learning. *British Journal of Educational Psychology*, *46*(2), 128–148.

Peterson, C. (2003). Bringing ADDIE to life: instructional design at its best. *Journal of Educational Multimedia and Hypermedia*, *12*(3), 1–5. https://doi.org/10.1017/CBO9781107415324.004

Phippen, A., Sheppard, L., & Furnell, S. (2004). A practical evaluation of Web analytics. *Internet Research*, *14*(4), 284–293. https://doi.org/10.1108/10662240410555306

Plomp, T., & Nieveen, N. (2013). Introduction to the collection of illustrative cases of educational design research. In T. Plomp & N. Nieveen (Eds.), *Educational Design Research* (pp. V–XX).

Polson, M. C., & Richardson, J. J. (2013). *Foundations of intelligent tutoring systems*. Psychology Press.

Popescu, E. (2008). *Dynamic adaptive hypermedia systems for e-learning*. University of Technology of Compiègne.

Popescu, E. (2009). Learning styles and behavioral differences in web-based learning settings. In *Advanced Learning Technologies, 2009. ICALT 2009. Ninth IEEE International Conference on* (pp. 446–450). IEEE.

Popescu, E. (2010a). A Unified Learning Style Model for Technology-Enhanced Learning: What, Why and How? *International Journal of Distance Education Technologies (IJDET)*, *8*(3), 65–81.

Popescu, E. (2010b). Learning Styles and Behavioral Differences in Web-based Learning Settings. In *The 5th International Conference on Virtual Learning ICVL 2010* (Vol. 1, pp. 206–213). IEEE. https://doi.org/10.1109/ICALT.2009.156

Popescu, E., Badica, C., & Moraret, L. (2010). WELSA: An Intelligent and Adaptive Web-based Educational System. In *Intelligent Distributed Computing III* (Vol. 34, pp. 175–185). Springer Berlin Heidelberg.

Popescu, E., Badica, C., & Trigano, P. (2008). Learning Objects' Architecture and Indexing in WELSA Adaptive Educational System. *Scalable Computing: Practice and Experience*, *9*(1), 11–20.

Popescu, E., Trigano, P., & Badica, C. (2007a). Evaluation of a Learning Management System for Adaptivity Purposes. In *2007 International Multi-Conference on Computing in the Global Information Technology (ICCGI'07)*. IEEE. https://doi.org/10.1109/ICCGI.2007.24

Popescu, E., Trigano, P., & Badica, C. (2007b). Towards a Unified Learning Style Model in Adaptive Educational Systems. In *Seventh IEEE International Conference on Advanced Learning Technologies*. IEEE Computer Society.

Provost, F., & Fawcett, T. (2013). Data Science and its Relationship to Big Data and Data-Driven Decision Making. *Big Data*, *1*(1), 51–59. https://doi.org/10.1089/big.2013.1508

Rayner, S., & Riding, R. (1997). Towards a categorisation of cognitive styles and learning styles. *Educational Psychology*, *17*(1–2), 5–27.

Reeves, T. (2006). Design research from a technology perspective. In *Educational design research* (pp. 64–78). Routledge.

Richey, R. C., Klein, J. D., & Tracey, M. W. (2011). The instructional design knowledge base: theory, research, and practice. *Journal of Chemical Information and Modeling*, *53*(9). https://doi.org/10.1017/CBO9781107415324.004

Riding, R., & Rayner, S. (2013). *Cognitive styles and learning strategies: Understanding style differences in learning and behavior*. David Fulton Publishers.

Rienties, B., Nguyen, Q., Holmes, W., & Reedy, K. (2017). A review of ten years of implementation and research in aligning learning design with learning analytics at the Open University UK. *Interaction Design and Architecture(s) Journal*, *33*, 134–154.

Ritzhaupt, A. D. (2010). Learning Object Systems and Strategy: A description and discussion. *Interdisciplinary Journal of E-Learning and Learning Objects*, *6*.

Roberts, L. D., Howell, J. A., Seaman, K., & Gibson, D. C. (2016). Student attitudes toward learning analytics in higher education: "The fitbit version of the learning world." *Frontiers in Psychology*, *7*(DEC), 1–11. https://doi.org/10.3389/fpsyg.2016.01959

Rodriguez, V., & Ayala, G. (2012). Adaptivity and Adaptability of Learning Objects Interface. *International Journal of Computer Applications*, *37*(1), 6–13. https://doi.org/10.5120/4570-6535

Roiger, R. J. (2017). *Data Mining: A Tutorial-Based Primer*. https://doi.org/10.1016/B978-1-4557-4376-6.00039-1

Romero, C., & Ventura, S. (2010). Educational Data Mining: A Review of the State of the Art. *IEEE Transactions on Systems, Man, and Cybernetics, Part C (Applications and Reviews)*, *40*(6), 601–618. https://doi.org/10.1109/TSMCC.2010.2053532

Romero, C., & Ventura, S. (2013). Data mining in education. *Wiley Interdisciplinary Reviews: Data Mining and Knowledge Discovery*, *3*(1), 12–27. https://doi.org/10.1002/widm.1075

Romero, C., Ventura, S., & García, E. (2008). Data mining in course management systems: Moodle case study and tutorial. *Computers and Education*, *51*(1), 368–384. https://doi.org/10.1016/j.compedu.2007.05.016

Saunders, M., Lewis, P., & Tornhill, A. (2007). *Research Methods for Business Students*. *Pearson Education Limited 2*. https://doi.org/10.1007/s13398-014-0173-7.2

Scotland, J. (2012). Exploring the philosophical underpinnings of research. *English Language Teaching*, *5*(9), 9–16. https://doi.org/10.5539/elt.v5n9p9

Seyal, A. H., & Rahman, M. N. A. (2015). Understanding Learning Styles, Attitudes and Intentions in Using e-Learning System: Evidence from Brunei. *World Journal of Education*, *5*(3), 61–72. https://doi.org/10.5430/wje.v5n3p61

Sharma, K. (2015). *Gaze Analysis methods for Learning Analytics*.

Shea, P., Pickett, A., & Pelz, W. (2003). A Follow Up Investigation of “Teacher Presence” in the SUNY Learning Network. *Journal of Asynchronous Learning Networks*, *7*(2), 61–80.

Siemens, G. (2013). Learning Analytics: The Emergence of a Discipline. *American Behavioral Scientist*, *57*(10), 1380–1400. https://doi.org/10.1177/0002764213498851

Siemens, G., & Baker, R. S. J. D. (2012). Learning analytics and educational data mining. *Proceedings of the 2nd International Conference on Learning Analytics and Knowledge - LAK ’12*, 252. https://doi.org/10.1145/2330601.2330661

Siemens, G., & Long, P. (2011). Penetrating the Fog: Analytics in Learning and Education. *EDUCAUSE Review*, *46*, 30–32. https://doi.org/10.1145/2330601.2330605

Simon, H. A. (1996). The Sciences of the Artificial. *Computers & Mathematics with Applications*, *33*(5). https://doi.org/10.1016/S0898-1221(97)82941-0

Sinclair, J., & Aho, A. (2016). Investigating the Experience of Moodle Adoption Through Expert Voices. In L. Uden, D. Liberona, & B. Feldmann (Eds.), *International Workshop on Learning Technology for Education in Cloud* (pp. 90–102). Cham: Springer International Publishing. https://doi.org/10.1007/978-3-319-42147-6_8

Sinclair, J., & Aho, A. (2017). Experts on super innovators: understanding staff adoption of learning management systems. *Higher Education Research & Development*, *37*(1), 158–172. https://doi.org/10.1080/07294360.2017.1342609

Slade, S., & Prinsloo, P. (2013). Learning Analytics: Ethical Issues and Dilemmas. *American Behavioral Scientist*, *57*(10), 1510–1529. https://doi.org/10.1177/0002764213479366

Smith, P. L., & Ragan, T. J. (1999). *Instructional design*. Wiley New York.

Souri, A., & Hosseini, R. (2018). A state-of-the-art survey of malware detection approaches using data mining techniques. *Human-Centric Computing and Information Sciences*, *8*(1), 3.

Steiner, C. M., Kickmeier-Rust, M. D., & Albert, D. (2016). LEA in Private: A Privacy and Data Protection Framework for a Learning Analytics Toolbox. *Journal of Learning Analytics*, *3*(1), 66–90. https://doi.org/10.18608/jla.2016.31.5

Thakurta, R., Mueller, B., Ahlemann, F., & Hoffmann, D. (2017). The State of Design – A Comprehensive Literature Review to Chart the Design Science Research Discourse. In *50th Hawaii International Conference on System Sciences* (pp. 4685–4694). https://doi.org/978-0-9981331-0-2

Tomlinson, C. A., Brighton, C., Hertberg, H., Callahan, C. M., Moon, T. R., Brimijoin, K., Conover, L. A., & Reynolds, T. (2003). Differentiating instruction in response to student readiness, interest, and learning profile in academically diverse classrooms: A review of literature. *Journal for the Education of the Gifted*, *27*(2–3), 119–145. https://doi.org/10.1177/016235320302700203

Tsai, Y.-S., & Gasevic, D. (2017). Learning analytics in higher education --- challenges and policies. *Proceedings of the Seventh International Learning Analytics & Knowledge Conference on - LAK '17*, (October), 233–242. https://doi.org/10.1145/3027385.3027400

Vaishnavi, V. K., & Kuechler, W. (2008). *Design science research methods and patterns: innovating information and communication technology*. Auerbach Publications. https://doi.org/10.1201/9781420059335

Van Den Akker, J. (1999). Principles and Methods of Development Research. In J. Van Den Akker, R. M. Branch, K. Gustafson, N. Nieveen, & T. Plomp (Eds.), *Design Approaches and Tools in Education and Training* (pp. 1–14). Springer Netherlands.

Van Den Akker, J., Bannan, B., Kelly, A. E., Nieveen, N., & Plomp, T. (2010). *An Introduction to Educational Design Research*. (T. Plomp & N. Nieveen, Eds.) (Vol. 3).

Van Seters, J. R., Ossevoort, M. A., Tramper, J., & Goedhart, M. J. (2012). The influence of student characteristics on the use of adaptive e-learning material. *Computers and Education*, *58*(3), 942–952. https://doi.org/10.1016/j.compedu.2011.11.002

Venable, J., Pries-Heje, J., & Baskerville, R. (2016). FEDS: A Framework for Evaluation in Design Science Research. *European Journal of Information Systems*, *25*(1), 77–89. https://doi.org/10.1057/ejis.2014.36

Verbert, K., Duval, E., Klerkx, J., Govaerts, S., & Santos, J. L. (2013). Learning Analytics Dashboard Applications. *American Behavioral Scientist*, *57*(10), 1500–1509. https://doi.org/10.1177/0002764213479363

Verdú, E., & Regueras, L. (2008). An analysis of the research on adaptive learning: the next generation of e-learning. *WSEAS Transactions on Information Science and Applications*, *5*(6), 859–868.

Vermunt, J. D. (1996). Metacognitive, cognitive and affective aspects of learning styles and strategies: A phenomenographic analysis. *Higher Education*, *31*(1), 25–50. https://doi.org/10.1007/BF00129106

Vermunt, J. D., & Donche, V. (2017). A Learning Patterns Perspective on Student Learning in Higher Education: State of the Art and Moving Forward. *Educational Psychology Review*, *29*(2), 269–299. https://doi.org/10.1007/s10648-017-9414-6

Viberg, O., Hatakka, M., Bälter, O., & Mavroudi, A. (2018). The Current Landscape of Learning Analytics in Higher Education. *Computers in Human Behavior*, *89*(December), 98–110. https://doi.org/10.1016/j.chb.2018.07.027

Vinay, A., Shekhara, V. S., Rituparna, J., Aggrawal, T., Balasubramanya Murthy, K. N., & Natarajan, S. (2015). Cloud based big data analytics framework for face recognition in social networks using machine learning. *Procedia Computer Science*, *50*, 623–630. https://doi.org/10.1016/j.procs.2015.04.095

Wadhawan, R. (2018). Prediction of Coronary Heart Disease Using Apriori algorithm with Data Mining Classification. *International Journal of Research in Science and Technology*, *3*(1), 1–15.

Waisberg, D. (2015). *Google Analytics Integrations*. John Wiley & Sons.

Wang, S. (2008). Ontology of Learning Objects Repository for Pedagogical Knowledge Sharing. *Interdisciplinary Journal of E-Learning and Learning Objects*, *4*(4), 1.

Webb, M., & Cox, M. (2004). A review of pedagogy related to information and communications technology. *Technology, Pedagogy and Education*, *13*(3), 235–286. https://doi.org/10.1080/14759390400200183

Wiggins, G. P., & McTighe, J. (2005). *Understanding by design*. (J. Houtz, Ed.), *Association for Supervision and Curriculum Development* (Second).

Wiley, D. (2002). Connecting learning objects to instructional design theory: A definition, a metaphor, and a taxonomy. In D. A. Wiley (Ed.), *The instructional use of learning objects* (1st ed.). Agency for Instructional Technology.

Willis, J. E., Slade, S., & Prinsloo, P. (2016). Ethical oversight of student data in learning analytics: a typology derived from a cross-continental, cross-institutional perspective. *Educational Technology Research and Development*, *64*(5), 881–901. https://doi.org/10.1007/s11423-016-9463-4

Witkin, H., & Moore, C. (1977). Field-dependent and field-independent cognitive styles and their educational implications. *Review of Educational Research*, *47*(1), 1–64.

Wolf, C. (2003). iWeaver: Towards learning style-based e-learning in computer science education. In *Fifth Australasian conference on Computing education* (pp. 273–279). Australian Computer Society, Inc.

Zajonc, R. . (1980). Feeling and thinking: Preferences need no inferences. *American Psychologist*, *35*(2), 151–175. https://doi.org/http://dx.doi.org/10.1037/0003-066X.35.2.151

Zhang, H., Rao, H., & Feng, J. (2018). Product innovation based on online review data mining: a case study of Huawei phones. *Electronic Commerce Research*, *18*(1), 3–22.

APPENDIX A: PUBLICATION STEMMING FROM THIS STUDY

SACJ 30(2) December 2018

Research Article

Process Model for Differentiated Instruction using Learning Analytics

Ronald G. Leppan[a], Johan F. van Niekerk[a, b], Reinhardt A. Botha[a]

[a]School of ICT, Nelson Mandela University, South Africa

[b]Noroff University College, Kristiansand, Norway

ABSTRACT

Higher education institutions seem to have a haphazard approach to harnessing the ubiquitous data that learners generate on online educational platforms, despite promising opportunities offered by this data. Several learning analytics process models have been proposed to optimise the learning environment based on this learner data. The model proposed in this paper addresses deficiencies in existing learning analytics models that frequently emphasises only the technical aspects of data collection, analysis and intervention, yet remain silent on ethical issues inherent in collecting and analysing student data and pedagogy-based approaches to the interventions. The proposed model describes how differentiated instruction can be provided based on a dynamic learner profile built through an ethical learning analytics process. Differentiated instruction optimises online learning through recommending learning objects tailored towards the learner attributes stored in a learner profile. The proposed model provides a systematic and comprehensive abstraction of a differentiated learning design process informed by learning analytics. The model emerged by synthesising steps of a tried-and-tested web analytics process with educational theory, an ethical learning analytics code of practice, principles of adaptive education systems and a layered abstraction of online learning design.

Keywords: learning analytics, web design, differentiated instruction, learning design

Categories: • **Applied computing ~ Computer assisted instruction**

Email:	**Article history:**
Ronald G. Leppan ronald.leppan@mandela.ac.za (CORRESPONDING)	Received: 01 Mar 2017
Johan F. van Niekerk johan.vanniekerk@noroff.no	Accepted: 17 Sep 2018
Reinhardt A. Botha reinhardta.botha@mandela.ac.za	Available online: 16 Dec 2018

1 INTRODUCTION

In response to the #FeesMustFall protests that halted classroom instruction in 2015 and 2016, many South African institutions adopted a blended learning approach to complete the academic year. The increase in the number of lecturers moving their courses online follows a global trend, resulting in the data generated by online student activity to escalate exponentially (Luna, Castro & Romero, 2017). The measurement, collection, analysis and reporting of data about learners and their contexts is called learning analytics (Siemens, 2013). Data is collected and analysed to optimise learning and the environment in which learning occurs. Exploring the challenges inherent in, and opportunities offered by learning analytics at two mega open distance learning institutions, Prinsloo, Slade and Galpin (2012) argue for a unified and holistic approach to learning analytics that involves all higher education stakeholders. Unlocking the potential of untapped data requires higher education institutions to embed a systematic, student-centric and ethical learning analytics process. One area where the potential of learning analytics can be harnessed is the provision of tailored instruction in response to a dynamic learner profile. Teachers who adapt pedagogy towards their learners' needs, do so out of a belief that a strategy that benefits one group of learners may potentially frustrate another (Brusilovsky, Wade & Conlan, 2007). Employing diverse teaching strategies, whether matched or mismatched to learner needs, could potentially keep learners adequately engaged or suitably challenged (Manning, Stanford & Reeves, 2010). There are several levels of tailored instruction that shares the same goal of modifying pedagogy but differ in how the profile is built and how the learning design is modified.

Differentiated Instruction is a teaching approach that tailors pedagogy towards the diverse needs of individuals or groups sharing similar characteristics (Tomlinson et al., 2003). In online learning, differentiated instruction can be achieved through proactively modifying and sequencing learning objects along preset pathways towards the same learning outcome. Learners are grouped according to shared attributes stored in a learner profile and guided to appropriate learning objects.

While not the primary focus of this paper, related terms need disambiguation since the proposed model incorporates some elements of each of the following levels of tailored instruction:

- Adaptive learning also tailors content and provides individualised pathways. However, unlike differentiated instruction, the profile is built and pedagogy adjusted in real-time through adaptation rules that conditionally include, hide or annotate learning objects (De Bra et al., 2003).
- Personalised learning, like adaptive learning, provides real-time profile building and adaptation but achieves a higher level of personalisation through incorporating initial diagnostic tests and providing learners direct control over their learning environment (Halim, Ali & Yahaya, 2011).
- Individualised learning is a teaching approach that allows learners to dictate their own pace and often set their learning agenda (Kop & Fournier, 2010), unlike the previous three levels of tailored instruction that generally works towards the same learning outcomes.

In the face-to-face classroom, a lecturer can tacitly identify a student's personal needs and adapt accordingly. It is accepted that students are more engaged with the learning material if the learning environment is matched to their attributes (Manning et al., 2010). However, attempting to cater for individual characteristics poses a challenge in face-to-face instruction, especially in large classes. This challenge gets manageable in the online learning environment. Still, lecturers do not directly interact with individual students in an online learning environment, so they need data to make a judgment call regarding student needs. The abundant data provided by Learning Management Systems provide an opportunity to create a learner profile of relevant learner attributes (Luna et al., 2017).

This paper proposes a process that can be used by lecturers who wish to capitalise on students' data generated through their online learning activities. Towards this aim, certain concepts need to be unpacked through a focused literature review (Section 2). Section 3 discusses the research approach followed to design the model. Section 4 synthesises all these related concepts into a comprehensive, systematic and data-driven model for the provision of differentiated instruction based on a dynamic learner profile. Section 5 provides recommendations for practice and future research.

2 BACKGROUND AND LITERATURE

Section 2.1 examines several learning analytics models to identify the typical steps of the learning analytics process. In response to deficiencies identified in existing process models, Section 2.2 describes an ethical learning analytics code of practice and Section 2.3 describes a layered approach to differentiated learning design as an example of a pedagogy-based approach to learning analytics interventions. Section 2.4 introduces the steps of a web analytics process model as an alternative to drive learning analytics interventions.

2.1 Existing learning analytics process models

The Learning Analytics research community uses Educational Data Mining techniques to understand and improve learning processes and learning environments (Siemens, 2013). Educational Data Mining is concerned with developing methods to explore complex data from educational contexts

(Romero & Ventura, 2010). The vision of Learning Analytics researchers is modest incremental interventions to complex educational problems (Merceron, Blikstein & Siemens, 2015). Several cyclical models have been proposed to abstract the steps in a typical learning analytics process.

In Chatti, Dyckhoff, Schroeder and Thüs (2012), the process is described as three steps: data collection and pre-processing, analytics and action, and post-processing (Figure 1). Data is gathered and aggregated from various educational platforms. This data is transformed into input for analysis using pre-processing techniques from the field of data mining. Learning analytics techniques are used to gain insight into strategies employed by learners navigating through online courses. The discovered knowledge about learners is used as a basis to inform suitable interventions and make informed recommendations. The final post-processing step is used to improve the analytics process.

In Clow (2012), learning analytics is described as a cycle that starts with learners participating in formal or informal online learning activities (Figure 2). Through their actions, learners generate large amounts of data that gets logged on online learning platforms. Raw data is processed into knowledge (metrics) about learning processes that can inform appropriate interventions.

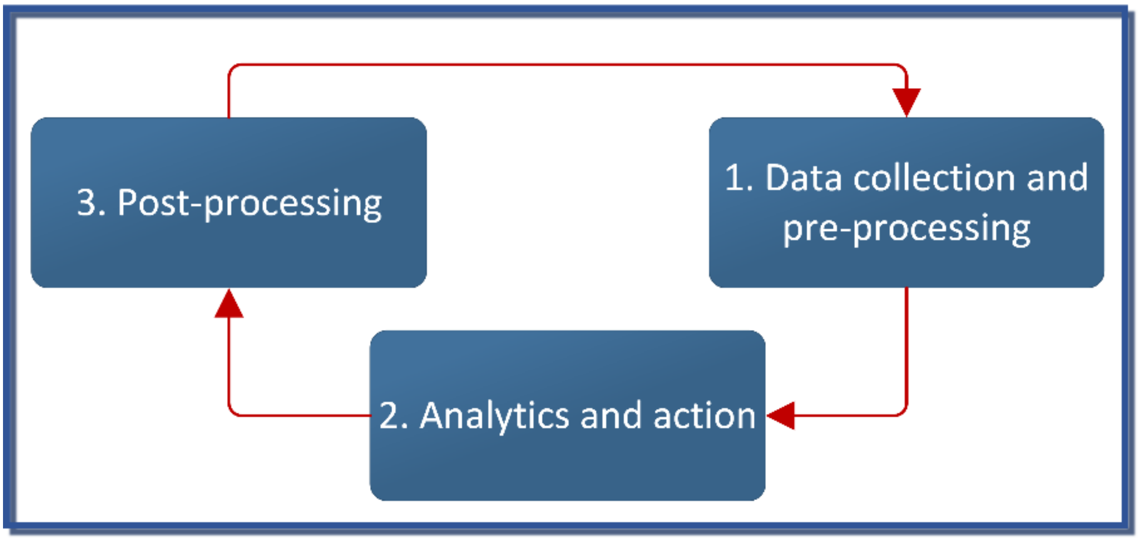


Figure 1: Learning analytics process (adapted from Chatti, Dyckhoff, Schroeder and Thüs (2012))

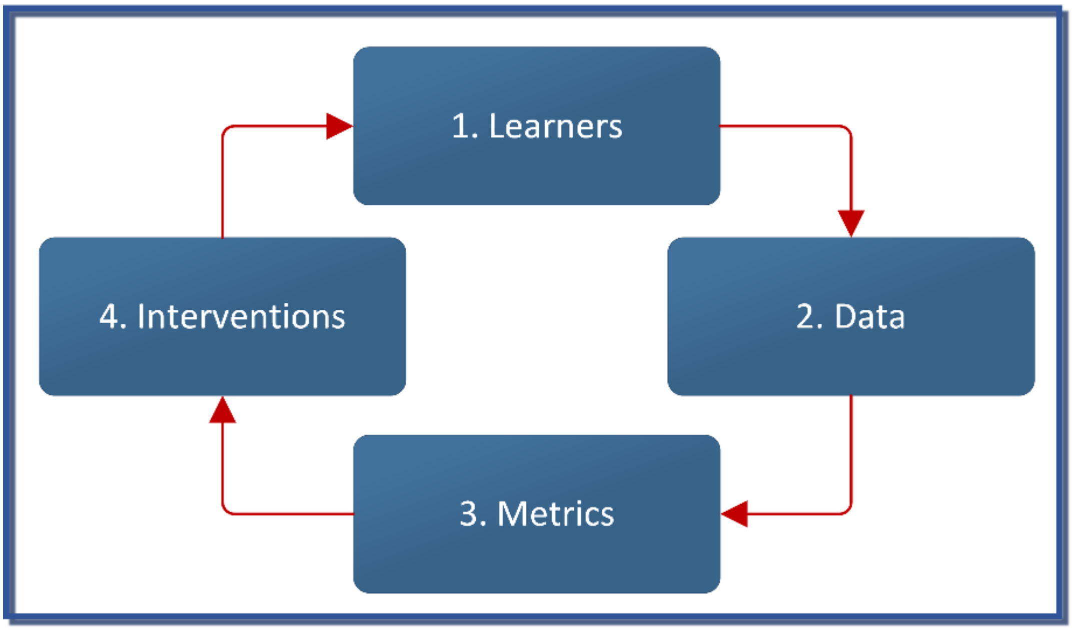


Figure 2: Learning analytics cycle (adapted from Clow (2012))

In (Hundhausen, Olivares & Carter, 2017) a learning analytics process model is used to design an Integrated Development Environment (IDE) capable of collecting data on learning strategies while programming and intervening where necessary. The process describes four steps (Figure 3): collecting data from the IDE, analysing the data to discover programming behaviours, designing the intervention and establishing an automated response to scaffold learners while learning how to code.

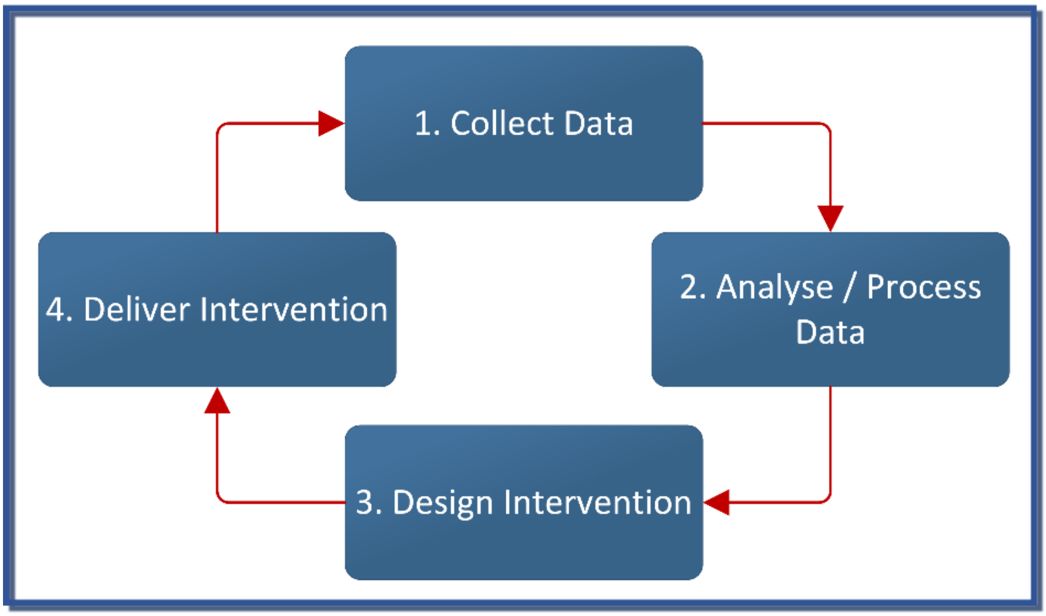


Figure 3: Process model for IDE-based learning analytics (adapted from Hundhausen, Olivares and Carter (2017))

The cyclic model of four stages in (Verbert, Duval, Klerkx, Govaerts & Santos, 2013) focus on the provision of a dashboard for learners to gain insight into their learning strategies (Figure 4). At the first stage, a dashboard will present data visually to the learner who can interrogate the data for self-reflection. After gaining a deeper understanding of their learning processes, the learners can decide whether it is in their best interest to act upon this new insight.

Learning analytics processes can also be used to turn raw data stored in Learning Management Systems into actionable information that can be used to enhance learning (Romero & Ventura, 2013; Romero, Ventura & García, 2008). Usage data of learners completing courses presented on a learning management system is stored in a database (Figure 5). This data needs to undergo a pre-processing phase to transform it into a format suitable for analysis. Data mining algorithms are used on the pre-processed data to create a learner model. Knowledge represented in the learner model can be interpreted and used to make improvements to the learning environment.

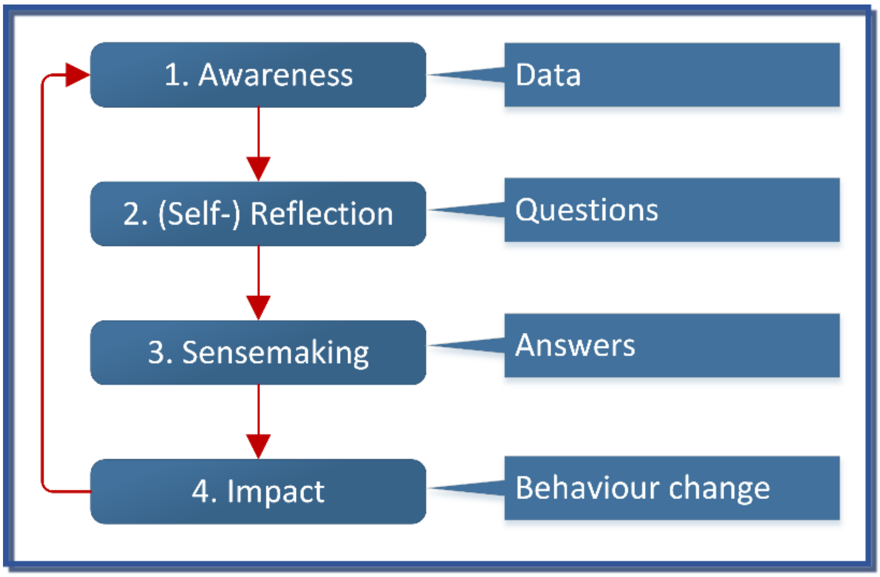


Figure 4: Learning analytics process model (adapted from Verbert, Duval, Klerkx, Govaerts and Santos (2013))

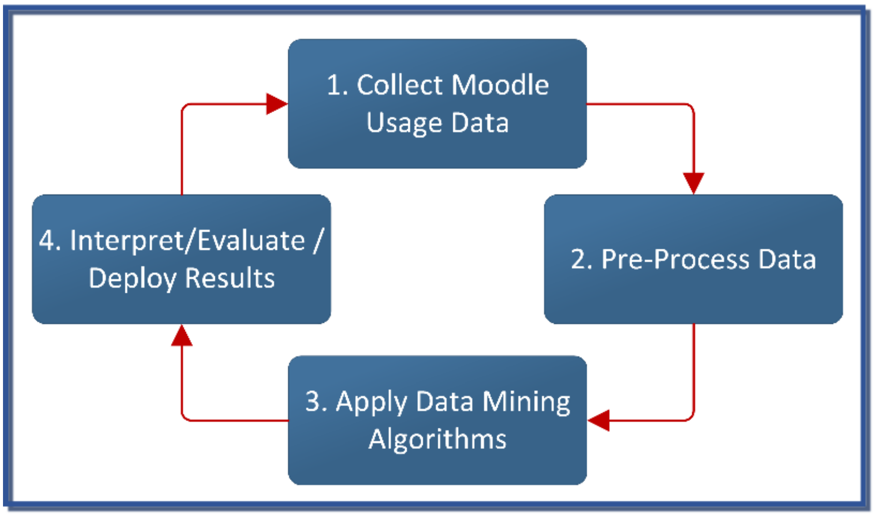


Figure 5: LA process model applied to LMS data (adapted from Romero, Ventura and García (2008))

The consensus from the learning analytics process models described above is that all of them are represented as a cyclic process that includes a data collection phase, a data analysis phase and a phase where action is taken based on the results of the data analysis. What is not explicitly mentioned in these models are:

1. An initial goal setting phase linked explicitly to educational theory
2. The form of the pedagogical intervention that can be taken based on analysis of the results
3. An explicit reflection on an ethical learning analytics process

The first two shortcomings are echoed by Tsai and Gasevic (2017) who also identified a lack of a pedagogy-based approach to learning analytics interventions. The third deficiency in the above list concurs with a concern raised by Viberg, Hatakka, Bälter and Mavroudi (2018), who found only 18% out of 252 papers published from 2012 to 2018 on learning analytics in higher education, reflected on the issue of ethics. Section 2.2 is aimed at addressing the first two concerns, i.e., the lack of an explicitly named pedagogical goal to initiate and conclude a learning analytics initiative, while Section 2.3 describes the relevant issues towards addressing the third concern (ethics).

2.2 Differentiated online learning design

At the core of this study is a belief that online Learning Design should be informed by behavioural patterns exhibited by learners as they navigate through the course material.

These behavioural patterns reveal a learner's cognitive processes (Sabine Graf & Kinshuk, 2008) and affective states (Desmarais & Baker, 2011) that influence the learning process. The learning design in an adaptive learning system that adapts to a learner profile is abstracted in a layered model (Atif, 2010). This type of layered abstraction makes it easier to define differentiation goals during learning design (Figure 6).

At the base is the domain layer that represents content knowledge as an ontology of relevant concepts and semantic relationships between these concepts. The domain model can be represented as a conceptual graph with nodes representing concepts and edges representing relationships between concepts (Melia & Pahl, 2009). Domain experts are responsible for preparing and structuring learning outcomes and related content. The domain layer should be pedagogically neutral.

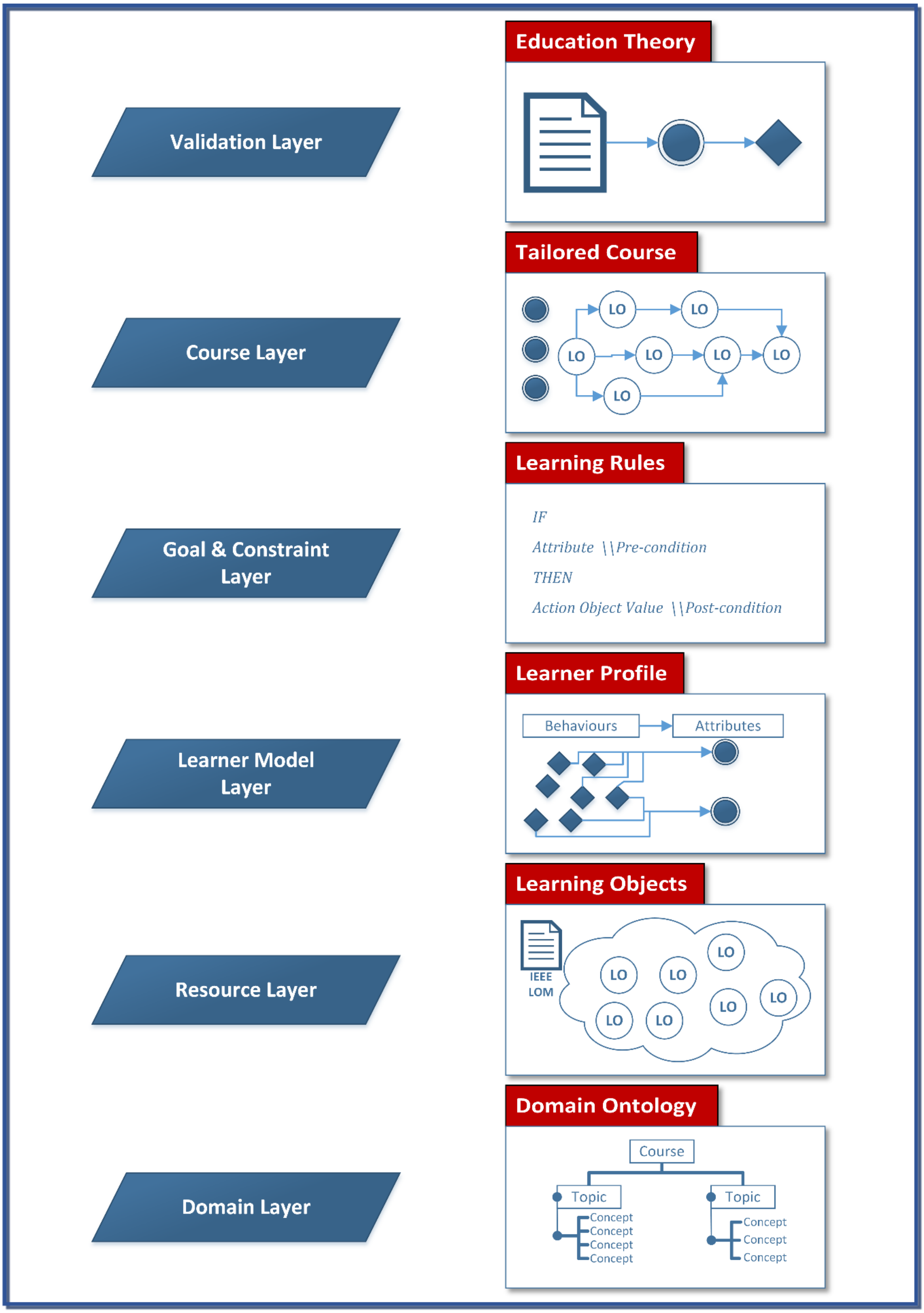


Figure 6: Layered online learning design model for differentiated instruction (adapted from Atif (2010))

The next layer, goal and constraint layer, overlays required competencies and instructional and pedagogical constraints by applying prerequisites and postconditions in the form of learning rules to the domain ontology. Instructional constraints lead to the sequencing of concepts based on whether knowledge of one concept is needed before the learner can move on to another concept. Pedagogical constraints must be defined on learners that are grouped based on prior knowledge and learning goals. Individual learner preferences described in learning style theories are also defined in the goal and constraint layer (Melia & Pahl, 2009).

The learner model layer represents the learner profile. The learner profile can be built explicitly by asking relevant questions to the learner, or implicitly through inferring relevant characteristics by analysing their behaviours (Graf, Kinshuk & Liu, 2008). The learner model can capture their knowledge progression from before, during and after instruction. The learner model can also record learner goals, needs and preferences. The learner model is built while learners work through the course material. In order to optimise online learning design, instructional designers need to build and maintain a dynamic learner profile that is used as a basis for learning interventions. Example categories of learning analytics-based interventions include (Baker & Yacef, 2009):

1. Predictive modelling to model something that cannot be directly observed
2. Structure discovery to find patterns in data that are not obvious
3. Relationship mining to discover or confirm meaningful connections between variables that affect learning
4. Distillation and preparation of data into meaningful information that teachers and learners can use to make informed decisions

Within the above taxonomy, one can identify several techniques within each category and call upon an extensive collection of algorithms to convert raw data into meaningful information. For example, clustering is a common technique categorised as structure discovery. Algorithm choices used for clustering include K-Means, Mean-Shift, or DBSCAN, among others. With so many choices of algorithms to analyse data, the learning analytics process needs to cater for the eventuality that different algorithms may be required if new hypotheses arose after the initial analysis.

The choice of analysis technique is just the technical aspect of building a learner profile. The process of building and maintaining a learner profile must also be conducted within ethical constraints. Some of the fundamental principles of an ethical code of practice for learning analytics is informed consent, transparency and trust (Section 2.3). Informed consent, transparency and trust will be achieved if the reason for the learning analytics initiative is clearly defined from the outset and the learner is made aware of these goals.

The resource layer focus on identifying, repurposing or constructing learning objects that represent the learning content. These learning objects are tagged with metadata based on a standard specification such as IEEE LOM (Atif, 2010). The resource model is, therefore, the layer where the basis is set for instructional design tailored towards the characteristics defined in the learner model. The focus of the adaptation is on the content and presentation of the learning object.

In the course layer, learning objects are sequenced based on the characteristics defined in the learner model. The learner's knowledge, goals, needs and preferences will ultimately dictate how the learner will traverse through the coursework as represented by the domain and goal and constraint models (Melia & Pahl, 2009).

The validation layer is used to examine the instruction design before course delivery (Melia & Pahl, 2009). Validation ensures that learning objects are logically constructed and appropriately sequenced. Validation criteria linked to educational theories must be applied to each layer of the learning design model. An explicit goal setting phase based on pedagogy is, therefore, necessary at the start of any learning analytics initiative.

2.3 Ethical learning analytics code of practice

Since the publishing of the Belmont Report (NCPHS, 1979), higher education institutions have established review committees to ensure research involving human subjects are carried out ethically (Willis, Slade & Prinsloo, 2016). The principles of ethical research upheld by these review committees include respect for persons, beneficence and justice.

Respect for persons is shown when the individual is given adequate information, and they can make informed judgements based on this information. Special care needs to be taken to protect individuals with diminished capacity, from harm. Informed consent by autonomous individuals or their legally authorised guardians should be sought for any ICT related research (Bailey, Dittrich, Kenneally & Maughan, 2012b). The principle of beneficence compels researchers to minimise risks associated with their research and maximise the potential benefits. Invasion of privacy is one of the major ethical dilemmas associated with learning analytics (Griffiths et al., 2016; Steiner, KickmeierRust & Albert, 2016). For any intervention based on learner data, the potential benefits must be weighed against the privacy concerns of the learners. The issue of privacy as it relates to learning analytics is further explored in Section 2.3.1. To ensure the principle of justice, all human subjects should have an equal chance to be selected as participants and receive equal benefits. The issue of equity as it relates to learning analytics is further explored in Section 2.3.2. 2.3.1 Privacy

To eliminate resistance to learning analytics interventions, custodians of data have an ethical and legal obligation to protect the privacy of learners (Hoel & Chen, 2016). Learners' privacy concerns, though, should not prohibit these data-driven initiatives. Admittedly, Slade and Prinsloo (2013) argues that it will be irresponsible to ignore the potential benefits of learning analytics to gain insight into complex learning processes. The issue of data privacy is, therefore, something that deserves careful consideration to ensure acceptance of learning analytics.

In the information age, data protection has become a critical issue related to informational privacy (Griffiths et al., 2016). This sentiment is echoed by Steiner et al. (2016) in the development of LEA's BOX, a learning analytics toolbox that addresses privacy concerns associated with data-driven learner interventions.

The LEA'S BOX privacy and data protection framework proposes eight principles that act as best practice guidelines for learning analytics research:

- Consent: Resistance to provide informed consent can be overcome when learners are provided with relevant information presented unambiguously (Drachsler & Greller, 2012). Necessary information includes, but is not limited to, assurance that their data will be protected, a description of the type of data collected and the purpose for analysing the data.

- Data protection: Learners need reassurance that their data will be protected from abuse. Strategies implemented, such as anonymisation of data and the use of the latest encryption standards, and privacy policies should be communicated to learners.

- Purpose and data ownership: The reason for collecting and analysing data should be published. Data ownership and access rights should be clearly defined and displayed throughout the entire learning analytics process.

- Transparency and trust: Transparency in learning analytics fosters trust in the process and inspires informed consent. An Open Learner Model as presented in (Bull & Kay, 2010) has the potential to build the trust necessary to acquire informed consent.

- Access and control: While transparency of Open Learner Models affords learners an opportunity to view their data and the inferences made from this data, they should also be allowed an opportunity to modify the data where feasible.

- Accountability and assessment: Stakeholders initiating learning analytics endeavours should have clearly defined roles and accountabilities throughout the process. Assigning accountabilities is done to ensure data sources and the analysis techniques are appropriate for the goal.

- Data quality: Data collected about the learner must be timely, precise, appropriate and consistent with the goal. While data quality alone will not guarantee accurate conclusions, poor data quality may undoubtedly contribute to incorrect inferences. All stakeholders have a responsibility to ensure the quality of the raw data collected and inferences made on the data.

- Data management and security: Policies for data management and security must be established at managerial and technical levels.

To minimise risks and maximise the benefits to be gained from learning analytics, these eight data privacy guidelines should underpin all data-driven initiatives. Adhering to these guidelines will support the principle of beneficence proposed in the Belmont Report (NCPHS, 1979).

2.3.2 Equity

To uphold the principle of justice, learning analytics must be applied fairly and equitably (Bailey, Dittrich, Kenneally & Maughan, 2012a). Unless there is a compelling reason, no learner or group of learners should be included (or excluded) from participating in data-driven interventions above others. Furthermore, if there are conflicts of interest between the educator and learner, these must be ethically managed. The actions taken as a result of the data analysis should be applied consistently to all participants (Roberts, Howell, Seaman & Gibson, 2016). To this end, special care needs to be taken to ensure models developed through learning analytics are validated. Any potential for bias must be accounted for in the development of the learner profiles. For example, if facial recognition data is analysed, data from male and female learners must be used to create the model. Models rarely have 100% accuracy, so automated interventions must be dealt with in a sensible way (Roberts et al., 2016). One possible solution to avoid mislabelling a learner through inaccurate models is the use of an open learner model as proposed in (Bull & Kay, 2010). Open learner models allow learners to identify potential misinterpretations made in the analysis process.

The ethical, privacy and equity restrictions placed on learning analytics should not deter educators from using learner data towards optimising the learning environment (Slade & Prinsloo, 2013). Instead, the learning analytics process should be accompanied by a carefully crafted code of practice to ensure buy-in from all stakeholders involved in the process.

2.4 Web analytics process model

The requirement of informed consent and beneficence imposed by an ethical learning analytics process, calls for "goal setting" to be an explicit step in any comprehensive learning analytics process model. "Learning optimisation" as a generic goal of learning analytics initiatives, is too vague and inadequate for building trust in a learner whose informed consent is required. The learning analytics process models described in Section 2.1 is mostly silent on the need for an explicit goal setting phase. With an overemphasis on data collection as the start of the process, we run the risk of taking a haphazard approach to learning analytics initiatives.

Furthermore, the extensive choice of analysis techniques and associated educational data mining algorithms that can be used to extract meaningful information from learner data, needs to be acknowledged in a comprehensive learning analytics model. New hypotheses may have emerged after the initial analysis, and while there was no change in the initial goal, these hypotheses need to be tested before action is taken based on the results of the analysis. The existing learning analytics processes (Section 2.1) also fail to acknowledge this intermediate step.

In online learning, and in particular if learning is delivered through Learning Management Systems, learners receive instruction in a web-based environment. These Learning Management Systems would typically record all learner interactions, thereby providing data that could potentially help us understand learners, their cognitive strategies or affective states. A systematic, comprehensive and student-centric learning analytics process is needed to avoid a hit-or-miss approach to harnessing this untapped learner data.

Learning Analytics and Web Analytics share the same generic goal of using data collection and analysis to understand users' online behaviours in order to optimise the websites with which they interact. It may, therefore, be worthwhile to examine web analytics models used in e-commerce.

Waisberg (2015) proposed a process of six steps that commercial website designers can use to optimise e-commerce websites under their control. An examination of this model reveals not only similar steps prescribed in the learning analytics process models described in Section 2.1, but also steps to overcome some of the limitations identified in existing models.

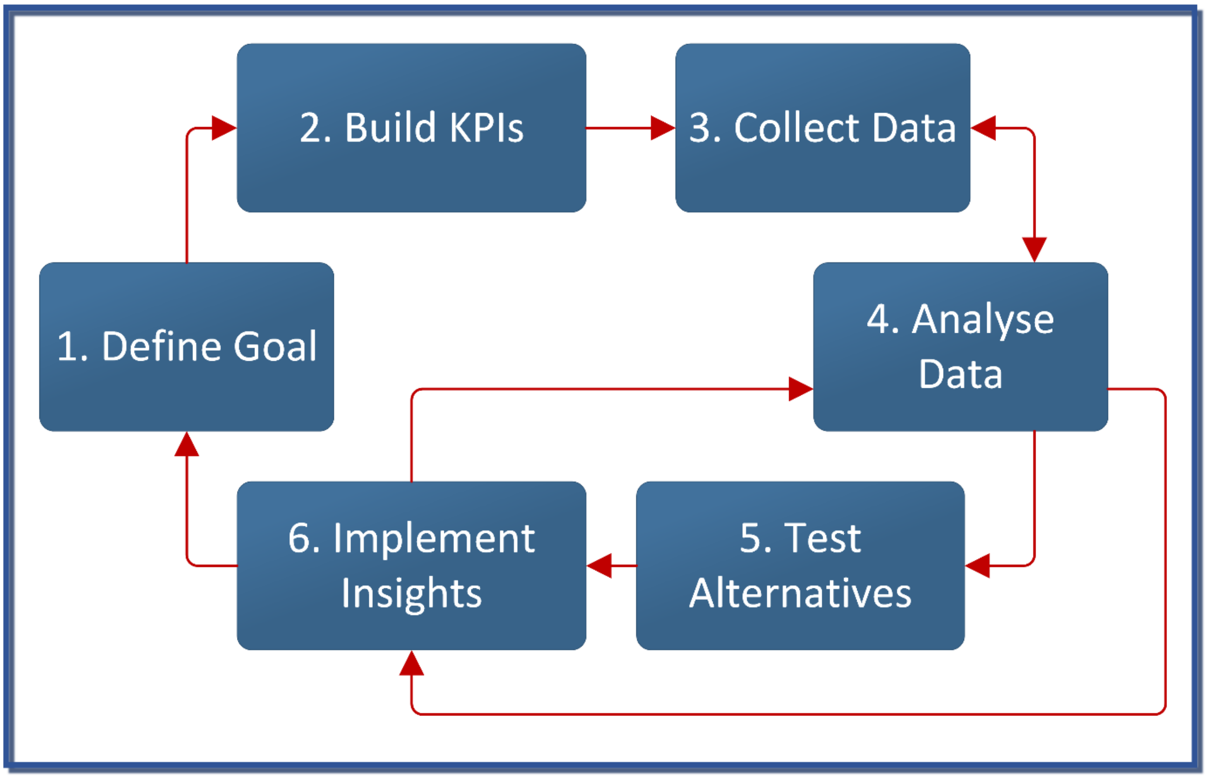


Figure 7: Web analytics process model (adapted from Waisberg (2015))

As can be seen from Figure 7, the web analytics process model is also represented as a cycle, and the following steps are congruent with the learning analytics process models described in Section 2.1:

- Step 3: Collect data,
- Step 4: Analyse data,
- Step 6: Implement insights

Having previously established the need for an explicit goal setting phase and an ability to evaluate alternative hypotheses post analysis, the web analytics model makes these steps explicit through the addition of Step 1 (Define Goal) and Step 5 (Test Alternatives).

These steps will be described in the context of the provision of a dynamic learner profile for differentiated instruction in Section 4. The next Section discusses the research approach and knowledge contribution of this paper in more detail.

3 RESEARCH APPROACH

Simon (1996) distinguishes research in the natural sciences with research in the "science of the artificial". The focus of research in natural science is on describing and explaining how objects in nature or society behave and interact, while research into human-made objects focus on how they are designed to meet predefined goals.

Building on the ideas of Simon (1996) and design research in other fields, Hevner, March, Park and Ram (2004) developed guidelines for conducting, evaluating and presenting design science research in the Information Systems discipline. Design Science Research produces technological artefacts as relevant solutions to problems identified in a specific context. These artefacts can take the form of a construct, model, method or instantiation. The artefact contributed in this study is the proposed model synthesised in Section 4. The model represents an abstracted process to optimise online learning environments through the provision of differentiated instruction based on a dynamic profile. The framework for a Design Science Research contribution (Gregor & Hevner, 2013) classifies artefacts according to solution and application domain maturity (Figure 8). A routine design exercise, in which known solutions are applied to known problems have no knowledge contribution and is therefore not suitable as a research inquiry. Based on this maturity model, knowledge contributions in design science can be classified as improvement, exaptation or invention. A knowledge contribution is classified as an invention if a new solution is developed for a previously unknown problem. An invention is a highly rare form of knowledge contribution, and examples in literature are scarce (Gregor & Hevner, 2013). A knowledge contribution is classified as an improvement if it is a new solution for a known problem and an exaptation if it adopts solutions from other fields to new problems.

When classifying a contribution on sliding scales from specific to abstract, limited to complete and less mature to more mature, three levels can be identified (Gregor & Hevner, 2013):

- Level 1: Situated implementation of the artefact, e.g. instantiation of a software product or application of a process to develop and evaluate the product.
- Level 2: Emerging design theory in the form of prescriptive knowledge, e.g., constructs, methods, models, design principles and technological rules.
- Level 3: Complete mid-range or grand design theories about embedded phenomena.

The process model, described in Section 4, was derived from an established web analytics process model used in business. As such it is a level 2 contribution of the exaptation type, since the proposed model extends a known solution customarily used in a business context (web analytics), to a problem in online learning design.

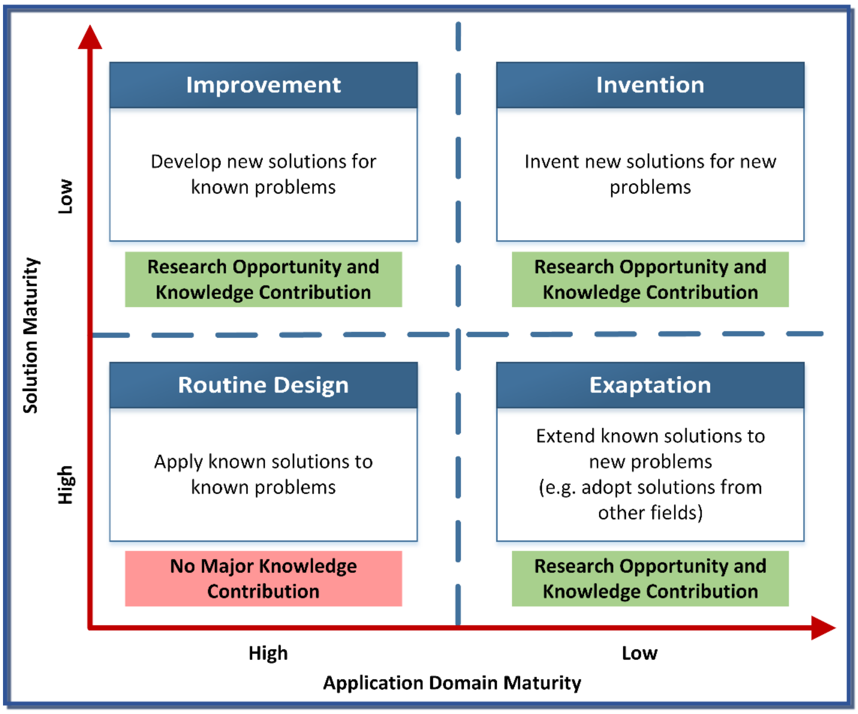


Figure 8: DSR knowledge contribution framework (adapted from Gregor and Hevner (2013))

4 A MODEL FOR DIFFERENTIATED ONLINE INSTRUCTION BASED ON A DYNAMIC LEARNER PROFILE

This Section synthesises a model for Differentiated Instruction based on a dynamic learner profile. The model is derived by integrating:

- the abstracted differentiated learning design layers to initiate and tailor the online course (Section 2.2),
- principles of an ethical learning analytics code of conduct (Section 2.3),
- the steps from the web analytics process to build a learner profile (Section 2.4).

The aim of building a learner profile to provide tailored instruction is shared by researchers who create automated adaptive education systems (AES). Two core phases of a typical AES are the learner modelling phase during which the learner profile is built, and an adaptive learning design phase, during which instruction is personalised based on the unique learner profiles (Brusilovsky & Millán, 2007). The proposed model consists of three phases (Figure 9):

- Preliminary Goal Setting Phase
- Learning Design Phase consisting of two distinct subphases
  - Initial Learning Design before learner modelling
  - Differentiated Learning Design after learner modelling
- Learner Modelling Phase

4.1 Preliminary goal setting phase

The preliminary goal setting phase in Figure 9 stems from steps 1 and 2 of the Web Analytics Process model (Figure 7). One of the abstracted learning design layers is the validation layer that proposes any learning design choice should be backed by recognised educational theories. One such theory, or group of theories, is the identification of learning styles and the tailoring of instruction based on unique learner attributes associated with the learning style model. The validation layer, therefore, maps onto the preliminary goal setting phase, since they both aim to initiate and conclude the learning analytics initiative based on pedagogy. This paper proposes a pragmatic approach of identifying relevant attributes from multiple learning style theories.

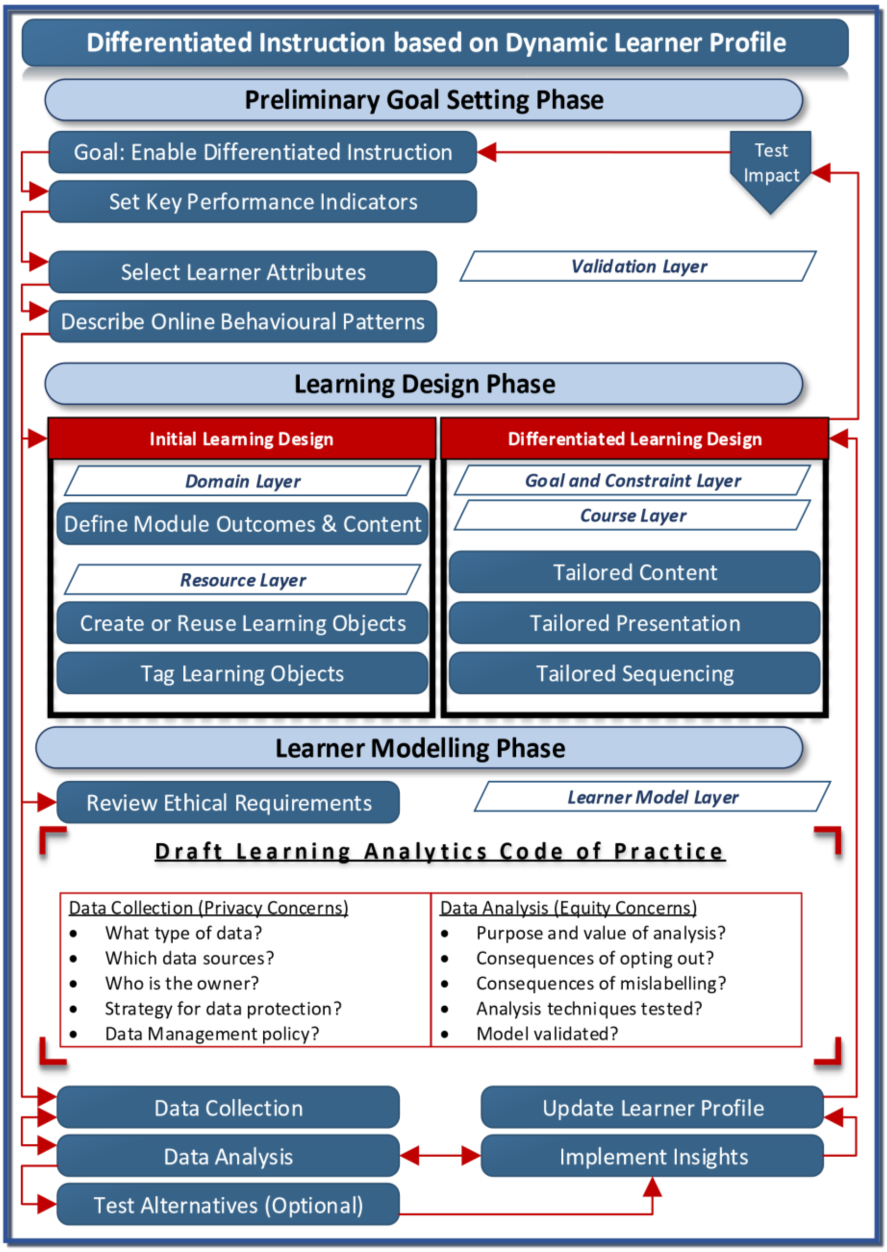


Figure 9: Differentiated instruction based on a dynamic learner profile

4.1.1 Identify goal and set Key Performance Indicators

The general goal proposed in the model (Figure 9) is enabling and optimising differentiated learning design based on learning style theories. Since differentiated instruction share common phases of tailored learning design and learner modelling with learning style based adaptive education systems, two sub-goals are identified:

- Correctly identifying relevant learner attributes from learners' online behaviours
- Appropriately tailoring instruction based on the identified learner attributes

With the goal identified as enabling differentiated instruction and optimising the learning design based on learner profiles, the Key Performance Indicators are linked to the two sub-objectives of the learner modelling phase and the learning design phase. Since the outcome of the learner modelling phase is a learner profile of attributes from selected learning style theories, a generic KPI for a successful learner modelling exercise can be "Attribute X is identified in Learner A". Similarly, the generic KPI to measure a successful learning design phase can be "Learning design is optimised for a learner with Attribute X". A model for evaluating the impact of the learning design is beyond the scope of this paper, but the step is included as part of the goal-setting phase of the proposed model (See the "Test Impact" shape on Figure 9). This impact study is necessary to measure whether the changes made to the learning design had the desired effect on learning. The results of the impact study will feed into further goals for optimising the learning environment and initiate a new cycle.

4.1.2 Select learner attributes and describe online behavioral patterns

One of the biggest challenges when integrating learning styles into adaptive learning systems is the selection of an appropriate learning style theory. Mounting criticism from some dissenting voices (Coffield, Moseley, Hall & Ecclestone, 2004; Cook, 2012; Kirschner, 2016) is pointing to theoretical incoherence, conceptual confusion, lack of scientific basis and seemingly never-ending overlapping characterisation of learner attributes. Further criticism is levelled at the questionnaires used to determine student attributes. This paper proposes that instead of focusing on the model of one particular theorist, we focus instead on the student attributes defined in various learning style theories. By limiting the content of the learner profile to only one learning style theory, we may be missing out on other attributes with an equally significant impact on teaching and learning. The following criteria should be applied to the selection of suitable attributes (Popescu, 2008):

- The learner attributes must influence the learning process in some way, based on an educational theory
- The learner attributes must have implications for differentiated learning design
- It should be possible to infer the learner attributes from metrics that represent online logged behaviours

The focus on collecting and analysing patterns of students' online behaviours to build a learner profile dynamically is precisely in response to the criticism against the use of questionnaires to determine student attributes. When using implicit learner modelling techniques, relevant metrics must be identified that describes the online behaviour of the learner. These metrics must be mapped onto the chosen learner attributes validated by existing educational theory.

4.2 Learning design phase

The learning design phase consists of two subsections, one performed before learner modelling (initial learning design) and one initiated in response to changes in the learner profile (differentiated learning design).

4.2.1 Initial learning design

During the initial learning design phase, the focus is on the domain layer and the resource layer.

For the domain layer, a theoretically sound online instructional design process should be followed to create a significant student-centric learning experience. Module outcomes need to be defined and matched with suitable content. At this stage, the content will be described and later instantiated when the focus shifts to the resource layer. The initial learning design can be represented in the form of a domain ontology.

The input for the resource layer is the learner attributes defined in the goal setting phase. The learning objects that will be presented to the students in the online environment should be linked to the stated module outcomes and be based on the pedagogic needs associated with the selected attributes. These learning objects must be tagged with educational metadata to record the teachers' pedagogic intention. IEEE LOM standards provide a suitable vocabulary for educational metadata (IEEE 1484.12.1, 2002).

4.2.2 Differentiated learning design

While learners navigate the course material, the learner modelling phase will continuously update a learner profile. This profile provides the input into the differentiated learning design subsection. During differentiated learning design the focus is on the goal and constraint layer and the course layer.

Learning rules are created in the goal and constraint layer. Pre- and post-conditions based on the learner profile are overlaid onto the domain ontology. These rules influence the sequencing, content and presentation of learning objects. Learning objects are differentiated based on pre-requisite knowledge, learner goals, cognitive and affective needs contained within the learner profile.

Rules for differentiated learning design based on learner attributes can be represented using IF statements of the format proposed in Popescu (2008):

IF Attribute THEN Action Object Value, where

- Action = Sort | Dim | Hide | Highlight | Trigger | Show
- Object = Metadata tag of Learning Object | UI element
- Value = Value of Metadata tag

The metadata tags of learning objects and their associated values are linked to the fields and values from the educational category of the IEEE LOM standard (IEEE 1484.12.1, 2002). The list of actions suggested above is not exhaustive. The mentioned actions are illustrative of the typical type of techniques used in adaptive education systems to tailor learning objects (Popescu, 2008):

- Sort represents the sequencing of LOs or UI elements
- Dim represents greying out or disabling an LO or UI element such as a button or hyperlink
- Hide represents the removal of an LO or UI element
- Highlight represents a recommendation of a particular LO or UI element
- Trigger represents an action such as the sending of an automated message
- Show represents displaying an LO or UI element such as a table of content or annotation

The learning rules designed in the goal and constraint layer are implemented in the course layer. The learning objects from the resource layer are tailored according to the rules defined in the goal and constraint layer. The learning objects can be differentiated on their sequence (Action: Sort), content (Actions: Dim, Hide, Highlight, Trigger, Show) or the presentation UI. The chosen educational theory will determine the form of the actions to be taken based on the learner attribute. Any tailored learning object must still guide the learners towards the same learning outcomes defined in the domain layer.

As can be seen from the IF statement, the identified attribute will be the trigger to inform the differentiated learning design choices. In the learner modelling phase, the online behaviour of learners will be used to infer relevant attributes to add to the learner profile. This learner modelling phase is described next.

4.3 Learner modelling phase

The steps in the learner modelling phase are based on steps 3—6 of the web analytics process model (Figure 7) and the learning analytics code of ethical practice described in Section 2.3. Also incorporated into the learner modelling phase are activities and techniques associated with learning style based adaptive education systems and educational data mining.

4.3.1 Review ethical requirements

Any learning analytics initiative must be conducted ethically, and practitioners must carefully address privacy (Section 2.3.1) and equity (Section 2.3.2) concerns. To ensure buy-in from learners, their privacy must be guaranteed during data collection, and they must be convinced that the benefits that will accrue from the data analysis outweigh potential risks. A learning analytics code of practice must be drafted and used to acquire informed consent from all participants whose data will be analysed and used for changes to the learning design. This code of practice must incorporate principles of ethical research, i.e., respect for persons, beneficence and justice (NCPHS, 1979).

4.3.2 Data collection

During the data collection step, metrics identified during the goal setting phase must be collected. All potential data sources that may supply these metrics need to be identified. In implicit modelling, these metrics represent learner cognitive and affective behaviours linked to learner attributes associated with educational theories. In explicit modelling, data can be elicited directly from learners responding to questions. During data collection, all privacy measures as drafted in the learning analytics code of practice must be implemented.

4.3.3 Data analysis

Learner attributes as identified during the goal setting phase are inferred during the data analysis step. The goal and the nature of the raw data collected in the previous step will determine the sequence of activities in the data analysis step. It may be possible, for example, to use simple inferential statistics if inferences and predictions are to be made on a small dataset. More complex goals and large datasets may require more advanced educational data mining techniques, such as listed below (Baker & Yacef, 2009):

- Predictive modelling to model something that cannot be directly observed by using readily available features as input (e.g., Classification, Latent Knowledge Estimation, Regression)
- Structure discovery to find patterns in data that are not obvious (e.g. Clustering, Factor Analysis, Social Network Analysis)
- Relationship mining to discover or confirm meaningful connections between variables that affect learning (e.g., Association Rule Mining, Correlation Mining, Sequential Pattern Mining, Causal Data Mining)
- Distillation and preparation of data into meaningful information that teachers and learners can use to make informed decisions (e.g., Data Visualisation, Text Mining)

Large data sets from disjoint sources may require pre-processing to prime data for analysis. Pre-processing can include data cleaning, integration, reduction or transformation. It is beyond the scope of this paper to report on all possible pre-processing techniques, but the following serve as an illustration of the potential strategies commonly applied to data mining:

- Data cleaning is responsible for removing inconsistencies and errors in the data. For example, there may be missing values, noisy, i.e., meaningless or unstructured data, outliers or inconsistent data.
- Data integration is responsible for consolidating data from multiple disjoint data sources. Learners frequently need to consult resources outside of the learning environment or perform offline activities. Alternatively, biometric data need to be integrated with online behavioural metrics in order to measure affect, for example. Metrics may, therefore, come from several sources and need to be combined sensibly.
- Data reduction focuses on deciding which data features to include or exclude for analysis. Data reduction aims to find a smaller dataset that can produce similar analytical results. Data reduction can be performed through several techniques such as:
  - Aggregation—combining two or more attributes
  - Sampling—selecting a subset from the population

- Feature subset reduction—removing redundant or irrelevant features

- Data transformation converts data into a different format. Conventional techniques to transform data include:
  - Normalisation—scaling values into a predetermined range
  - Smoothing—the removal of outliers
  - Aggregation—preparing data into a summarised format
  - Generalisation—substituting data points into hierarchical layers

When data is ready, analysis can proceed through a suitable educational data mining technique. Data pre-processing and analysis is concluded by evaluating the results of the analysis. Evaluation methods will depend on the data mining technique used and are necessary to measure the quality of the learner model that results from the data analysis. 4.3.4 Test alternatives and implement insights

The educational data mining step may reveal unexpected results that need further investigation. The proposed model allows an optional step to generate new hypotheses that may require:

- Exploration of different data sources
- Addition of new attributes/features
- Application of different educational data mining techniques, for example
  - Trying different algorithms
  - Tweaking clusters
  - Using the results of one analysis technique as input into another
- Applying negotiated learner modelling to seek the learners' approval of the conclusions made in the data analysis step
- Making a quick change to the learning design and conducting a small-scale pilot study to measure the effect of the change

Once satisfactory results are achieved the necessary action can be taken ("Implement Insights"). The subsequent action involves a two-part process:

- Updating the learner profile with inferred information
- Initiating the differentiated learning design in response to the changes in the learner profile (Section 4.2.2)

Evaluation of the impact of the differentiated learning design on learner satisfaction, learning effectiveness and efficiency closes the process model loop. This step is represented in the model as an off-page reference since this step is yet to be modelled as part of future work.

5 CONCLUSION AND FURTHER WORK

Existing learning analytics process models suffer from a too narrow focus on the data collection and analysis steps of learning interventions. This myopic view on the technical aspects of learning analytics often results in interventions lacking pedagogical validation and ethical reflection. When the first step of the learning analytics process is data collection, there is likely to be a lack of clarity on the goal of the intervention. An ethical learning analytics code of practice requires participants to be explicitly made aware of the goal of the data collection, analysis and intervention. A learning analytics process model also needs to acknowledge the fact that more questions may arise after the initial analysis is done. There is, therefore, a need for a more comprehensive abstraction of the learning analytics process.

Regarding Design Science Research, the knowledge contribution made in this paper is that of an emerging model that addresses limitations in existing learning analytics models. The proposed solution can be classified as an exaptation of a tried-and-tested model used in e-commerce and applying it to the online learning application domain.

The process model proposed in this paper emerged by incorporating steps of an established web analytics process with educational theory, an ethical learning analytics code practice and a layered abstraction of online learning design. The pedagogical aspects of the model are derived from the concept of differentiated instruction, a teaching approach that prescribes modifying instruction based on the diverse needs of individuals sharing similar attributes. The online learning design is abstracted through several layers that systematically guides instructional designers through the process of designing and developing tailored learning objects to satisfy a range of diverse learner needs. The learner modelling phase prescribes a review of ethical requirements, drafting of an ethical code of practice and implementation of mechanisms to ensure principles of data privacy and equity are upheld throughout data collection, analysis and intervention. The learner modelling phase also provides an optional step to test new hypotheses should they arise after initial analysis.

Many education institutions are adopting Learning Management Systems as the online learning environment. However, Learning Management Systems mostly suit a one-size-fits-all approach to teaching. Future work includes instantiating the learning design phase and learner modelling phase in a Learning Management System to determine whether it is possible to provide differentiated instruction and maintain a dynamic learner profile based on the data logged by the system. The ultimate goal for the proposed model is to enable the discovery of relevant learner cognitive and affective attributes that influence online learning behaviours. While the contribution of this paper is on how learning analytics can inform learning design, a model to measure the impact of the changes to the learning design also remain future work.

References

Atif, Y. (2010). An architectural specification for a system to adapt to learning patterns. *education and Information Technologies*, *16*(3), 259–279. 10.1007/s10639-010-9125-9

Bailey, M., Dittrich, D., Kenneally, E. & Maughan, D. (2012a). *Applying ethical principles to information and communication technology research: A companion to the Department of Homeland Security Menlo report*. US Department of Homeland Security.

Bailey, M., Dittrich, D., Kenneally, E. & Maughan, D. (2012b). The Menlo report. *IEEE Security & Privacy*, *10*(2), 71–75. 10.1109/MSP.2012.52

Baker, R. S. & Yacef, K. (2009). The state of educational data mining in 2009: A review and future visions. *Journal of Educational Data Mining*, *1*(1), 3–16. Last accessed 23 Oct 2018. Retrieved from https://jedm.educationaldatamining.org/index.php/JEDM/article/view/8

Brusilovsky, P. & Millán, E. (2007). User models for adaptive hypermedia and adaptive educational systems. In P. Brusilovsky, A. Kobsa & W. Nejdl (Eds.), *The adaptive web*. 10.1007/978-3-540-72079-9_1

Brusilovsky, P., Wade, V. & Conlan, O. (2007). From learning objects to adaptive content services for e-learning. In C. Pahl (Ed.), *Architecture solutions for e-learning systems* (pp. 243–261). IGI Global.

Bull, S. & Kay, J. (2010). Open learner models. In R. Nkambou, J. Bourdeau & R. Mizoguchi (Eds.), *Studies in computational intelligence* (Vol. 308). 10.1007/978-3-642-14363-2_15

Chatti, M. A., Dyckhoff, A. L., Schroeder, U. & Thüs, H. (2012). A reference model for learning analytics. *International Journal of Technology Enhanced Learning*, *4*(5/6), 318–331. 10.1504/ijtel.2012.051815

Clow, D. (2012). The learning analytics cycle. In *Proceedings of the 2nd international conference on learning analytics and knowledge* (pp. 134–138). 10.1145/2330601.2330636

Coffield, F., Moseley, D., Hall, E. & Ecclestone, K. (2004). *Learning styles and pedagogy in post-16 learning: A systematic and critical review*. Learning and Skills Research Centre London.

Cook, D. A. (2012). Revisiting cognitive and learning styles in computer-assisted instruction. *Academic Medicine*, *87*(6), 778–784. 10.1097/ACM.0b013e3182541286

De Bra, P., Aerts, A., Berden, B., De Lange, B., Rousseau, B., Santic, T., . . . Stash, N. (2003). AHA! The adaptive hypermedia architecture. In *Proceedings of the fourteenth acm conference on hypertext and hypermedia* (pp. 81–84). 10.1145/900065.900068

Desmarais, M. C. & Baker, R. (2011). A review of recent advances in learner and skill modeling in intelligent learning environments. *User Modeling and User Adapted Interaction*, *22*(1-2), 9–38. 10.1007/s11257-011-9106-8

Drachsler, H. & Greller, W. (2012). The pulse of learning analytics understandings and expectations from the stakeholders. In *Proceedings of the 2nd international conference on learning analytics and knowledge* (pp. 120–129). 10.1145/2330601.2330634

Graf, S., Kinshuk, K. & Liu, T.-C. (2008). Identifying learning styles in learning management systems by using indications from students' behaviour. In *2008 eighth ieee international conference on advanced learning technologies* (pp. 482–486). 10.1109/ICALT.2008.84

Graf, S. [Sabine] & Kinshuk, T. (2008). Analysing the behaviour of students in learning management systems with respect to learning styles. In M. Wallace, M. Angelides & P. Mylonas (Eds.), *Advances in semantic media adaptation and personalization*. 10.1007/978-3-540-76361_3

Gregor, S. & Hevner, A. R. (2013). Positioning and presenting design science research for maximum impact. *MIS Quarterly*, *37*(2), 337–355. 10.25300/misq/2013/37.2.01

Griffiths, D., Drachsler, H., Kickmeier-Rust, M., Steiner, C., Hoel, T. & Greller, W. (2016). *Is privacy a show-stopper for learning analytics? A review of current issues and solutions*. 23 Oct 2018. Retrieved from http://www.laceproject.eu/learning-analytics-review/files/2016/04/LACE-review-6%7B%5C_%7Dprivacy-show-stopper.pdf

Halim, N. D. A., Ali, M. B. & Yahaya, N. (2011). Personalized learning environment: Accommodating individual differences in online learning. In *2011 international conference on social science and humanity* (Vol. 5, pp. 398–400). 23 Oct 2018. Retrieved from http://www.ipedr.com/vol5/no2/88-H10220.pdf

Hevner, A., March, S., Park, J. & Ram, S. (2004). Design science in information systems research. *MIS Quarterly*, *28*(1), 75–105. https://doi.org/10.2307/25148625

Hoel, T. & Chen, W. (2016). Privacy-driven design of learning analytics applications: Exploring the design space of solutions for data sharing and interoperability. *Journal of Learning Analytics*, *3*(1), 139–158. 10.18608/jla.2016.31.9

Hundhausen, C., Olivares, D. & Carter, A. (2017). IDE-based learning analytics for computing education: A process model, critical review and research agenda. *ACM Transactions on Computing education*, *17*(3), 1–26. 10.1145/3105759

IEEE 1484.12.1. (2002). IEEE standard for learning object metadata. *Learning Technology Standards Committee of the IEEE*, *1484.12.1*-(July), 1–44. 10.1109/IEEESTD.2002.94128

Kirschner, P. A. (2016). Stop propagating the learning styles myth. *Computers & education*, *106*, 166–171. 10.1016/j.compedu.2016.12.006

Kop, R. & Fournier, H. (2010). New dimensions to self-directed learning in an open networked learning environment. *International Journal of Self-Directed Learning*, *7*(2), 1–18.

Luna, J. M., Castro, C. & Romero, C. (2017). MDM tool: A data mining framework integrated into Moodle. *Computer Applications in Engineering education*, *25*(1), 90–102. 10.1002/cae.21782

Manning, S., Stanford, B. & Reeves, S. (2010). Valuing the advanced learner: Differentiating up. *The Clearing House*, *83*(4), 145–149. 10.1080/00098651003774851

Melia, M. & Pahl, C. (2009). Constraint-based validation of adaptive e-Learning courseware. *IEEE Transactions on Learning Technologies*, *2*(1), 37–49. 10.1109/TLT.2009.7

Merceron, A., Blikstein, P. & Siemens, G. (2015). Learning analytics: From big data to meaningful data. *Journal of Learning Analytics*, *2*(3), 4–8. 10.18608/jla.2015.23.2

NCPHS. (1979). The Belmont report: Ethical principles and guidelines for the protection of human subjects of research. *OPPR Reports*, *44*(76), 1–5. 10.1002/9780471462422.eoct093. arXiv: DHEWPublicationNo.(OS)78-0012

Popescu, E. (2008). *Dynamic adaptive hypermedia systems for e-learning* (Doctoral Dissertation, University of Technology of Compiègne).

Prinsloo, P., Slade, S. & Galpin, F. (2012). Learning analytics: Challenges, paradoxes and opportunities for mega open distance learning institutions. *Proceedings of the 2nd International Conference on Learning Analytics and Knowledge - LAK '12*, 130–133. 10.1145/2330601.2330635

Roberts, L. D., Howell, J. A., Seaman, K. & Gibson, D. C. (2016). Student attitudes toward learning analytics in higher education: "The fitbit version of the learning world". *Frontiers in Psychology*, *7*(DEC), 1–11. 10.3389/fpsyg.2016.01959

Romero, C. & Ventura, S. (2010). educational data mining: A review of the state of the art. *IEEE Transactions on Systems, Man, and Cybernetics, Part C (Applications and Reviews)*, *40*(6), 601–618. 10.1109/TSMCC.2010.2053532

Romero, C. & Ventura, S. (2013). Data mining in education. *Wiley Interdisciplinary Reviews: Data Mining and Knowledge Discovery*, *3*(1), 12–27. 10.1002/widm.1075

Romero, C., Ventura, S. & García, E. (2008). Data mining in course management systems: Moodle case study and tutorial. *Computers & education*, *51*, 368–384.

Siemens, G. (2013). Learning analytics: The emergence of a discipline. *American Behavioral Scientist*, *57*(10), 1380–1400. 10.1177/0002764213498851

Simon, H. A. (1996). The sciences of the artificial. *Computers & Mathematics with Applications*, *33*(5). 10.1016/S0898-1221(97)82941-0. arXiv: 0262191938

Slade, S. & Prinsloo, P. (2013). Learning analytics: Ethical issues and dilemmas. *American Behavioral Scientist*, *57*(10), 1510–1529. 10.1177/0002764213479366

Steiner, C. M., Kickmeier-Rust, M. D. & Albert, D. (2016). LEA in private: A privacy and data protection framework for a learning analytics toolbox. *Journal of Learning Analytics*, *3*(1), 66–90. 10.18608/jla.2016.31.5

Tomlinson, C. A., Brighton, C., Hertberg, H., Callahan, C. M., Moon, T. R., Brimijoin, K., Reynolds, T. (2003). Differentiating instruction in response to student readiness, interest, and learning profile in academically diverse classrooms: A review of literature. *Journal for the education of the Gifted*, *27*(2-3), 119–145. 10.1177/016235320302700203

Tsai, Y.-S. & Gasevic, D. (2017). Learning analytics in higher education — challenges and policies. *Proceedings of the Seventh International Learning Analytics & Knowledge Conference on - LAK '17*, (October), 233–242. 10.1145/3027385.3027400

Verbert, K., Duval, E., Klerkx, J., Govaerts, S. & Santos, J. L. (2013). Learning analytics dashboard applications. *American Behavioral Scientist*, *57*(10), 1500–1509. 10.1177/0002764213479363

Viberg, O., Hatakka, M., Bälter, O. & Mavroudi, A. (2018). The current landscape of learning analytics in higher education. *Computers in Human Behavior*, *89*(December), 98–110. 10.1016/j.chb.2018.07.027

Waisberg, D. (2015). *Google analytics integrations*. John Wiley & Sons.

Willis, J. E., Slade, S. & Prinsloo, P. (2016). Ethical oversight of student data in learning analytics: A typology derived from a cross-continental, cross-institutional perspective. *educational Technology Research and Development*, *64*(5), 881–901. 10.1007/s11423-016-9463-4

APPENDIX B: LEARNER MODELLING FROM MOODLE LOGS

A pilot study of Moodle data was conducted to evaluate the expected practicality and utility of the learner modelling phase in a Learning Management System (Section 5.5). The intention of the pilot study was to investigate the Moodle logs generated by learners in an actual course. The investigation focused on the data provided in a standard Moodle log and no intervention on the learners in the course was intended. To ensure privacy during data collection, all identifying fields (User full name and affected user) were removed by the lecturer. The system generated user id was extracted from the Description field and used to represent unique learners. In addition, the main investigator has no direct teacher relationship with the learners whose Moodle data was explored.

In particular, a standard Moodle log was pre-processed, and clustering applied in order to determine whether there are discernible differences in the way learners access course material. To enable successful analysis, the initial learning design had to be done in a way prescribed by the proposed model.

1. Initial Learning Design

Since the objective of the pilot study was merely to investigate the data produced in a standard Moodle log file, the actual metadata tags used at this stage are arbitrary and interchangeable. The intention was to analyse the format of the log file and determine what pre-processing is required to prime data ready for analysis.

Initially, the tag functionality was used but it was found not to be visible in the report. When the Tag block is activated in Moodle, a word cloud shows the tags used in the course (Figure B-1). Moodle tags can be used to search for relevant resources. However, this functionality was not used by the learners, most likely since this functionality was not demonstrated to the class. Moodle allows the teacher to add a description of resources and activities. This Description field was also used to describe the pedagogic intent behind the resource in more detail. This enabled learners to select the resources which they would like to learn from, especially if resources describe similar content, but in different ways. The description field is also not in the report, but its usefulness is based on the depth of the information that can be shared with the learners.

Figure B-1. Moodle Tag Block

Due to the tags and description fields not being visible in the log file, the decision was made to introduce a special naming convention to tag resources with relevant metadata (<educationalTag>:<Title>). This naming convention simultaneously guide learners to appropriate resources and activities and ensure the necessary information is available in the standard log in the field "Event Context". However, the naming convention has the limitation that not all relevant tags can be used. Most resources and activities require multiple tags to adequately describe the content and pedagogic intent behind the Learning Object.

Most of the resources used in the pilot study were links to web pages or external URLs. These resources used in this pilot study all represented content in the form of "Text" or "Video" media elements. In total there were 195 text resources and 124 video resources. Initial access was during a 2-hour proctored session in a computer laboratory, during which learners could use a chat facility to communicate with the lecturer and primary investigator to ask questions. The material was available to access for 20 days prior to a semester test. After the 20 days, the log file was downloaded for the entire period. The raw data was pre-processed to prepare data for analysis and to investigate the practicality of Moodle with regards to the analysis phase of the proposed model. The next Section describes how the log file was manipulated to get it ready for clustering. The goal of clustering applied to this data was to determine if there are discernible differences based on preferences for a specific media type. Excerpts of the log are shown next to demonstrate how the log file changed after each pre-processing step.

2. Raw Data from Moodle Log

Total number of records in the initial log = 21355. This includes all interaction data generated by teacher and site administrator roles.

Time	User full name	Affected user	Event context	Component	Event name	Description	Origin	IP address
18/05/18, 08:35			URL: Text: What is an API?	URL	Course module viewed	The user with id '60090' viewed the 'url' activity with course module id '171695'.	web	10.102.130.10
18/05/18, 08:35			Module: CNW2110/SSI2000: Operating Systems II	System	Course viewed	The user with id '57231' viewed the section number '17' of the course with id '516'.	web	10.102.138.236
18/05/18, 08:35			Assignment: Assignment: Discuss how operating s	Assignment	The status of the submission has been viewed.	The user with id '60089' has viewed the submission status page for the assignment with course module id '171717'.	web	10.102.131.204
18/05/18, 08:35			Module: CNW2110/SSI2000: Operating Systems II	System	Course viewed	The user with id '60089' viewed the section number '17' of the course with id '516'.	web	10.102.131.204
18/05/18, 08:36			URL: Text: What is an API?	URL	Course module viewed	The user with id '47556' viewed the 'url' activity with course module id '171695'.	web	185.94.190.17
18/05/18, 08:36			Module: CNW2110/SSI2000: Operating Systems II	System	Course viewed	The user with id '47556' viewed the section number '17' of the course with id '516'.	web	185.94.190.17
18/05/18, 08:36			Module: CNW2110/SSI2000: Operating Systems II	System	User profile viewed	The user with id '47617' viewed the profile for the user with id '47617' in the course with id '516'.	web	10.102.129.110
18/05/18, 08:36			URL: Video: What is an API?	URL	Course module viewed	The user with id '39721' viewed the 'url' activity with course module id '171693'.	web	10.102.137.36
18/05/18, 08:36			URL: Video: What is an API?	URL	Course module viewed	The user with id '39721' viewed the 'url' activity with course module id '171693'.	web	10.102.137.36
18/05/18, 08:36			URL: Video: What is an API?	URL	Course module viewed	The user with id '52307' viewed the 'url' activity with course module id '171693'.	web	10.102.139.142
18/05/18, 08:36			Module: CNW2110/SSI2000: Operating Systems II	System	Course viewed	The user with id '52307' viewed the section number '17' of the course with id '516'.	web	10.102.139.142
18/05/18, 08:36			Module: CNW2110/SSI2000: Operating Systems II	System	Course viewed	The user with id '51525' viewed the course with id '516'.	web	10.102.137.117
18/05/18, 08:36			Module: CNW2110/SSI2000: Operating Systems II	System	Course viewed	The user with id '51525' viewed the section number '17' of the course with id '516'.	web	10.102.137.117
18/05/18, 08:36			URL: Video: What is an API?	URL	Course module viewed	The user with id '51525' viewed the 'url' activity with course module id '171693'.	web	10.102.137.117
18/05/18, 08:36			URL: Video: What is an API?	URL	Course module viewed	The user with id '51525' viewed the 'url' activity with course module id '171693'.	web	10.102.137.117
18/05/18, 08:36			Module: CNW2110/SSI2000: Operating Systems II	System	Course viewed	The user with id '51525' viewed the section number '17' of the course with id '516'.	web	10.102.137.117
18/05/18, 08:36			URL: Text: What is an API?	URL	Course module viewed	The user with id '48356' viewed the 'url' activity with course module id '171695'.	web	10.102.129.196
18/05/18, 08:36			URL: Text: What is an API?	URL	Course module viewed	The user with id '48356' viewed the 'url' activity with course module id '171695'.	web	10.102.129.196
18/05/18, 08:36			Chat: Session 1 Chat	System	Course activity completion updated	The user with id '48391' updated the completion state for the course module with id '171720' for the user with id '48391'.	web	10.102.130.14
18/05/18, 08:36			Chat: Session 1 Chat	System	Course activity completion updated	The user with id '48391' updated the completion state for the course module with id '171720' for the user with id '48391'.	web	10.102.130.14
:					:			:
7/06/18, 00:26			Quiz: Revision Quiz 1: Elements of an Operating Sy	Quiz	Quiz attempt summary viewed	The user with id '47491' has viewed the summary for the attempt with id '1817824' belonging to the user with id '47491' f	web	10.122.171.177
7/06/18, 00:26			Quiz: Revision Quiz 1: Elements of an Operating Sy	Quiz	Quiz attempt viewed	The user with id '47491' has viewed the attempt with id '1817824' belonging to the user with id '47491' for the quiz with c	web	10.122.171.177
7/06/18, 00:26			Quiz: Revision Quiz 1: Elements of an Operating Sy	Quiz	Quiz attempt viewed	The user with id '47491' has viewed the attempt with id '1817824' belonging to the user with id '47491' for the quiz with c	web	10.122.171.177
7/06/18, 00:26			Quiz: Revision Quiz 1: Elements of an Operating Sy	Quiz	Quiz attempt viewed	The user with id '47491' has viewed the attempt with id '1817824' belonging to the user with id '47491' for the quiz with c	web	10.122.171.177
7/06/18, 00:26			Quiz: Revision Quiz 1: Elements of an Operating Sy	Quiz	Quiz attempt viewed	The user with id '47491' has viewed the attempt with id '1817824' belonging to the user with id '47491' for the quiz with c	web	10.122.171.177
7/06/18, 00:26			Quiz: Revision Quiz 1: Elements of an Operating Sy	Quiz	Quiz attempt viewed	The user with id '47491' has viewed the attempt with id '1817824' belonging to the user with id '47491' for the quiz with c	web	10.122.171.177
7/06/18, 00:26			Quiz: Revision Quiz 1: Elements of an Operating Sy	Quiz	Quiz attempt viewed	The user with id '47491' has viewed the attempt with id '1817824' belonging to the user with id '47491' for the quiz with c	web	10.122.171.177
7/06/18, 00:25			Quiz: Revision Quiz 1: Elements of an Operating Sy	Quiz	Quiz attempt viewed	The user with id '47491' has viewed the attempt with id '1817824' belonging to the user with id '47491' for the quiz with c	web	10.122.171.177
7/06/18, 00:25			Quiz: Revision Quiz 1: Elements of an Operating Sy	Quiz	Quiz attempt started	The user with id '47491' has started the attempt with id '1817824' for the quiz with course module id '172101'.	web	10.122.171.177
7/06/18, 00:25			Quiz: Revision Quiz 1: Elements of an Operating Sy	Quiz	Course module viewed	The user with id '47491' viewed the 'quiz' activity with course module id '172101'.	web	10.122.171.177
7/06/18, 00:23			Quiz: Revision Quiz 1: Elements of an Operating Sy	Quiz	Quiz attempt reviewed	The user with id '47491' has had their attempt with id '1817801' reviewed by the user with id '47491' for the quiz with cou	web	10.122.171.177
7/06/18, 00:23			Quiz: Revision Quiz 1: Elements of an Operating Sy	Quiz	Quiz attempt submitted	The user with id '47491' has submitted the attempt with id '1817801' for the quiz with course module id '172101'.	web	10.122.171.177
7/06/18, 00:23			Module: CNW2110/SSI2000: Operating Systems II	System	User graded	The user with id '-1' updated the grade with id '6140601' for the user with id '47491' for the grade item with id '53762'.	web	10.122.171.177
7/06/18, 00:23			Module: CNW2110/SSI2000: Operating Systems II	System	User graded	The user with id '47491' updated the grade with id '6756659' for the user with id '47491' for the grade item with id '60329'	web	10.122.171.177
7/06/18, 00:23			Quiz: Revision Quiz 1: Elements of an Operating Sy	Quiz	Quiz attempt summary viewed	The user with id '47491' has viewed the summary for the attempt with id '1817801' belonging to the user with id '47491' f	web	10.122.171.177
7/06/18, 00:23			Quiz: Revision Quiz 1: Elements of an Operating Sy	Quiz	Quiz attempt viewed	The user with id '47491' has viewed the attempt with id '1817801' belonging to the user with id '47491' for the quiz with c	web	10.122.171.177
7/06/18, 00:17			Quiz: Revision Quiz 1: Elements of an Operating Sy	Quiz	Quiz attempt viewed	The user with id '47491' has viewed the attempt with id '1817801' belonging to the user with id '47491' for the quiz with c	web	10.122.171.177
7/06/18, 00:16			Quiz: Revision Quiz 1: Elements of an Operating Sy	Quiz	Quiz attempt viewed	The user with id '47491' has viewed the attempt with id '1817801' belonging to the user with id '47491' for the quiz with c	web	10.122.171.177
7/06/18, 00:15			Quiz: Revision Quiz 1: Elements of an Operating Sy	Quiz	Quiz attempt viewed	The user with id '47491' has viewed the attempt with id '1817801' belonging to the user with id '47491' for the quiz with c	web	10.122.171.177
7/06/18, 00:15			Quiz: Revision Quiz 1: Elements of an Operating Sy	Quiz	Quiz attempt viewed	The user with id '47491' has viewed the attempt with id '1817801' belonging to the user with id '47491' for the quiz with c	web	10.122.171.177
7/06/18, 00:15			Quiz: Revision Quiz 1: Elements of an Operating Sy	Quiz	Quiz attempt viewed	The user with id '47491' has viewed the attempt with id '1817801' belonging to the user with id '47491' for the quiz with c	web	10.122.171.177
7/06/18, 00:13			Quiz: Revision Quiz 1: Elements of an Operating Sy	Quiz	Quiz attempt viewed	The user with id '47491' has viewed the attempt with id '1817801' belonging to the user with id '47491' for the quiz with c	web	10.122.171.177
7/06/18, 00:13			Quiz: Revision Quiz 1: Elements of an Operating Sy	Quiz	Quiz attempt started	The user with id '47491' has started the attempt with id '1817801' for the quiz with course module id '172101'.	web	10.122.171.177
7/06/18, 00:06			Quiz: Revision Quiz 1: Elements of an Operating Sy	Quiz	Course module viewed	The user with id '58114' viewed the 'quiz' activity with course module id '172101'.	web	196.23.22.138
7/06/18, 00:06			Module: CNW2110/SSI2000: Operating Systems II	System	Course viewed	The user with id '58114' viewed the section number '17' of the course with id '516'.	web	196.23.22.138

Figure B-2. Unprocessed data from a Standard Moodle Log

3. Pre-processed Data

a.) In **Sampling**, records are removed that is not relevant to the goal of the analysis. In this pilot study, sampling was performed on user roles and event types.

(1) The first step was to remove records based on user role. Only learner data is relevant for the clustering exercise. Total number of learner records remaining = 15205. This excludes actions performed by teacher and site administrator roles.

(2) The next step was to remove records based on events, i.e. the actions performed on the Resources and Activities (Appendix C). This required studying learner actions on the resources recorded in the log file. Sorting by component and removing all duplicates resulted in Appendix C.

The following represented components and events relevant to distinguishing between text and video.

- Component: Page; Event name: Course module viewed
- Component: URL; Event name: Course module viewed

After removing all but the Page and URL components, 2855 records remained.

b.) **Feature Subset Reduction** removes unnecessary fields. For this investigation, the following fields were removed from the Moodle log file: User full name, Affected user, Component, Event name, Origin, IP address. The following fields remain.

Time	Event context	Description
18/05/18, 08:55	URL: Video: What is an API?	The user with id '47617' viewed the 'url' activity with course module id '171693'.
18/05/18, 08:57	URL: Video: BIOS and UEFI	The user with id '47617' viewed the 'url' activity with course module id '171700'.
18/05/18, 08:57	URL: Text: Windows kernel	The user with id '47617' viewed the 'url' activity with course module id '171713'.
18/05/18, 08:58	URL: Video: What is an API?	The user with id '47617' viewed the 'url' activity with course module id '171693'.
18/05/18, 08:58	URL: Video: What is an API?	The user with id '47617' viewed the 'url' activity with course module id '171693'.
18/05/18, 09:05	URL: Video: What is an API?	The user with id '47617' viewed the 'url' activity with course module id '171694'.
18/05/18, 09:08	URL: Text: What is an API?	The user with id '47617' viewed the 'url' activity with course module id '171695'.

Figure B-3. Fields after Feature Subset Reduction

c.) **Feature Construction and Aggregation** develops new fields that are necessary for analysis.

(1) The following new fields were required and extracted from existing fields:

- User ID – extracted from the Description field
- Day – extracted from the existing Time field
- Time – extracted from the existing Time field
- MediaElement (Text, Video, Animation, Graphic) – extracted from the Event context field named according to the convention <MediaElement>: <Resource Title>
- Title – extracted from the Event context field named according to the convention <MediaElement>: <Resource Title>

After the media elements were extracted from the Event Context field, further records were removed that did not relate to any of the four media elements, leaving 1713 records from 88 learners to use for analysis. The following fields remained:

UserID	Day	Time	MediaElement	Title
57967	18/05/18	08:00	Video	What is an API?
56082	18/05/18	08:01	Text	How to use an API
56082	18/05/18	08:01	Text	What is an API?
56082	18/05/18	08:01	Video	What is an API?
57281	18/05/18	08:10	Text	Windows kernel
48356	18/05/18	08:14	Video	What is an API?
57629	18/05/18	08:14	Video	What is an API?
47560	18/05/18	08:15	Animation	What is an API?
47560	18/05/18	08:15	Text	What is an API?
47560	18/05/18	08:15	Text	What is an API?
48356	18/05/18	08:15	Animation	What is an API?
50652	18/05/18	08:15	Animation	What is an API?

Figure B-4. Fields after Feature Construction and Aggregation

(2) The following information was aggregated and summarized per UserID and MediaElement type:

- NoAccess – Total number of times a learner viewed a resource of MediaElement “Text” or a resource of MediaElement “Video”, “Graphic”, or “Animation”.
- %Access – The ratio of the number of times a resource of a particular MediaElement type was viewed divided by the total number of times all resources were viewed.

d.) **Discretization** was used to categorize preferred sequence of access to a particular MediaElement type, based on the following rule:

- Examine the time of first access for all occurrences where the same concept was explained using resources of different MediaElement types.
  - For each learner (UserID), assign a 1 in the column of the MediaElement accessed first and 0 for the others.
  - Determine the frequency of 1’s over all concepts and aggregate to a 1 for the max frequency.

e.) **Smoothing** was used to remove outliers. The only set of outliers removed were occurrences where the log showed the same timestamp for the same resource. These duplicate entries may have been due to the learner clicking on a link, exiting and clicking back on the same link within a minute. Or the duplicates may have been due to automatic updating the completion status of the resource. This would typically result in an entry that shows the resource was viewed and a system report stating the completion status. Learner errors or completion status updates are not relevant for the analysis phase, so these outliers in the log files had to be removed. After removing these duplicates, 1576 records remained for analysis.

The amount of data passed to the analysis phase was minimized through the pre-processing steps described above. The resultant .CSV file contained only the relevant records (instances) and fields (attributes) that was passed through a clustering algorithm.

4. Clustering

The intention behind this part of the pilot study was to determine whether the WEKA Data Mining tool (Waikato Environment for Knowledge Analysis) can produce a list of learners and the groups to which they should be assigned. WEKA was used to perform clustering on the pre-processed log file, using the SimpleKMeans algorithm.

Since the aim was to create two groups of learners, each exhibiting behaviour associated with two dichotomous attributes, then number of clusters were set at two (k=2). In this instance the attributes of interest were a preference for text-based resources or a preference for video\graphical\animation-based resources. The dataset loaded into WEKA contained the UserID, and fields representing the number of times text (NoText) and video/graphic/animation (NoVisual) resources were accessed as well as the predominant sequence of access (TextVis and VisText). WEKA calls these fields "Attributes", not to be confused with the final groupings which are also called learner attributes in this thesis. The WEKA attributes correspond to the behavioural metrics used to distinguish between different learner characteristics.

Before running the SimpleKMeans algorithm, the number of clusters must be set to two. In addition, the UserID field must be ignored, since they do not represent an attribute that should be used by the clustering algorithm. The algorithm produced two clusters of 27 (Cluster 0) and 61 (Cluster 1) learners respectively, using the following (random) starting points:

- Cluster 0: 9, 11, 0, 1
- Cluster 1: 0, 12, 0, 1

The AddCluster filter can be applied on the results to see which learners belong to which group. This filter creates a new attribute (cluster) that shows the information in Figure B-5. The problem with this output, though, is that the generic names "cluster 1" and "cluster 2" are assigned to the two groups. We need a way to map these two clusters onto "RW" or "Visual" attributes.

No.	1: ï»¿UserID Numeric	2: NoText Numeric	3: NoVisual Numeric	4: TextVis Numeric	5: VisText Numeric	6: **cluster** Nominal
1	16179.0	18.0	26.0	1.0	0.0	cluster2
2	31265.0	2.0	11.0	0.0	1.0	cluster1
3	38925.0	1.0	4.0	0.0	1.0	cluster1
4	39266.0	0.0	2.0	0.0	1.0	cluster1
5	39267.0	1.0	0.0	1.0	0.0	cluster2
6	39721.0	4.0	5.0	0.0	1.0	cluster1
7	40288.0	7.0	2.0	1.0	0.0	cluster2
8	44751.0	2.0	1.0	0.0	1.0	cluster1
9	46778.0	8.0	15.0	0.0	1.0	cluster1
10	47269.0	13.0	6.0	1.0	0.0	cluster2

Figure B-5. Learners divided into two groups (cluster 1, cluster 2)

One way to do this, is to add a class attribute to the dataset. By examining the data in the pre-processed log, one may be able to make an educated guess as to each learner's attribute. For this dataset, the attribute (called Class) was added and populated with "RW" for all those learners mostly exhibiting a preference for text and "Visual" for those mostly exhibiting a preference for video, graphics and animations. For this new dataset, the UserID and the Class attributes were ignored. Applying the AddCluster filter shows the following output (Figure B-6):

No.	1: ï»¿UserID Numeric	2: NoText Numeric	3: NoVisual Numeric	4: TextVis Numeric	5: VisText Numeric	6: Class Nominal	7: **cluster** Nominal
1	16179.0	18.0	26.0	1.0	0.0	RW	cluster2
2	31265.0	2.0	11.0	0.0	1.0	Visual	cluster1
3	38925.0	1.0	4.0	0.0	1.0	Visual	cluster1
4	39266.0	0.0	2.0	0.0	1.0	Visual	cluster1
5	39267.0	1.0	0.0	1.0	0.0	RW	cluster2
6	39721.0	4.0	5.0	0.0	1.0	Visual	cluster1
7	40288.0	7.0	2.0	1.0	0.0	RW	cluster2
8	44751.0	2.0	1.0	0.0	1.0	RW	cluster1
9	46778.0	8.0	15.0	0.0	1.0	Visual	cluster1
10	47269.0	13.0	6.0	1.0	0.0	RW	cluster2

Figure B-6. Learners divided into named clusters

Note from Figure B-6, the majority of cluster2 maps onto the “RW” learner attributes and the majority of cluster1 maps onto the “Visual” attribute. This shows that the educated guess was mostly correct. In cases where learners have been assigned to a different cluster than was otherwise predicted may simply mean the learner exhibits a more balanced behaviour, in this case not showing a distinct difference in selecting text-based resources over graphical-based resources. The results in Figure B-5 and Figure B-6 can be saved as a .CSV file. This file, with a bit of tweaking can be used to quickly create Cohorts in Moodle as described in Section 5.4.3.2.

This pilot study has, therefore, shown that it is technically possible to pre-process data from a standard Moodle log and that the WEKA data mining tool can produce the desired groups in the .CSV file that is necessary to create Cohorts in Moodle. However, performing manual pre-processing on the data is extremely time-consuming. Consequently, a plugin is needed to accommodate the pre-processing needs of the model proposed in this thesis.

APPENDIX C: ACTIONS PERFORMED ON MOODLE RESOURCES AND ACTIVITIES

Table C-1. Actions performed on Moodle Resources and Activities

Component	Event Name	Description
Assignment	A submission has been submitted.	The user with id '16179' has submitted the submission with id '959484' for the assignment with course module id '171717'.
Assignment	Course module instance list viewed	The user with id '39721' viewed the instance list for the module 'assign' in the course with id '516'.
Assignment	Grading table viewed	The user with id '5678' viewed the grading table for the assignment with course module id '171717'.
Assignment	Submission confirmation form viewed.	The user with id '16179' viewed the submission confirmation form for the assignment with course module id '171717'.
Assignment	Submission form viewed.	The user with id '16179' viewed their submission for the assignment with course module id '171717'.
Assignment	The status of the submission has been viewed.	The user with id '12979' has viewed the submission status page for the assignment with course module id '171717'.
Assignment	The user has accepted the statement of the submission.	The user with id '16179' has accepted the statement of the submission with id '959484' for the assignment with course module id '171717'.
Book	Chapter created	The user with id '5678' created the chapter with id '1916' for the book with course module id '173258'.
Book	Chapter deleted	The user with id '5678' deleted the chapter with id '1916' for the book with course module id '173258'.
Book	Chapter updated	The user with id '5678' updated the chapter with id '1916' for the book with course module id '173258'.
Book	Chapter viewed	The user with id '12979' viewed the chapter with id '1923' for the book with course module id '173258'.
Book	Course module viewed	The user with id '12979' viewed the 'book' activity with course module id '173258'.
Chat	Course module instance list viewed	The user with id '39721' viewed the instance list for the module 'assign' in the course with id '516'.
Chat	Course module viewed	The user with id '12979' viewed the 'book' activity with course module id '173258'.
Chat	Message sent	The user with id '47269' has sent a message in the chat with course module id '171720'.
Chat	Sessions viewed	The user with id '16179' has viewed the sessions of the chat with course module id '171720'.
Choice	Choice answer added	The user with id '43359' has added the option with id '4193' for the user with id '43359' from the choice activity with course module id '124880'.
Choice	Choice answer deleted	The user with id '48920' has deleted the option with id '4193' for the user with id '48920' from the choice activity with course module id '124880'.

Table C-1. Actions performed on Moodle Resources and Activities (Continued)

Component	Event Name	Description
Choice	Choice report downloaded	The user with id '12979' has downloaded the report in the 'xls' format for the choice activity with course module id '124880'
Choice	Choice report viewed	The user with id '12979' has viewed the report for the choice activity with course module id '124880'
Choice	Course module instance list viewed	The user with id '39721' viewed the instance list for the module 'assign' in the course with id '516'.
Choice	Course module viewed	The user with id '12979' viewed the 'book' activity with course module id '173258'.
Database	Course module instance list viewed	The user with id '39721' viewed the instance list for the module 'assign' in the course with id '516'.
Database	Course module viewed	The user with id '12979' viewed the 'book' activity with course module id '173258'.
MS Excel spreadsheet	XLS grade exported	The user with id '12979' exported grades using the xls export in the gradebook.
Feedback	Course module instance list viewed	The user with id '39721' viewed the instance list for the module 'assign' in the course with id '516'.
Feedback	Course module viewed	The user with id '12979' viewed the 'book' activity with course module id '173258'.
Feedback	Response submitted	The user with id '31265' submitted response for 'feedback' activity with course module id '173050'.
File	Course module viewed	The user with id '12979' viewed the 'book' activity with course module id '173258'.
File submissions	A file has been uploaded.	The user with id '16179' has uploaded a file to the submission with id '959484' in the assignment activity with course module id '171717'.
File submissions	Submission created.	The user with id '16179' created a file submission and uploaded '1' file/s in the assignment with course module id '171717'.
File submissions	Submission updated.	The user with id '16179' updated a file submission and uploaded '1' file/s in the assignment with course module id '171717'.
Forum	Course module instance list viewed	The user with id '39721' viewed the instance list for the module 'assign' in the course with id '516'.
Forum	Course module viewed	The user with id '12979' viewed the 'book' activity with course module id '173258'.
Forum	Discussion created	The user with id '5678' has created the discussion with id '27794' in the forum with course module id '173022'.
Forum	Discussion subscription created	The user with id '5678' subscribed the user with id '5678' to the discussion with id '27794' in the forum with the course module id '173022'.

Table C-1. Actions performed on Moodle Resources and Activities (Continued)

Component	Event Name	Description
Forum	Discussion viewed	The user with id '47558' has viewed the discussion with id '26618' in the forum with course module id '68816'.
Forum	Some content has been posted.	The user with id '5678' has posted content in the forum post with id '66987' in the discussion '27794' located in the forum with course module id '173022'.
Forum	Subscription created	The user with id '3251' subscribed the user with id '3251' to the forum with course module id '171724'.
Glossary	Comment created	The user with id '3251' added the comment with id '6290' to the glossary activity with course module id '172074'.
Glossary	Comment deleted	The user with id '3251' deleted the comment with id '6290' from the glossary activity with course module id '172074'.
Glossary	Course module instance list viewed	The user with id '39721' viewed the instance list for the module 'assign' in the course with id '516'.
Glossary	Course module viewed	The user with id '12979' viewed the 'book' activity with course module id '173258'.
Glossary	Entry has been approved	The user with id '5678' has approved the glossary entry with id '2130' for the glossary activity with course module id '172074'.
Glossary	Entry has been created	The user with id '3251' has created the glossary entry with id '2129' for the glossary activity with course module id '172086'.
Glossary	Entry has been deleted	The user with id '5678' has deleted the glossary entry with id '2112' in the glossary activity with course module id '172074'.
Glossary	Entry has been disapproved	The user with id '5678' has disapproved the glossary entry with id '2130' for the glossary activity with course module id '172074'.
Glossary	Entry has been updated	The user with id '5678' has updated the glossary entry with id '2112' in the glossary activity with course module id '172074'.
Grader report	Grader report viewed	The user with id '12979' viewed the grader report in the gradebook.
Live logs	Live log report viewed	The user with id '5678' viewed the live log report for the course with id '516'.
Logs	Log report viewed	The user with id '5678' viewed the log report for the course with id '516'.
Media collection	Collection deleted	The user with id '5678' has deleted the collection with id '57' in the Media collection with course module id '173257'.
Media collection	Course module viewed	The user with id '12979' viewed the 'book' activity with course module id '173258'.

Table C-1. Actions performed on Moodle Resources and Activities (Continued)

Component	Event Name	Description
Media collection	Gallery created	The user with id '5678' has created the gallery with id '51' in the Media collection with course module id '173257'.
Media collection	Gallery viewed	The user with id '5678' has viewed the gallery with id '51' in the Media collection with course module id '173257'.
Overview report	Grade overview report viewed	The user with id '57070' viewed the overview report in the gradebook.
Page	Course module viewed	The user with id '12979' viewed the 'book' activity with course module id '173258'.
Quiz	Course module instance list viewed	The user with id '39721' viewed the instance list for the module 'assign' in the course with id '516'.
Quiz	Course module viewed	The user with id '12979' viewed the 'book' activity with course module id '173258'.
Quiz	Quiz attempt abandoned	The user with id '48356' has had their attempt with id '1803003' marked as abandoned for the quiz with course module id '172101'.
Quiz	Quiz attempt preview started	The user with id '5678' has had their attempt with id '1800605' previewed by the user with id '5678' for the quiz with course module id '172101'.
Quiz	Quiz attempt reviewed	The user with id '16179' has had their attempt with id '1711389' reviewed by the user with id '16179' for the quiz with course module id '168285'.
Quiz	Quiz attempt started	The user with id '16179' has started the attempt with id '1813692' for the quiz with course module id '172101'.
Quiz	Quiz attempt submitted	The user with id '16179' has submitted the attempt with id '1813692' for the quiz with course module id '172101'.
Quiz	Quiz attempt summary viewed	The user with id '16179' has viewed the summary for the attempt with id '1813692' belonging to the user with id '16179' for the quiz with course module id '172101'.
Quiz	Quiz attempt viewed	The user with id '16179' has viewed the attempt with id '1813692' belonging to the user with id '16179' for the quiz with course module id '172101'.
Quiz	Quiz edit page viewed	The user with id '12979' viewed the edit page for the quiz with course module id '171718'.
Quiz	Quiz report viewed	The user with id '12979' viewed the report 'overview' for the quiz with course module id '168285'.
Recycle bin	Item created	Item created with ID 2036.
Recycle bin	Item deleted	Item with ID 1859 deleted.

Table C-1. Actions performed on Moodle Resources and Activities (Continued)

Component	Event Name	Description
Respondus 4.0 Web Services	Questions published from Respondus	The user with id '12979' published questions from Respondus to the question category with id '41104'.
Single view	Grade single view report viewed.	The user with id '12979' viewed the singleview report in the gradebook.
System	Calendar event created	The user with id '12979' created the event 'Revision Quiz 1: Elements of an Operating System (Quiz closes)' with id '68086'.
System	Calendar event deleted	The user with id '12979' deleted the event 'Revision Quiz 1: Elements of an operating system (Quiz opens)' with id '67852'.
System	Calendar event updated	The user with id '5678' updated the event 'Assignment 2: Find technical specifications of a CPU' with id '68484'.
System	Course activity completion updated	The user with id '12979' updated the completion state for the course module with id '171686' for the user with id '12979'.
System	Course module created	The user with id '12979' created the 'quiz' activity with course module id '172096'.
System	Course module deleted	The user with id '12979' deleted the 'quiz' activity with course module id '171718'.
System	Course module instance list viewed	The user with id '39721' viewed the instance list for the module 'assign' in the course with id '516'.
System	Course module updated	The user with id '12979' updated the 'quiz' activity with course module id '172101'.
System	Course Section updated	The user with id '5678' updated Section number '17' for the course with id '516'
System	Course user report viewed	The user with id '57618' viewed the user report for the course with id '516' for user with id '57618'.
System	Course viewed	The user with id '12979' viewed the course with id '516'.
System	Grade deleted	The user with id '12979' deleted the grade with id '6715624' for the user with id '12979' for the grade item with id '60242'.
System	Recent activity viewed	The user with id '16179' viewed the recent activity report in the course with id '516'.
System	Role assigned	The user with id '3251' assigned the role with id '5' to the user with id '3251'.
System	Role unassigned	The user with id '5678' unassigned the role with id '5' from the user with id '3251'.
System	Tag added to an item	The user with id '5678' added the tag with id '4351' to the item type 'course_modules' with id '172083'.

Table C-1. Actions performed on Moodle Resources and Activities (Continued)

Component	Event Name	Description
System	Tag removed from an item	The user with id '5678' removed the tag with id '4389' from the item type 'course_modules' with id '172309'.
System	User enrolled in course	The user with id '3251' enrolled the user with id '3251' using the enrolment method 'self' in the course with id '516'.
System	User graded	The user with id '-1' updated the grade with id '5882146' for the user with id '47556' for the grade item with id '53762'.
System	User list viewed	The user with id '39721' viewed the list of users in the course with id '516'.
System	User profile viewed	The user with id '16179' viewed the profile for the user with id '16179' in the course with id '516'.
Turnitin Assignment	List Turnitin assignments	User viewed the Turnitin assignment list for course 516
Turnitin Assignment	View Turnitin assignment	User viewed assignment 'Submit Chapter 1 Case Study'
URL	Course module viewed	The user with id '12979' viewed the 'book' activity with course module id '173258'.
User report	Grade user report viewed	The user with id '16179' viewed the user report in the gradebook.
Wiki	Comments viewed	The user with id '5678' viewed the comments for the page with id '589' for the wiki with course module id '173205'.
Wiki	Course module viewed	The user with id '12979' viewed the 'book' activity with course module id '173258'.
Wiki	Wiki diff viewed	The user with id '5678' viewed the diff for the page with id '589' for the wiki with course module id '173205'.
Wiki	Wiki history viewed	The user with id '5678' viewed the history for the page with id '589' for the wiki with course module id '173205'.
Wiki	Wiki page created	The user with id '5678' created the page with id '589' for the wiki with course module id '173205'.
Wiki	Wiki page locks deleted	The user with id '5678' deleted locks for the page with id '589' for the wiki with course module id '173205'.
Wiki	Wiki page map viewed	The user with id '5678' viewed the wiki map for the page with id '589' for the wiki with course module id '173205'.
Wiki	Wiki page updated	The user with id '5678' updated the page with id '589' for the wiki with course module id '173205'.
Wiki	Wiki page viewed	The user with id '5678' viewed the page with id '589' for the wiki with course module id '173205'.

APPENDIX D: A MODEL FOR USING LEARNERS' ONLINE BEHAVIOUR TO INFORM DIFFERENTIATED INSTRUCTIONAL DESIGN IN A LEARNING MANAGEMENT SYSTEM