MULTIMODAL SOFTWARE FOR AFFECTIVE EDUCATION: USER INTERACTION DESIGN AND EVALUATION

by

Reza GhasemAghaei

A thesis submitted to
the Faculty of Graduate Studies and Research
in partial fulfillment of
the requirements for the degree of

DOCTOR OF PHILOSOPHY

School of Computer Science

at

CARLETON UNIVERSITY

Ottawa, Ontario
May, 2017

Contents

List of Tables

List of Figures

Abstract

Multimodal systems allow user interaction through various methods, or modalities, such as gesture, gaze, speech or head and body movements. There have been many frameworks proposed for the design of multimodal systems, but they focus principally on issues of technical architecture. Some applications for multimodal systems, including education, however, involve issues relating to affective (emotional) user experience.

The thesis presents our experience with creating and applying a framework for multimodal software design and evaluation, with a focus on affect (i.e. emotion) in education. Our goal is to support the system designer, and indirectly also teachers and students. We outline our conceptual basis, adaptation of interaction design and evaluation techniques, and our experience over several case studies. In particular, we describe how our techniques were refined after each case study, and how the evaluation techniques led us to revisit our design techniques. We also gained new insight about the role of the teacher, leading us to develop a new affect visualization dashboard for teachers.

Acknowledgements

First and foremost, I would like to thank my supervisors:
To Robert Biddle: words cannot express how grateful I am for all your vision, wisdom, support, kindness and mentorship.
To Ali Arya: thank you for your valuable guidance, insight, perception, intelligence and advice during my thesis.
Please accept my heartfelt gratitude. I am always grateful for your support. This work would have not been possible without your supervision.

Thanks to the members of my committee, Andrew Patrick, Anthony Whitehead, Emil Petriu, Gabriel Wainer, Nicholas Graham, and the defence chair Adrian Chan for their expertise, and for offering different perspectives, all in which have helped shape this work.

Thanks to my colleagues in the HotSoft lab and Interactive Media Group for their friendship, and who helped with studies, and offered valuable feedback and insight throughout the process.

Thanks to the participants who took part in our user studies to test our design and evaluation techniques.

I would like to express my appreciation to my family, especially my Dad and Mom, who were everyday waiting for this moment, for their unwavering support and confidence throughout this journey.

I am thankful for assistance from Discovery Grants through the Natural Sciences and Engineering Research Council of Canada (NSERC), and from the Industry Canada GRAND Network of Centres of Excellence.

Chapter 1

Introduction

1.1 Motivation

According to Oviatt [109], multimodal systems are "two or more combined user input modes— such as speech, pen, touch, manual gestures, gaze, and head and body movements— in a coordinated manner with multimedia system output." They provide interaction with the virtual and physical environment through natural modes of communication based on the five human senses. Such systems began to be developed in the early 1980s [11], but have recently become more widespread as sensor technology has become more advanced and less costly. Frameworks for multimodal systems have been created to address many issues relating to engineering and technical matters. Guidelines for interaction design, however, and methodologies for evaluation, are missing from such frameworks. These methodologies are especially vital because they can consider the emotional experience of users when they interact with a multimodal system. This thesis focuses on methodologies for affective multimodal software, with a focus on education as a specific domain.

The ability to communicate emotionally, as well as cognitively, plays an important role in human-computer interaction (HCI) in the context of education [105]. A positive mood can induce a different kind of thinking, because emotions are linked to other mental functions such as attention, perception, memory, decision-making and learning [67,115]. Our challenge is how practical techniques of HCI can address affective design and evaluation in multimodal systems. In the design of a multimodal educational system, we can consider the following four elements: **affect**, **multimodality**, **education**, and **software design**. Research has informed each of these four elements. For example, James Russell's circumplex model of emotion [129] proposes that all affective states arise from two dimensions, one related to valence (a pleasure–displeasure continuum) and the other to arousal (sensation of energy).

Chang et al. [17] have demonstrated how a multimodal system can support affect by developing a system involving a "Haptic Creature" to support users experiencing trauma. Toby Hede and Andy Hede [62] addressed learning that involves simultaneous interaction with multimodal media elements. Helene Gelderblom and Paula Kotzé [54] have suggested guidelines for software design based on theories of cognitive development for young children. However, as yet no research has made an explicit practical connection between all of these elements. Our goal is not to create a technical package or a toolkit, but to consider the context of education, affect and multimodality and how to support the system designer. We take an approach similar to that of User-centered Design, adapting well-known techniques for design and evaluation.

Affect is a complex subject, and there is much ongoing study to determine its nature. In a paper devoted to clarifying the issues, Russell [130] addresses "core affect" as follows:

> At the heart of emotion, mood, and any other emotionally charged event are states experienced as simply feeling good or bad, energized or enervated. These states— called core affect—influence reflexes, perception, cognition, and behavior and are influenced by many causes internal and external... [130]

He goes on to provide this definition: "Core affect: A neurophysiological state that is consciously accessible as a simple, nonreflective feeling that is an integral blend of hedonic (pleasure–displeasure) and arousal (sleepy–activated) values" [130]. He then provides some related definitions such as this: "Affective quality: The ability to cause a change in core affect" [130]. In this thesis, these are the definitions we accept.

In education, affective objectives were addressed in "Bloom's Taxonomy" as "Objectives which emphasize a feeling tone, an emotion, or a degree of acceptance or rejection. Affective objectives vary from simple attention to selected phenomenon to complex but internally consistent qualities of conscience" [80]. In the taxonomy, a wide range of objective are discussed, including those related to affective subject matter itself. Our affective education focus is on affective objectives that support cognitive objectives. In recent literature, this is sometimes called "affective teaching", as described by Zhang and Lu:

> So, affective teaching firstly aims to develop students' affective quality. Secondly,

> the brain areas involving emotion and cognition highly overlap. This means more than that they just influence each other. None of the ingredients that we deal with in education, such as "concept" and "emotions" and "behaviors," is separate. They influence and shape each other [16]. Affective teaching ought to play a key role of emotions to optimize cognition and enhance students' integrative quality. [156]

1.2 Research Goals

In investigating the role of affective strategies in designing and evaluating multimodal educational systems, we set out to answer the following research questions.

- What are the principles for a framework to support affective multimodal education?
- How might these principles be leveraged for design?
- How might these principles be leveraged for evaluation?

It is worth noting that while our research focused on the particular domain of education—with the goal of laying the groundwork for more specific future applications—we also hope that our findings might later also be applied to other fields, in a more general way. This thesis intends to contribute to the existing body of work on the topic, giving it the potential to affect the broader study of multimodal design and evaluation. Our investigation was pragmatic, requiring an operational understanding of cognitive and affective communication in the context of HCI.

1.3 Research Approach

Our intent was to support better design and evaluation of the user experience, when using multimodal software for affective education. The overall approach that guided the development of this thesis is outlined below.

We first reviewed existing frameworks for multimodal software and found an emphasis on

technical issues, but very little on affect. In HCI, there are well-established design processes and techniques for usability in general, but again very little relating to affect.

We then considered the issue of affect in the education domain, using Bloom's taxonomy for learning [10]. This suggested how affective learning might be considered in relation to cognitive and psychomotor learning. We identified a framework called Multimodal Affect for Design and Evaluation (MADE), which considers cognitive and affective strategies with multiple sensory modalities and learning objectives. Our framework connects the various elements, and specifies how sensory modalities might support educational learning objectives and affective strategies. In particular, we identify the stakeholders, consisting of several educator roles, and their relationship with the system designer. We proposed that HCI design methods could be adapted to help the system designer for affective multimodal education systems.

We next considered how to apply MADE to design and evaluation. We focused on several of the most widely used techniques in HCI, and adapted them as appropriate. For the design guidelines for affective multimodal education, we adapted personas from Goal-Directed Design (GDD) and we adapted Essential Use Cases from Usage-Centered Design (UsageCD). For the evaluation we adapted techniques of usability evaluation, taking into consideration several aspects. These included heuristic evaluation, and various forms of walkthrough evaluation: cognitive, affective and pluralistic. As a result, we propose two new evaluation methods: the Affective Walkthrough, and the Affective Heuristic Evaluation.

To apply and refine the techniques we developed, we conducted several case studies. We started by designing and implementing a multimodal education software system, and then conducted two studies that evaluated other systems. This process enabled us to assess and improve our techniques. The evaluation studies, in particular, helped us to better understand the challenges of design, and especially the role of the teacher. This led us to propose and implement a visualization system designed to support the teacher's management of affective learning.

This approach did have limitations, most importantly that our studies did not involve participants in the target audience, either learners (e.g. children) or domain experts (e.g. teachers). However, we believe our work is indicative, and further research involving such groups is needed.

1.4 Thesis Structure

As illustrated in Figure 1.1, this thesis is structured in accordance with its key elements.

- **Chapter 1** introduces the purpose and structure of the study.
- **Chapter 2** provides some background (including a literature survey) of frameworks for multimodal software, such as theories, guidelines and toolkits.
- **Chapter 3** identifies the theoretical background for multimodal affective education.
- **Chapter 4** proposes a design and evaluation framework for educational applications. We then outline the various stakeholders involved. We suggest learning objectives, affective strategies, cognitive strategies and modalities should inform the system designer, therefore, supporting the teacher to help learners engage more easily.
- **Chapter 5** examines the process of adapting existing design methodologies to the MADE framework. These include principles of GDD, including personas, and of UsageCD, including Essential Use Cases.
- **Chapter 6** describes our process of adapting existing evaluation methods to the MADE framework, to ensure that this software supports affective use of multimodal design. We propose an affective adaptation of the cognitive walkthrough, plus a modification to the pluralistic usability walkthrough by adding personas. We also proposed an affective adaptation of heuristic evaluation.

The next four chapters examine how we apply the MADE interaction design framework to three case studies, plus an exploration of how teachers could monitor their students' affective states. These case studies allow us to assess how well our proposed approach works. We

refined our design and evaluation techniques on the basis of our experience from the case studies.

- **Chapter 7: Case Study 1, MADE Ratio.** In the context of using sensory modalities in education, we look at the affective strategies that instructors use in a classroom to motivate students to learn mathematical concepts. The **purpose** of this case study is to explore our proposed design techniques; and the **plan of action** is to apply affective Essential Use Cases and affective personas in the design of a multimodal system for mathematical learning.

- **Chapter 8: Case Study 2, Museum Explorer.** We check the affective value of narrative-based interaction, a framework suitable for a classroom or a museum. The **purpose** of this case study is to explore our proposed evaluation techniques with an existing software application developed by colleagues; and the **plan of action** is to use affective walkthrough and affective heuristic evaluation, and then propose refinements.

- **Chapter 9: Case Study 3, Minecraft Hour of Code**. We evaluate an educational software application, checking the affective value of a learning environment using blocks of code. The **purpose** of this case study is to evaluate a professional and widely used multimodal software application; and the **plan of action** is to apply the affective heuristic evaluation and the affective walkthrough, and then propose refinements.

- **Chapter 10: The MADE Teacher's Dashboard.** The case studies of evaluation let us to identify improvements to our design case study. This chapter revisits that design and proposes better support for the teacher. We then propose a system of data visualization that allows instructors to monitor online learner emotions.

- **Chapter 11: Conclusions.** provides a summary of the thesis work, identifies the key contributions, and outlines opportunities for future work.

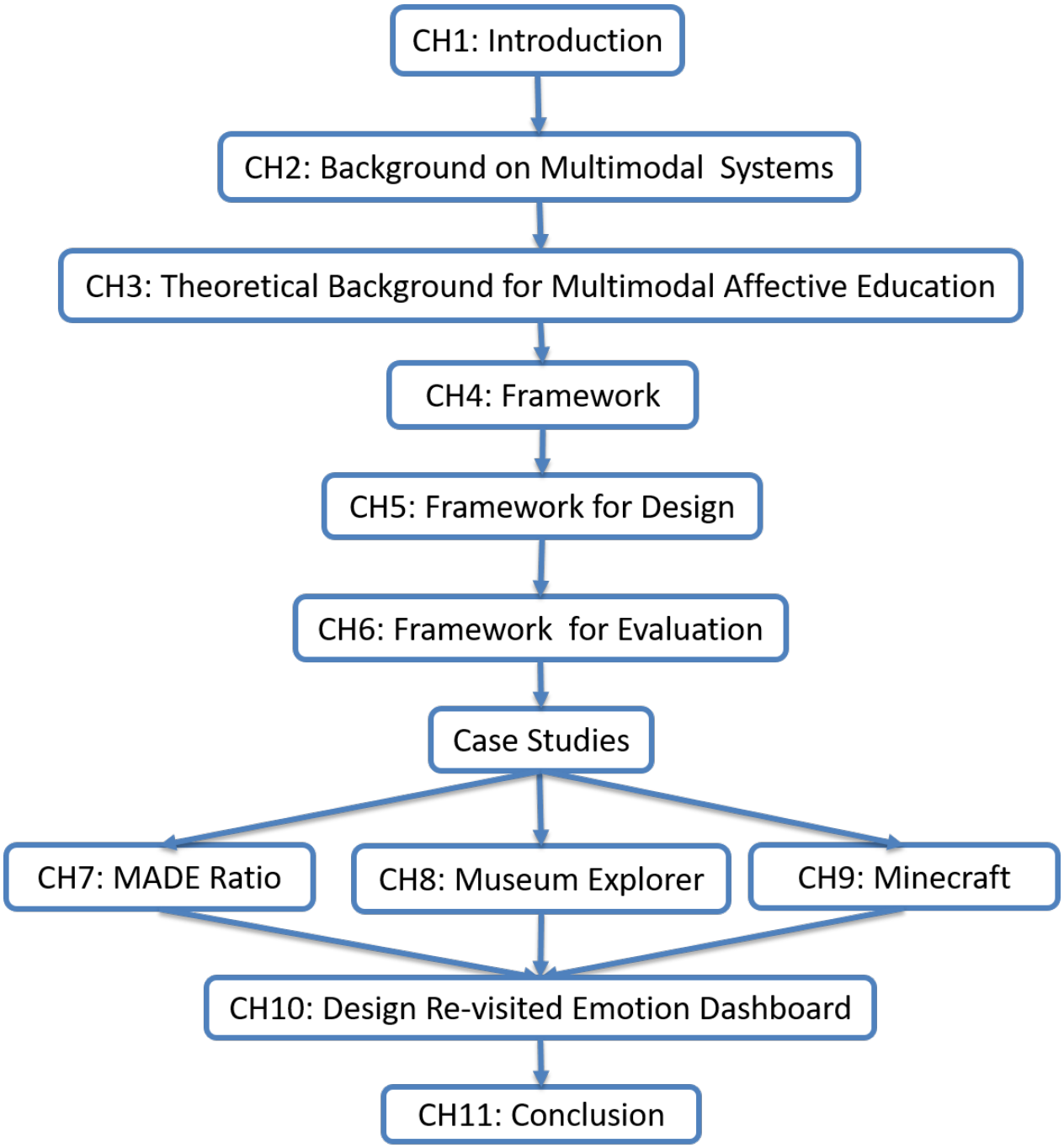


Figure 1.1: Thesis flowchart, showing how the chapters organize the various elements of this research.

1.5 Contributions

This thesis focuses on three main areas of research: multimodal interaction, affective education, and UI design and evaluation.

This thesis makes four contributions to the body of research on the use of affective multimodal systems for education, as outlined below.

1. We have created a **framework** to represent theories of multimodal interaction, based on multiple sensory modalities and quasi-sensory modalities. It offers a compact way to address affective multimodal systems.

    - We have utilized the domains of Bloom's taxonomy, considering that multi-sensory modality can help the affective and cognitive domains.

    - We have developed a new framework, showing the relationship between educator roles and the system designer. We suggest learning objectives and strategies come first, and inform the system designer in their work to support the teacher in facilitating student activities.

2. We have provided a **design** process to support the framework: one that facilitates learning and engagement in a multimodal system for the educational domain. We developed a set of guidelines based on the principles of GDD and UsageCD.

    - We adapted GDD personas for educational scenarios, and adapted UsageCD Essential Use Cases for affective multimodal education.

    - We conducted a case study to validate and refine our design techniques.

3. We created two **evaluation** approaches, that will enable us to better explore the effectiveness of multimodal education systems.

    - We created two novel evaluation methods, one adapted from the cognitive walkthrough: the Affective Walkthrough, and the other adapted from the heuristic evaluation: the Affective Heuristic Evaluation.

- We conducted two case studies, with findings that validated our evaluation techniques, and allowed us to refine them.

4. We created an analytical software system, the Teacher's Dashboard, that allows instructors to keep track of the emotions of their students.

   - The software will help teachers see how students' emotions change in real time, allowing them to determine whether learners are bored, frustrated or anxious; or positive, challenged or triumphant. This feedback lets teachers gauge the effectiveness of their strategies.

   - One novel feature of the web learning application is to show a visualization based on recognition of emotion on the learner's face, indicating their emotional state as they process an activity.

1.6 Publications

The following papers were published as part of this thesis work. With the exception of the first, all are the work of Reza GhasemAghaei, with the advice of his co-supervisors. Throughout the thesis, the term "we" is used accordingly.

- Farhadi-Niaki F, GhasemAghaei R, Arya A. Empirical study of a vision-based depth-sensitive human-computer interaction system. In *Proceedings of the 10th Asia Pacific Conference on Computer-Human Interaction, ACM*, 2012, pp. 101-108.

  – *This paper describes the result of our user study on part of a vision-based, depth-sensitive input system that allows users to perform typical desktop tasks through arm gestures. This study provided motivation and interest in multimodal software.*

- GhasemAghaei R, Biddle R, Arya A. The MADE Framework: Multimodal Software for Affective Education. In Proceedings of EdMedia: World Conference on Educational Media and Technology, AACE, 2015.

  – *This paper, the basis of Chapter 4, describes our MADE framework.*

- GhasemAghaei R, Biddle R, Arya A. Multimodal Software for Affective Education: UI Design. In Proceedings of EdMedia: World Conference on Educational Media and Technology, AACE, 2015.
  - *This paper, the basis of Chapter 5, describes our proposed design methodologies.*
- GhasemAghaei R, Biddle R, Arya A. Multimodal Software for Affective Education: UI Evaluation. In Proceedings of EdMedia: World Conference on Educational Media and Technology, AACE, 2015.
  - *This paper, the basis of Chapter 6, describes our proposed evaluation inspection techniques.*
- GhasemAghaei R, Arya A, Biddle R. Design practices for multimodal affective mathematical learning. International Symposium on Computer Science and Software Engineering, IEEE, CSSE, 2015, pp. 1-7.
  - *This paper, the basis of Chapter 7, describes our case study "MADE Ratio."*
- GhasemAghaei R, Arya A, Biddle R. MADE Ratio: Affective Multimodal Software for Mathematical Concepts, Lecture Notes in Computer Science (LNCS), Proceedings of the 18th International Conference on Human-Computer Interaction International, HCII'16, Toronto, Canada, 2016.
  - *This paper, the basis of Chapter 7, performs a user study on our case study "MADE Ratio."*
- GhasemAghaei R, Arya A, Biddle R. Evaluating Software for Affective Education: A Case Study of the Affective Walkthrough. Published at Lecture Notes in Computer Science (LNCS), In International Conference on Human-Computer Interaction, Springer International Publishing, 2016, pp. 226-231.
  - *This paper, the basis of Chapter 8, evaluates our case study "Museum Explorer using Affective Walkthrough."*
- GhasemAghaei R, Arya A, Biddle R. Evaluating Software for Affective Education: A

Case Study of Affective Heuristics. In EdMedia: World Conference on Educational Media and Technology, 2016, Vol. 2016, No. 1, pp. 567-574.

  - *This paper, the basis of Chapter 8, evaluates our case study "Museum Explorer using Affective Heuristics."*

- GhasemAghaei R, Arya A, Biddle R. A Dashboard for Affective E-learning: Data Visualization for Monitoring Online Learner Emotions. In EdMedia: World Conference on Educational Media and Technology, AACE, 2016, Vol. 2016, No. 1, pp. 1530-1537.

  - *The paper above, the basis of Chapter 10, describes the Teacher's Dashboard software.*

- GhasemAghaei R, Arya A, Biddle R. Affective Walkthroughs and Heuristics: Evaluating Minecraft Hour of Code, International Conference on Human-Computer Interaction, Springer International Publishing, 2017.

  - *This paper, the basis of Chapter 9, describes our case study "Minecraft Hour of Code."*

- GhasemAghaei R, Arya A, Biddle R. Experience with MADE: A Framework for Multimodal Affect for Design and Evaluation, Educational Technology Research and Development, Springer, 2017 (to be submitted).

  - *This paper presented our experience with creating and applying our framework for multimodal software design and evaluation, with a focus on affect in education.*

Chapter 2

Background on Multimodal Systems

This thesis focuses on multimodal interaction, particularly on using multimodal software for affective education. Many researchers and engineers have developed many different types of systems and applications using multiple modalities. This chapter examines the foundations of such interaction, and introduces its conceptual background—beginning with a brief history of multimodal systems, as well as their motivations, principles and conceptual models. Next, we describe some technical architectures and frameworks. We then introduce the concepts of HCI and user interface design related to multimodal software. Finally, we discuss multimodal systems, education and affect.

2.1 A Brief History

Multimodal interaction can be viewed as an expanding of the traditional desktop experience: from pure text, to the graphical WIMP user interface (windows, icons, menus, pointing device), to new ideas to allow users to interact with the system—and each other, and the world—more naturally using both verbal and non-verbal tools [144]. Such multimodal systems that expand the traditional experience support rich and natural ways to communicate, allowing users to interact through methods that are more natural and comfortable for them, as opposed to a fixed interaction system. They offer users two or more system input modes (such as gestures, speech and touch), and coordinate these with the multimedia system output [109].

Well-known early examples of multimodal systems include the "Put That There" demonstration developed by Richard Bolt and his colleagues at MIT in the early 1980s [11]; and the QuickSet tablet PC interface developed by Cohen et al. in 1997 [20]. In the first example,

the user communicated with a large screen display via speech and pointing gestures. The QuickSet was a pen/voice system running on an early tablet, which was a U.S. Marine Corps training simulator (see Figure 2.1). Neither of these multimodal systems considered affective states in their design, and neither explained their design and evaluation methodologies.

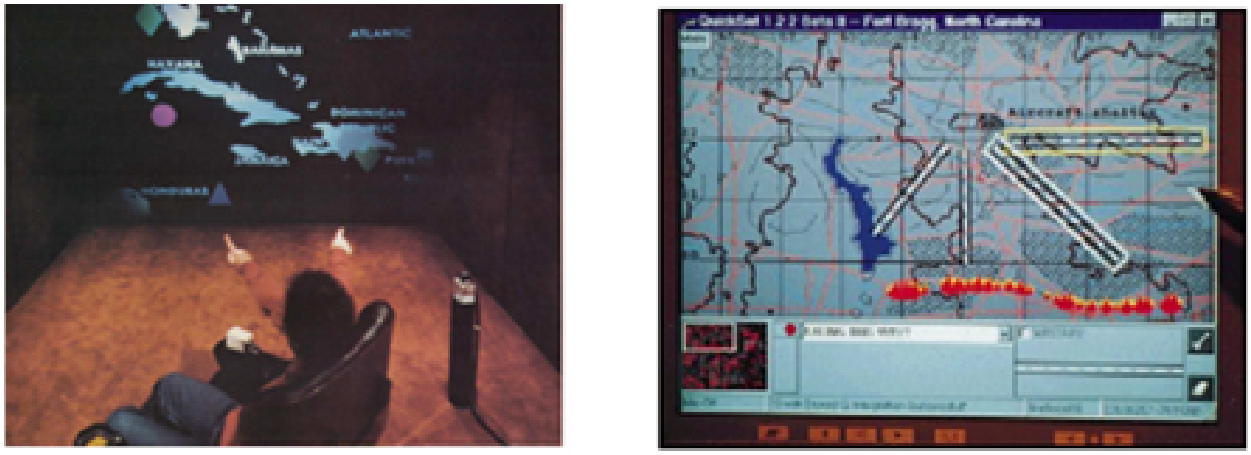


Figure 2.1: Left side: Talking and pointing to items on the Media Room's large screen [11]. Right side: The QuickSet tablet PC interface [20].

2.2 Motivation for Multimodal Systems

Between people, multimodal interaction is part of our everyday discourse. We communicate with one another by speaking, moving, gesturing, changing our facial expressions, and shifting our gaze [108]. In terms of HCI, the traditional GUI is limited in its ability to support the range of interactions that users need. What is necessary is a more natural, intuitive, adaptive and unobtrusive interface that can accommodate a wider range of tasks, scenarios and preferences. Multimodal interfaces offer that range by allowing users to input their commands through two or more combined modes, such as speech, touch, gestures, gaze, head and body movements [128, 143].

The advantages of multimodal systems include reducing users' cognitive load, and improving communication between them—since they can process information faster when it's presented in multiple modalities [109, 148]. Multimodal approaches provide a framework for the collection and analysis of visual, aural, embodied, and spatial aspects of interaction with the

environment, and the relationship between them [68,81]. These systems use recognition-based technologies to identify forms of human language and behaviour [144].

According to Oviatt [110] and Ruiz [128], other potential advantages of multimodal systems include robustness, flexibility, and the ability to minimize errors. Such systems can accommodate a wider range of tasks, situations, and users (including those with permanent or temporary disabilities); they also permit more flexible use of input modes, including alternation and integrated use. They can help to prevent people from over-using any individual mode by giving them alternative ways to interact. They can adapt to continuously changing environmental conditions. They increase performance, enhancing error avoidance and ease of error resolution. Multimodal inputs also increase robustness—the quality of communication between user and system—because conveying similar or related information through different modalities increases the likelihood of correct recognition. This can result in a high degree of naturalness. Another benefit is flexibility: users can perceive and structure their communication in different ways for specific contexts [110].

The growth of the multimodal interaction research community has been helped by the commercial explosion of smartphones, which has accelerated the jump to a post-WIMP world. These powerful mobile devices are well suited for multiple modes of interaction, since most have speech and gesture recognition. The introduction of 3-D vision sensors, such as the Microsoft Kinect or the Leap Motion Controller, offer further motivations to strengthen these features [144].

2.3 Principles and Conceptual Models of Multimodal Systems

The influential W3C multimodal interaction framework identifies these basic components [83]: users, input and output modes, interaction manager, application functions, session components, and system and environment (see Figure 2.2).

- *Users* enter data into the *system*, then observe and/or hear the information it presents.
- An interactive system uses multiple *input* modes such as audio and speech; and one or more *output* modes, such as audio, speech, text, graphics, files, or animation.

- The *interaction manager* is the logical component that coordinates data, and manages information flow from the input and output modalities.
- The *application functions* provide multiple modes for users to provide input, and receive output; these modes can be switched.
- The *session components* support state management, and temporary and ongoing sessions for applications (such as a multiplayer game, a chat room or a meeting room, etc.).
- The *system and environment* component enables the interaction manager to detect and respond to changes in device capabilities, environmental conditions, and user preferences—for aspects such as audio input, supporting colour, and other capabilities and configurations.

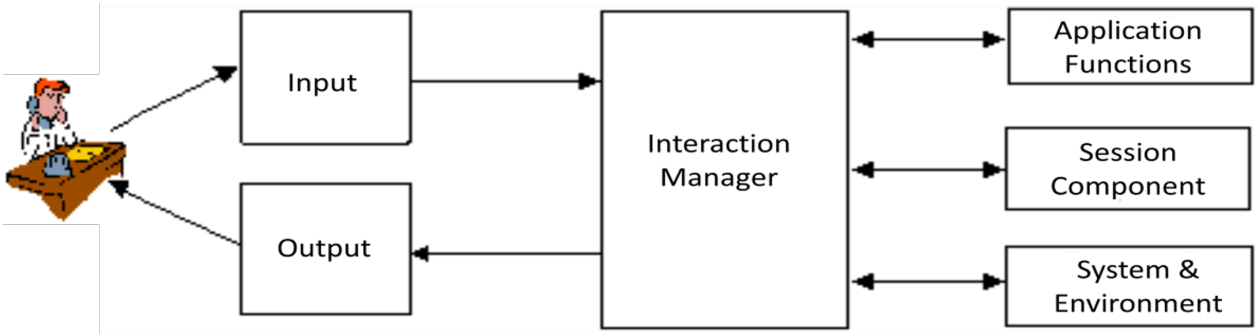


Figure 2.2: The multimodal interaction framework [83].

2.3.1 Fusion of Input Modalities, and Fission of Output Modalities

A comprehensive survey of multimodal system architecture has emphasized the integration of multimodal inputs (fusion) and outputs (fission) [44,51]. Figure 2.3 illustrates the interaction from the computer side, with the system's major software components. These components—a fusion engine, a fission module, a dialog manager and a context manager—together form the *integration committee*. As one can see, this figure shows the processing flow between the input and output modalities, the integration committee, and the potential applications. Input modalities are first perceived though a range of recognizers, and the *fusion engine*

communicates with the dialog manager to provide an output result from a common interpretation of the inputs. Finally, the *fission engine* communicates with the user through the most appropriate modality, depending on context and user preferences.

For incoming data, *fusion techniques* are needed to integrate input from different modalities. Figure 2.4 shows three levels for this fusion [135]. Each analyzes the same modality channel at a different level. The first, *data-level fusion*, deals with multiple signals coming from a similar modality sources. The second, *feature-level fusion*, is common when dealing with tightly coupled or time-synchronized modalities such as speech and lip movements. The third, *decision-level fusion*, is the most common type [44].

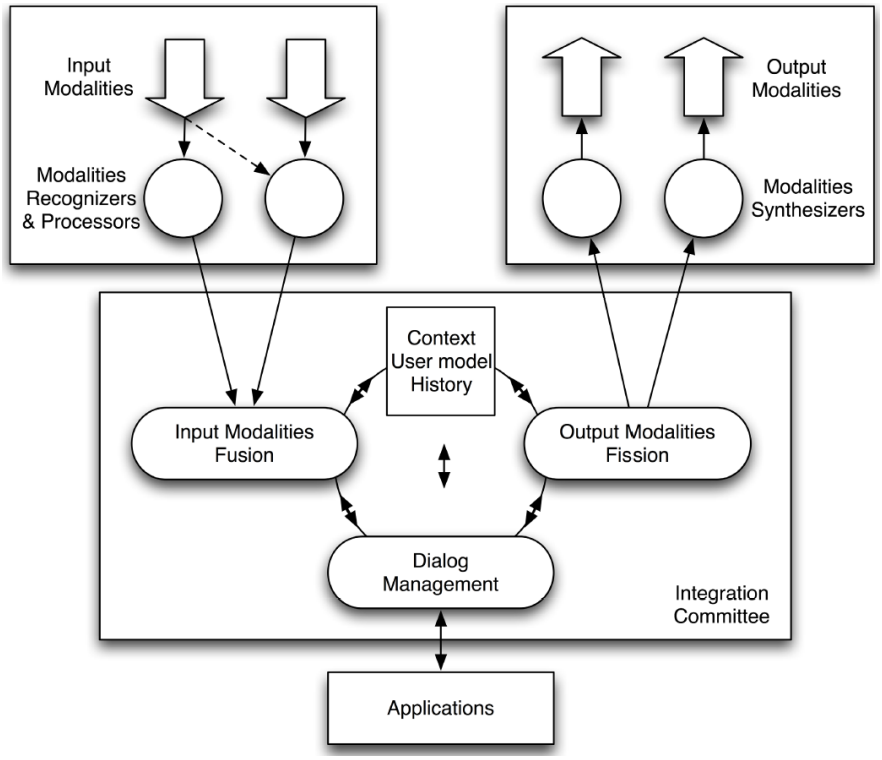


Figure 2.3: The architecture of a multimodal system, with the central integration committee and its major software components [44].

When users interact with a multimodal system, their communication can be divided into four different states: decision state, action state, perception state and interpretation state (as outlined in Figure 2.5) [44].

1. **Decision state:** the *user* prepares their message—either consciously and intentionally,

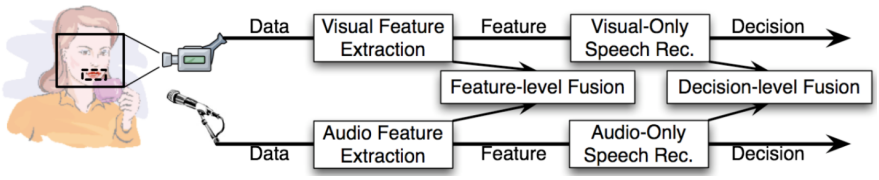


Figure 2.4: The various levels of multimodal fusion [44].

or unconsciously for attentional content or emotions.

2. **Action state:** the *user* selects their format for transmitting the message to the system, such as speech, gestures, writing, or facial expressions.

3. **Perception state:** the *system* receives information from one or more sensors, at one or more levels, and interprets the user's message.

4. **Interpretation state:** the *system* performs fusion of the messages, giving meaning to the information it gathered in the perception state.

5. **Computational state:** the system generates and transmits an answer (in the *action state*), based on the meaning it extracted in the *interpretation state*. A fission engine determines the most relevant way to return the message.

Figure 2.5 shows the four main states of a multimodal HCI loop, as described by Dumas [44]. While this is a valuable interpretation, it discusses affective aspects of communication (such as emotion and attention) only very briefly, it does not lead to design and evaluation methodologies.

2.3.2 Conceptual Models

A conceptual model, formed through experience, training or instruction, gives the user a useful mechanism for mapping interaction elements to a familiar scenario [143]. Many authors have created such models, and we will discuss several of them in this section. Rousseau et al. [125], for instance, proposed a model based on four concepts—What, Which, How and Then (WWHT)—to describe the process by which a multimodal system presents information (see Figures 2.6 and 2.7). The authors created an application platform designed to allow the

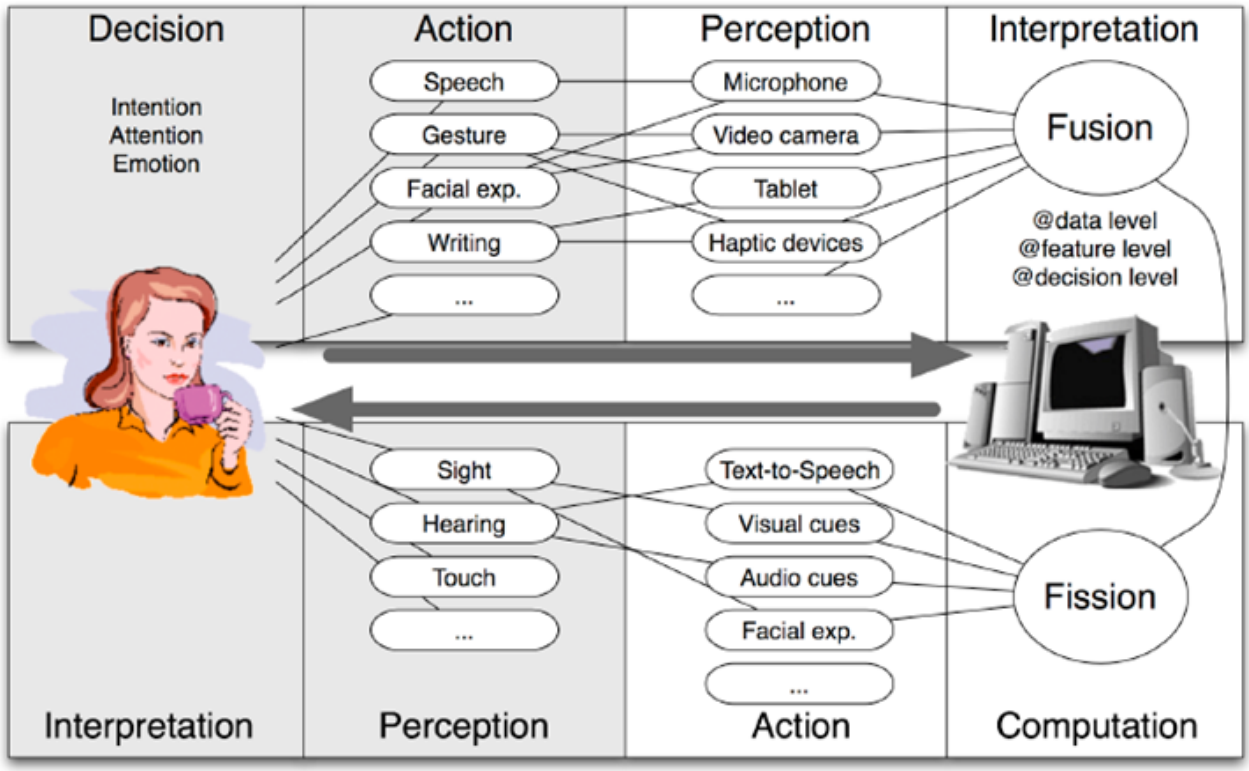


Figure 2.5: A representation of multimodal user system interaction loop, multimodal fusion (input interpretation), multimodal fission (output presentation) [44].

system's output [125]. The four elements are these:

- *What* is the information to present?
- *Which* modalities should be used?
- *How* should the information be presented with these modalities?
- *Then* how should the resulting presentation be handled?

The first three elements concern the initial presentation design, while the last concerns the presentation evolution.

Figure 2.7 illustrates the design process for adapting a multimodal presentation. The conceptual framework described by Rousseau et al., identifying the different steps, is based on four elements: information to present, interaction components, interaction context, and user behaviour. Again, however, the authors did not consider affective strategies, nor suggest

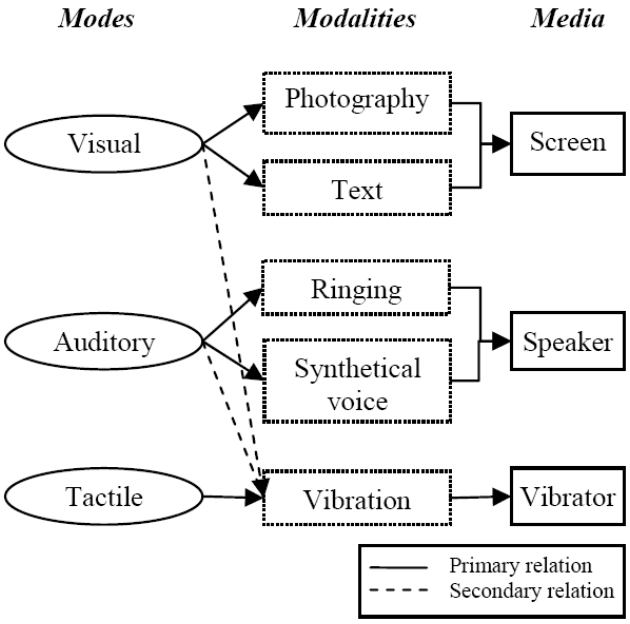


Figure 2.6: Interaction components diagram [125].

affective design and evaluation methodologies.

Another theoretical framework, based on types and goals for building multimodal systems, was proposed in 1998 by Martin el al. [91]. It postulates six basic types of cooperation between modalities (see Figure 2.8):

- **Transfer:** information produced by one modality is used by another.
- **Equivalence:** information may be transmitted by more than one modality.
- **Specialization:** a specific kind of information is always transmitted by the same modality.
- **Redundancy:** the same information is processed by a number of modalities.
- **Complementarity:** different chunks of information belonging to the same command are transmitted over more than one modality.

In 2009, Turri et al. [145] also provided a new conceptual framework for the design of multimodal user interfaces (UIs), using embodied cognition. The authors discussed the user's cognitive strategies when interacting with multimodal systems, and briefly explained some

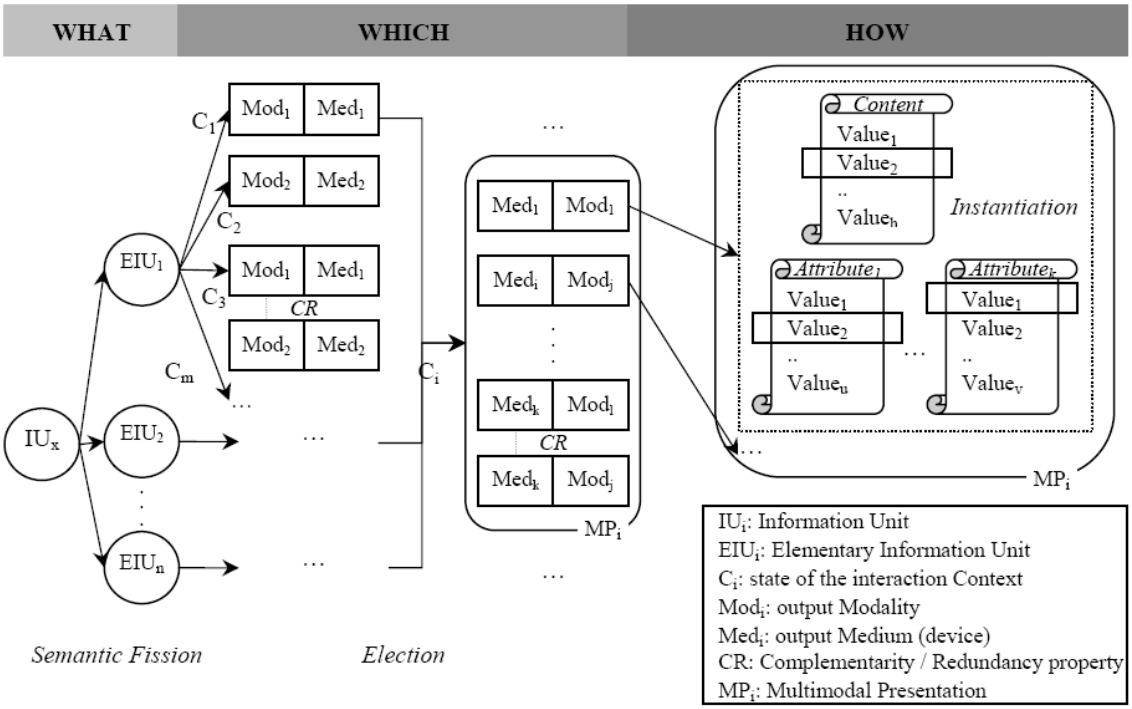


Figure 2.7: Design process of an adapted multimodal presentation [125].

affective strategies, such as emotion and attention, but went no further.

2.4 Technical Architectures and Frameworks

Many specialized toolkits and frameworks have been developed enabling their users to specify certain functionality with the intention of facilitating the creation of multimodal systems. In this section, we discuss three kinds: fusion-fission, rapid development, and controller supporting.

2.4.1 Fusion-Fission Support

Cuenca et al. [33] show that many toolkits are designed to facilitate the use of the fusion-fission architecture discussed earlier (section 2.3.1). They use a general model as a basis for comparison as shown in Figure 2.9. In this architecture, user inputs are recognized by recognizers, which are a group of particular software components. Examples of these

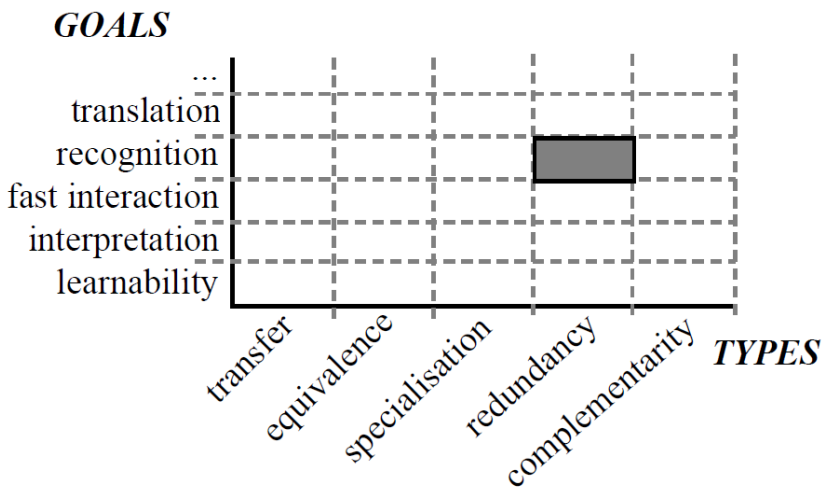


Figure 2.8: A theoretical framework for studying multimodality [91].

components are gestures, handwriting and voice recognizers. Recognizers interpret a stream of user inputs, and then inform the fusion engine. Fusion engines merge the information provided by all the recognizers in order to interpret the user's request. The dialog manager decides how to handle the user's request, and passes on to the fission component, which then chooses the synthesizers. The system output generation is the responsibility of the fission component. This model is used to analyse a range of different toolkits, including CoGenIVE, HephaisTK and ICon [33]. They show a strong commonality amongst toolkits, highlighting how the formal support for various elements in general model result in varying usefulness in each of the toolkits. For example, CoGenIVE [32] and HephaisTK [33] have a supported fusion engine, but no fission engine; ICon has neither.

- *User inputs*—such as gestures, handwriting and voice—are interpreted by a group of software components called *recognizers*.
- The *recognizers* inform the *fusion engines,* which merge the information provided by the recognizers.
- The *dialog manager* decides how to handle the user's request, and passes on the information to the *fission component*.
- That component then chooses the *synthesizers*, programs that control the most suitable *rendering devices*.

- The *fission component* is also responsible for generating the *system outputs*.
- Finally, information about a user's profile or interaction history is stored in data centers called *knowledge sources* [33].

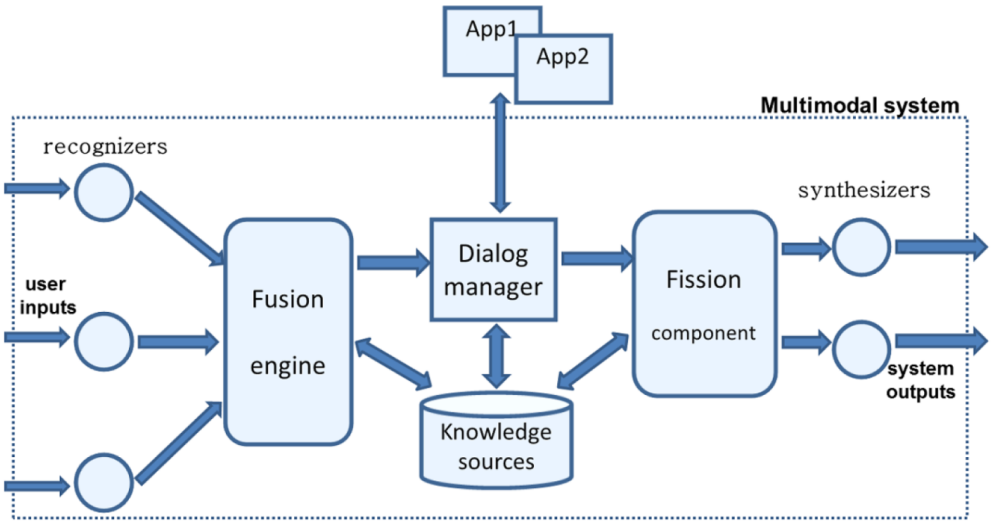


Figure 2.9: General architecture of a multimodal system [33].

2.4.2 Rapid Development

Cuenca et al. [33] created a toolkit for rapidly prototyping multimodal systems; it consists of a framework and a graphical editor. One source of this prototype was the "Put That There" interaction technique [11], which displays a series of objects on a touch-sensitive screen. Users can move these objects by using speech and pointing: for example, using the voice command "zoom here" while touching a region area of the screen (as shown on the left side of Figure 2.10). Both these functions can also be sent to the system via a visual model (as shown on the right side of Figure 2.10 [32, 33]. However, the authors of this project did not consider affective and cognitive strategies, or design and evaluation methodologies.

Another multimodal interaction framework—called i*Chameleon, developed by Tang et al. [141]—makes it possible for programmers to prototype and test new interactive devices or interaction modes. This framework allows users to customize their desktop environments, enabling interaction beyond the usual KVM (keyboard, video and mouse) devices. They

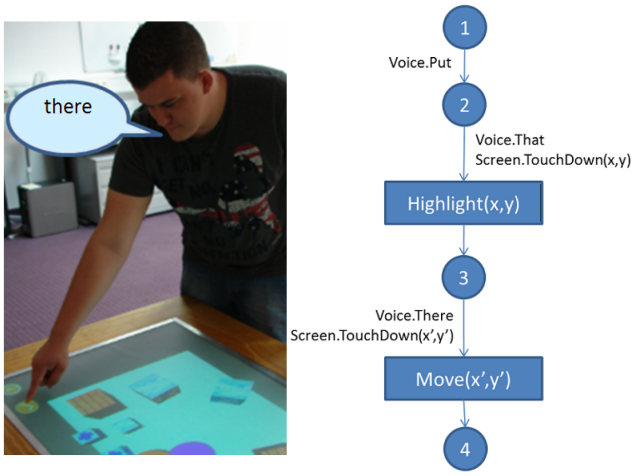


Figure 2.10: Left side: End user interacting with a multimodal system. Right side: Visual model used for specifying human-machine interaction [32].

may, for example, develop new applications with interactions such as distributed applications in collaborative environments, or robot control. Figure 2.12 shows the architecture of the i*Chameleon framework, designed to be modular, extensible and scalable [141].

Figure 2.11: User controlling Bing Maps and Google Earth with hand gesture and voice command based on i*Chameleon [141].

The research team of Van de Camp and Stiefelhagen [146] presented a generic framework called GlueTK, used to build different applications. It supports a wide range of input modalities beyond point- and touch-like interactions, such as speech, person tracking, head-pose estimation, and face identification. This allows interfaces to be designed for a variety of

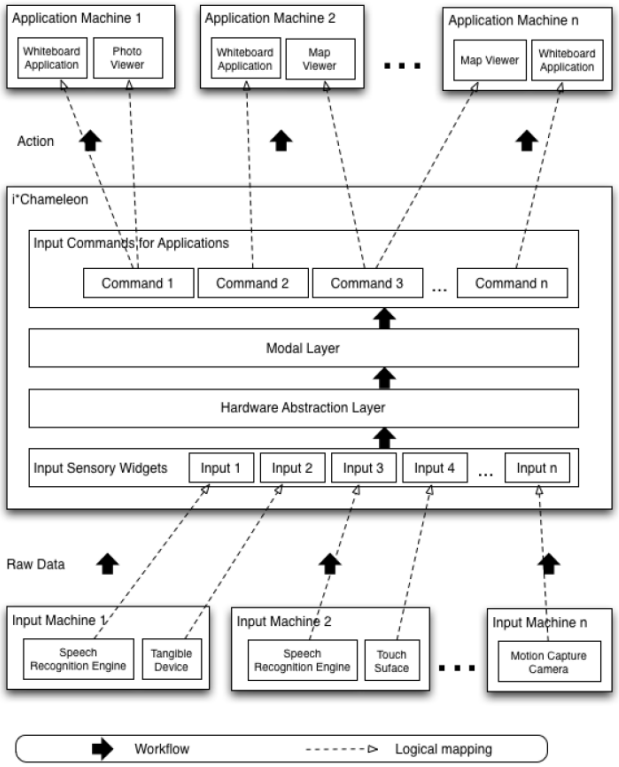


Figure 2.12: The i*Chameleon architecture [141].

display sizes. As an example, the GlueTK authors presented a controlled interactive room application (see Figure 2.13) [146]. However, they considered technical architecture only.

A novel algorithm for fusing input modalities and managing input data was presented in 2012 by Dumas et al. [45]. Their framework, called HephaisTK, is a toolkit for helping developers to create multimodal interfaces (see Figure 2.14). It focuses on the study of fusion algorithms, and on the high-level modelling of human-machine interaction. HephaisTK uses Synchronized Multimodal User Interfaces Modelling Language (SMUIML) to describe HCI dialogue, including scenarios and integration mechanisms. It is linked to the CARE aspects of multimodal interaction—Complementarity, Assignment, Redundancy, and Equivalence—that may occur between interaction techniques [29, 45].

Figure 2.13: Top: Room setup with back projection video wall and digital table. Below: Left. Visual feedback indicating the face identification in progress. Middle: Pointing gesture to select items in the locked menu. Right: Map layer selection with pointing gesture on the digital table [146].

2.4.3 Controller Support

Another kind of framework for multimodal systems relates to management of control devices. The examples we discuss are OpenNI [87], TUIO [71], OpenRemote [65] and OpenURC [22]. While these frameworks are influential and in active use, there is much current discussion on how they might lead to a general framework for the Internet of Things [95]. OpenNI (Open Natural Interaction) is a 3D sensing development open source framework. The original framework was acquired by Apple and closed down [151], and a later version created to carry on the work. It is partially developed by PrimeSence, which is one of the creators of the Kinect. [87]. Its open source SDK is the de-facto standard for developing computer vision middleware and 3D solutions. It can support programming with RGB-D devices for NUI applications. It is commonly used to work with the Microsoft Kinect, but can also support PrimeSense 3D sensors. OpenNI can support multiple sensors in the same system, allowing a composite 3D model of the interactive environment (see Figure 2.15).

TUIO (Tangible User Interface Objects) is an open framework that defines a common protocol and API for tangible multitouch surfaces. The TUIO protocol allows the transmission of an

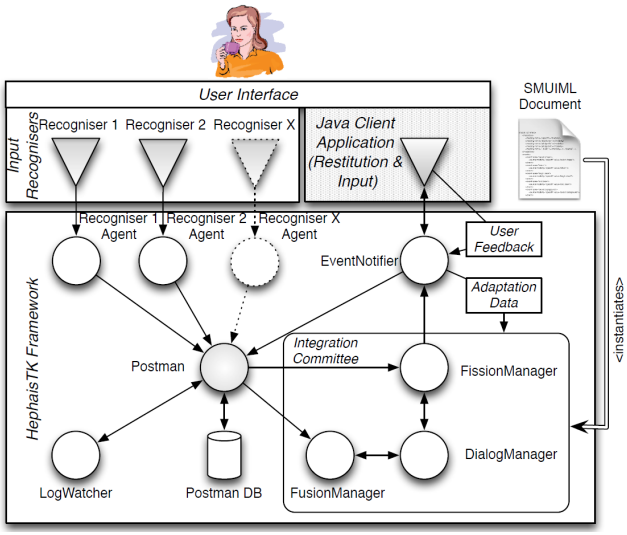


Figure 2.14: The HephaisTK architecture [45].

abstract description of interaction events from surfaces, including touch events and tangible object states. This support allows software to be written independently of specific control techniques (see Figure 2.16).

OpenRemote is a centralized open-source platform designed to address the challenges of integration between many different protocols and solutions available for home automation. The architecture has three elements: configuration tools, a local runtime controller that supports automation roles, and user control panels [94] (see Figure 2.17).

OpenURC is an initiative to promote the concept of a Universal Remote Console (URC). There are associated ISO standards for its application in products. This facilitates user interfaces that are simple and intuitive to use, where remote consoles are independent from the products they control, and inter-operate across a range of products. This supports accessibility and also should support use of future technologies such as natural language and computer brain interaction, among others. Early versions emphasized stand-alone connections, but more recent development addresses network and cloud connectivity (see Figure 2.18) [22].

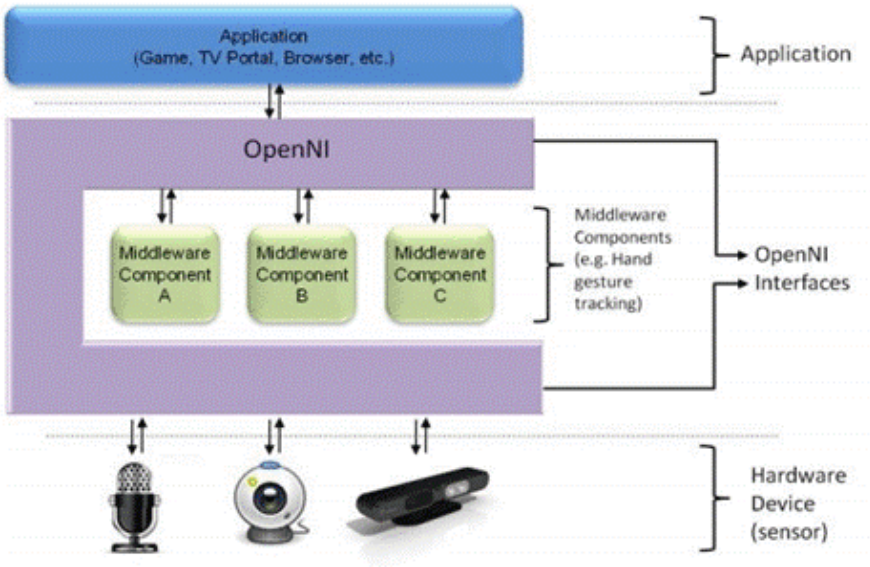


Figure 2.15: The OpenNI scheme architecture [87].

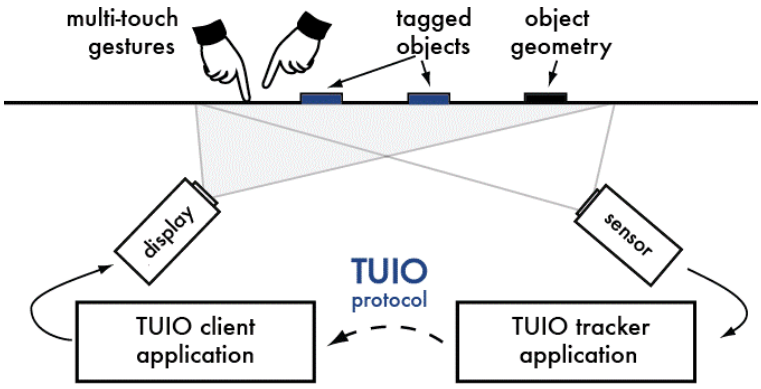


Figure 2.16: Tangible user interface objects protocol: a common protocol and API for tangible multitouch surfaces [71].

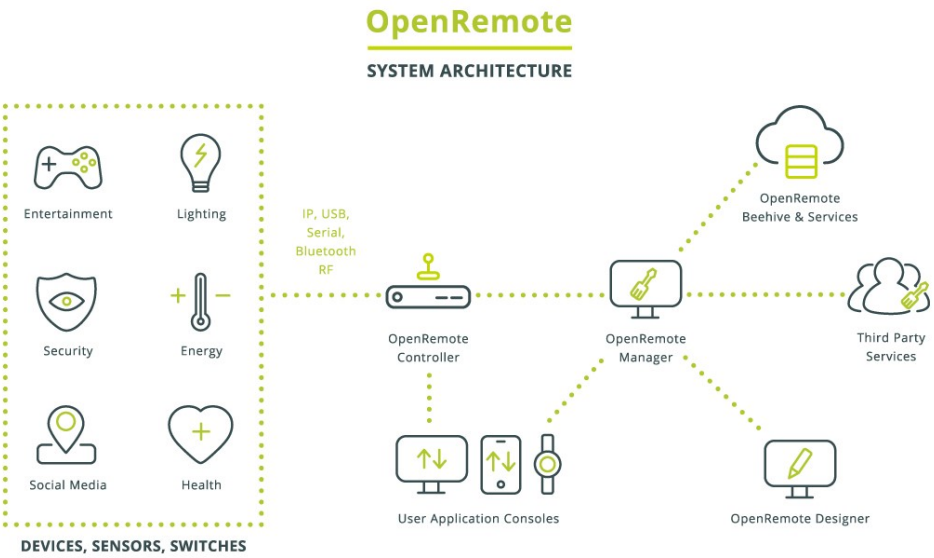


Figure 2.17: System architecture of OpenRemote, an open source framework for the Internet of Things [65].

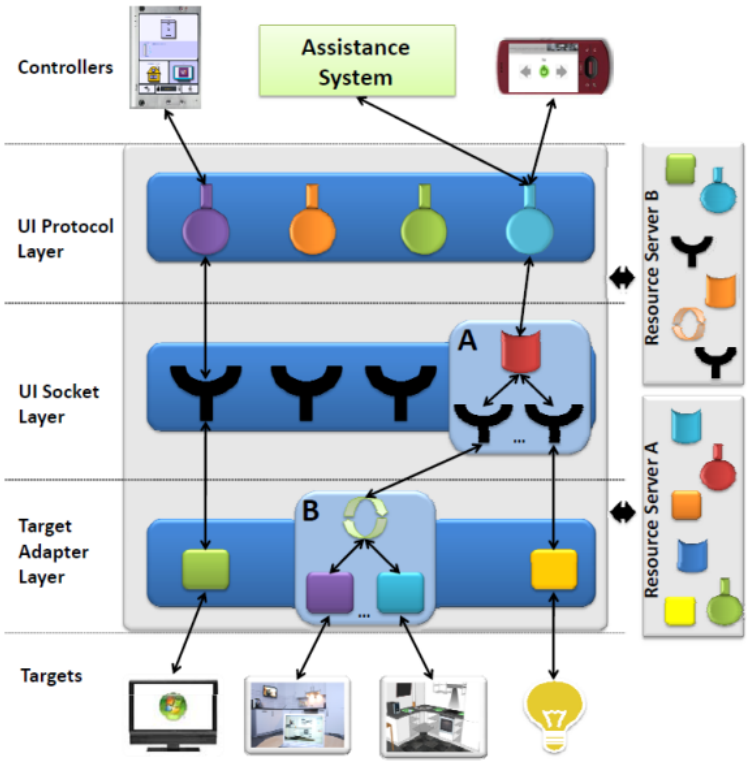


Figure 2.18: Conceptual view of a URC driven infrastructure [22].

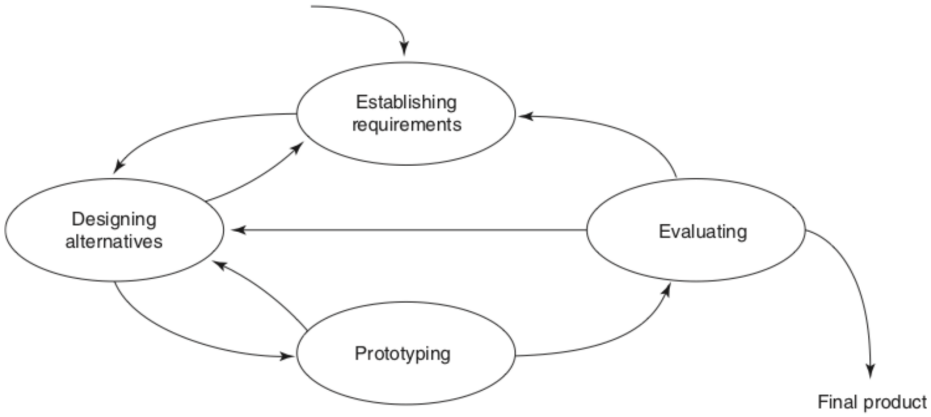


Figure 2.19: A simple interaction design lifecycle model [117].

Table 2.1: Preece et al.'s usability goals [117].

Usability goals	
Memorability	Once learned, how easy is it for a user to remember how to use the system?
Learnability	How easy is it for a user to learn how to use a system?
Utility	Does the system provide the functionality a user needs?
Effectiveness	How well the system performs its intended task?
Efficiency	How well does the system supports users in carrying out a task?
Safety	How well does the system protects a user from an undesirable (or even dangerous) situation?

2.5 UI Design for Multimodal Systems

The goal of effective HCI is to allow a wide range of users, of various ages and abilities, to interact with computers on many tasks and environments. As a result, when we design software, the aim is to improve users' experience with the system. To this end, many authors—such as Preece et al. [117] and Norman [106] —have defined some usability goals and design principles. The first of these are shown in Table 2.1. Table 2.2 shows the design principles that are vital for creating positive user experiences [106].

Best practice HCI requires a user-centred approach to design, engineering, and testing. This includes considering four major aspects: design, human factors, devices, and software [143]. Designing interactive systems requires a deep understanding of users' individual personalities,

Table 2.2: Summary of Don Norman's Design Principles [106].

Design principles
Visibility
Feedback
Affordance
Mapping
Constraints
Consistency

emotions and behaviours, and the conditions in which they will interact with the system [37]. The principles and techniques used in traditional GUIs do not necessarily apply to multimodal systems, since many other important aspects must also be considered: inputs and outputs, information design, fusion and fission of modalities, adaptability, consistency, error handling, etc. [100]. Reeves et al. [119] discussed six main categories of design guidelines and principles; these are outlined below in Table 2.3). As can be seen, these are at a very general level.

The user-centered approach is always regarded as iterative, involving a cycle of design and evaluation. This is made specific in standards such as ISO 13407 and 9241 [69]. Is also commonly highlighted in textbooks on HCI, such as Sharp et al. [117], see Figure 2.19.

To apply this approach, we need techniques for usability evaluation. There are two ways to measure the effectiveness of a user experience: we can work with either system experts, or the users themselves. Each brings certain strengths: the experts can do heuristic evaluations [104] or cognitive walkthroughs [150], while the user's experience is likely to be more direct.

In the former, evaluators use a set of predetermined usability principles based on knowledge of the typical user; in the latter, users participate in studies, typically in a laboratory setting. Such data-gathering can include interviews, surveys, video-recordings, and instrumented measurements. A more specific type is the field study, done in an informal setting rather than a lab.

While this whole approach has become widely adopted for UI design in general, there is little or no work on how to apply the approach to the design of multimodal affective systems.

Table 2.3: Guidelines for multimodal UI design [119]. In section 2.6 we show some multimodal systems in education domain.

Guidelines for multimodal user interface design		
1	Requirements specification	Design for broadest range of users and contexts of use Address privacy and security issues
2	Designing multimodal input and output	Maximize human cognitive and physical abilities
3	Adaptivity	Adapt to the needs and abilities of different users, as well as different contexts of use
4	Consistency	Presentation and prompts share common features and refer to a common task including using the same terminology across modalities
5	Feedback	Users are aware of their current connectivity and know which modalities are available Users are aware of alternative interaction options without being overloaded by lengthy instructions that distract from the task
6	Error prevention/handling	User errors are minimized and error handling improved by providing clearly marked exits from a task, modality, or the entire system, allowing users to undo a previous action or command

2.6 Multimodal Systems and Education

One environment where multimodal systems are increasingly being adopted is education, where successful teaching and learning depend on the important factor of engagement. This can best be provided through interactivity, which can make learning stimulating and even fun. Technologically advanced learning environments that include such systems have been shown

to offer new opportunities for physically mediated interaction. However, the relationship between learning, interaction and embodiment is a challenging one, as is the task of developing effective methodologies for examining real-time cognition and action.

Several authors have attempted this task: for instance, Price and Jewitt [118] offer an approach to analyzing embodied interaction in a tangible learning environment, using pairs of students. These authors considered cognitive learning strategies, but not affective strategies. In another work, Toby Hede and Andy Hede [62] addressed learning that involves simultaneous interaction with multimodal media elements, as illustrated in Figure 2.20. As the system input increases, the learner must compensate by increasing the amount of cognitive information-processing, and the working memory, which is applied to the learning situation. The learner's motivation, attention, and time spent on the task are essential [62]. Again, these authors considered the cognitive aspects but not the affective ones; they also did not consider design and evaluation methodologies.

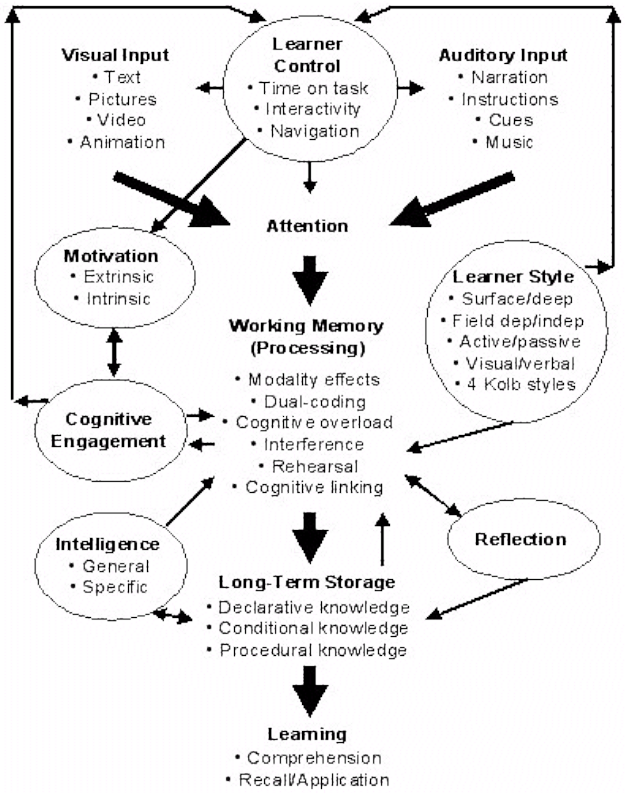


Figure 2.20: Integrated model of multimedia effects on learning [62].

Lee et al. [85] designed a principled "edutainment" system, called Dreamware, that provides sensibility treatment and intelligence training for children with developmental disabilities. The toy-like multimedia content of this authoring tool included a real-time monitoring tool that allows instructors to observe the children's learning, and provides training results. It also offers the children the opportunity to use their visual, auditory and tactile senses (see Figure 2.21). The authors considered some cognitive and affective aspects of learning for motivating children with disabilities, but these were not applied to all kind of users. The authors also did not consider design and evaluation methodologies.

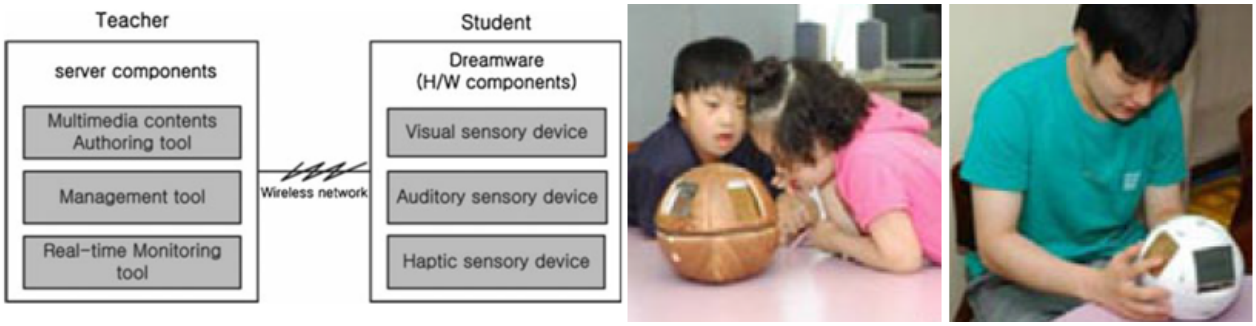


Figure 2.21: Left side. Configuration of our edutainment system, Right side. Playing and learning children with Dreamware [85].

2.7 Multimodal Systems and Affect

A central interest of this thesis is how multimodal systems can best be designed for learning in the domain of affect—a term that refer to the positive or negative valence of an emotional experience (see Figure 2.22). Many findings in different disciplines illustrate that affect, intertwined with cognition, can result in rational behavior, decision-making, creativity and better learning [115]. The role typically seen for multimodality, however, simply relates to sensors that can detect emotion. This approach has been applied to education, where the sensors feed into an artificial intelligence system that attempts to learn appropriate affective responses [154].

An emotion can be defined as an internal mental and affective state. Emotions may be thought of as a purely organic attribute. But as Rosalind Picard pointed out in 1997: "If we

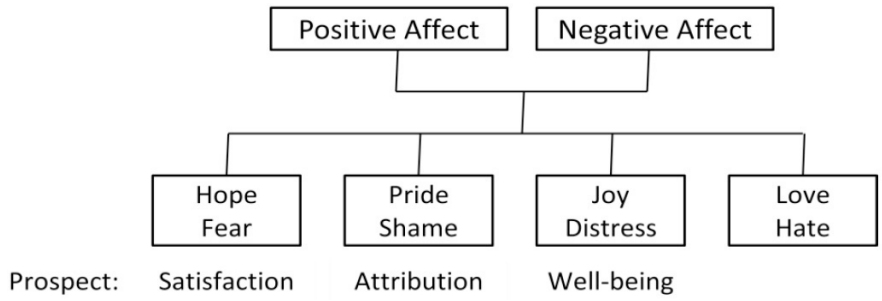


Figure 2.22: Positive or negative valence of an emotional experience prototype model.

want computers to be genuinely intelligent, and to interact naturally with us, we must give them the ability to recognize" emotions—and even, she posits, "to have and express" them. Emotion, after all, can be likened to other functions such as learning, attention, perception, memory and decision-making. Theories on emotion generally have two components, *cognitive* (mental) and *physical* (bodily) [114,116]. James and Sebe [67] have grouped vision techniques according to the human body (see Figure 2.23). Body movements, gestures, and gaze analysis are used for tasks such as emotion recognition in affective interaction for different applications. The authors explained affective HCI and considered affective strategies in their system, but did not talk about design and evaluation methodologies.

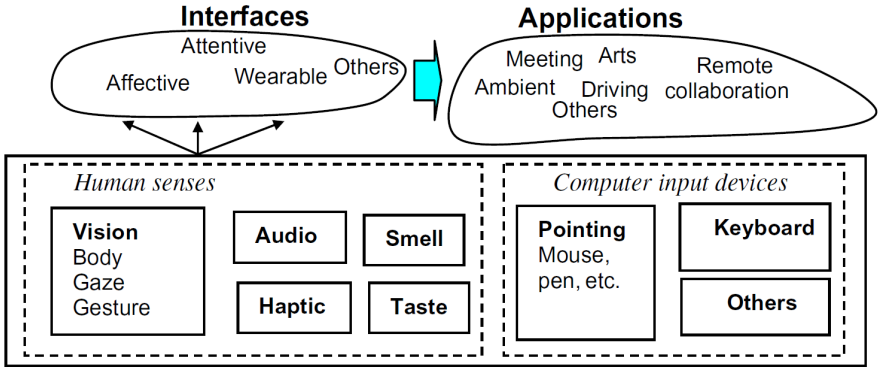


Figure 2.23: Overview of multimodal interaction using a human-centered approach [67].

An example of a multimodal system using affect is the animated robot called the Haptic Creature, created by Chang, MacLean, and Yohanan [17] to investigate affective touch. Informed by studies of human-to-human social touch, which communicates or evokes emotion, the Creature features an improved sensor system. The authors described their first-generation

gesture-recognition engine; but they did not consider cognitive strategies, or design and evaluation methodologies. In a similar project, Stiehl et al. [140] created the Huggable Creature. Its sensitive skin has four modalities of human (and animal) somatic perception: touch, pain, temperature, and kinesthetic information. Equipped with small video cameras, microphones, and an inertial measurement unit, the body "knows" how it is being held and touched [140]. By designing a robotic companion to illustrate the positive benefits of companion animal therapy, the authors considered affective aspects; but they did not consider the user's cognitive issues, nor design and evaluation methodologies (see Figure 2.24).

Figure 2.24: Participant interacting with the Haptic Creature during study [155].

A similar project by D'Mello et al. [47] developed an affect-sensitive Intelligent Tutoring System called AutoTutor (see Figure 2.25). As students learned physics, computer skills, and critical thinking with this system, it first recognized affect and then responded to it. The AutoTutor detected the emotions of a learner by monitoring their facial features, conversational cues, and body language. While it considered learners' affective and cognitive states, and embodied pedagogical agents synthesizing affective responses through animated facial expressions and speech, it did not talk about design and evaluation methodologies.

A similar system—the Multimodal Affective and Reactive Character, called MARC—was developed by Courgeon et al. in 2008 [28]. The authors explained how this agent (see Figure 2.26) integrated image-processing of a user's facial expressions with the FaceReader analysis tool, to demonstrate its real-time capabilities [28]. They considered affective interaction between the user and virtual agent, but not users' cognitive issues, or design and evaluation

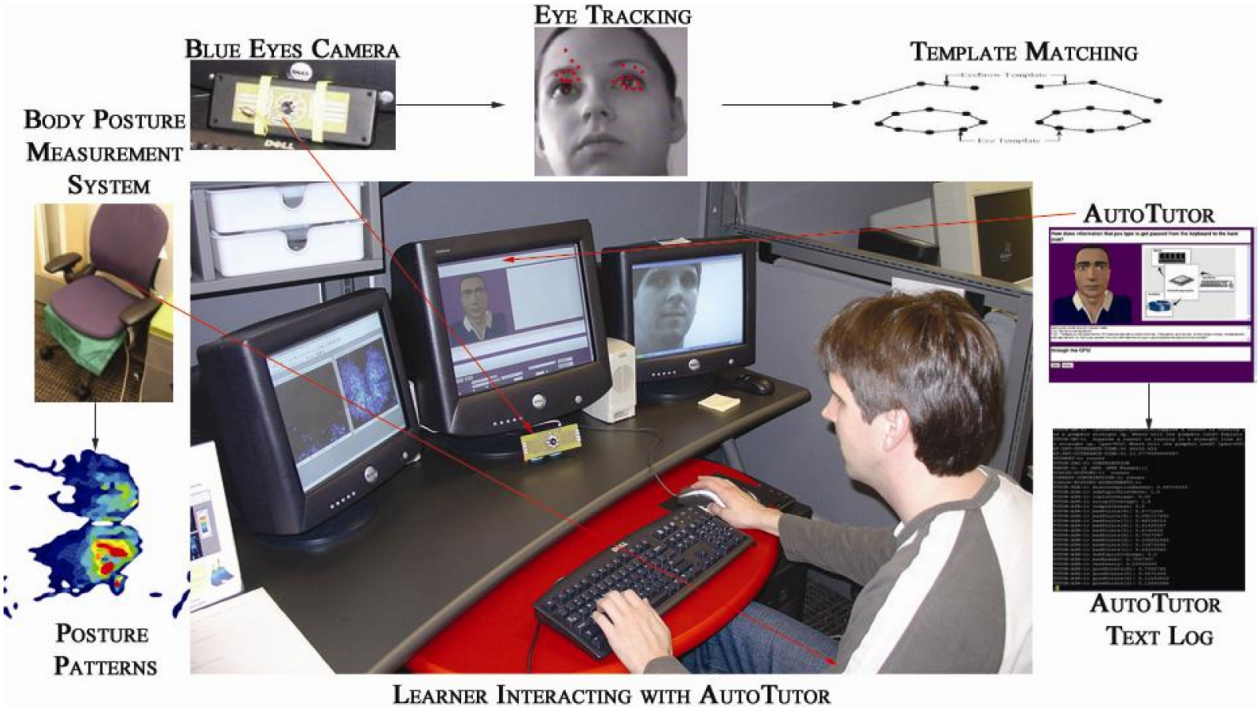


Figure 2.25: Sensors used for affect detection as learner interacts with AutoTutor [47].

methodologies. Affect-aware tutors [154] is another example of a system that recognizes and responses to student affect in teaching mathematics.

2.8 Summary

In this chapter, we discussed earlier work and developments in the general field of multimodal systems. By far most of the frameworks and models address technical issues, rather than affect or UI design. Even where affect or UI design is addressed, there is no focus on UI design for affective learning.

Table A.2 in appendix A provides a summary of each of the frameworks we reviewed, their contributions and the gaps we identified. All of these frameworks are missing affective design methodologies and evaluation techniques in their work. Many of these systems consider technical architecture only, and do not address affective issues in their design, or emotional aspects in their multimodal architecture.

Figure 2.26: Right side: Recognition of a blend of anger and disgust in user's facial expression by image processing. Left side: facial expression of this blend by the MARC agent [28].

One of the most related research projects was the "AutoTutor" system [47], which considered the learner's affective and cognitive states. It gave affective responses through animated facial expressions and speech. However, HCI design and evaluation techniques were not addressed.

Chapter 3

Theoretical Background for Multimodal Affective Education

3.1 Introduction

The world is becoming increasingly computationally-enabled, thanks to technologies such as mobile devices, smart objects, and the Internet of things (IoT). New forms of interaction such as gesture and speech, and those including cognition and emotion, are rapidly emerging, and the computation is moving beyond traditional desktop [7].

In this chapter we explore the background for multimodal affective systems for education. From a design perspective, over the past decade there has been an increasing interest in studying the task of learning—whether as a behaviour, as concept-building and experience, or as social communication. There seems to be a need for a new conceptualization of learning that emphasizes affective and cognitive concepts.

As we have described earlier, multimodal systems, which coordinate two or more user input modes—such as speech, pen, touch, manual gestures, gaze, and head and body movements [109]—are a new class of interface. They aim to recognize natural forms of language and behaviour, and incorporate recognition-based technologies such as speech, vision, or handwriting [110]. Such interfaces seek to be flexible and adaptable, as well as powerful. But to accommodate a wider range of scenarios, tasks, users and preferences, interfaces must become more intuitive, adaptive and unobtrusive. Accomplishing this goal is a primary motivation for developers [143].

Nobody denies the role of affect in learning. Certainly teachers know that aspects such as motivation, emotion, interest, reward and attention all play a crucial role. Research has demonstrated, for example, that even a slightly positive mood does not just make students

feel better; it also induces in them a different *kind* of thinking [115]. This is because our emotions are intricately linked to other mental functions such as attention, perception, memory, decision-making and learning [67]. To capitalize on this fact, the need for more precise theory is being driven by growing efforts to build technologies that truly interact with learners—motivating, engaging, assisting and challenging them in new ways [115]. HCI systems that can sense affective states such as stress, inattention, anger, boredom, etc., and that can adapt and respond to these, are likely to be perceived as more natural, trustworthy and effective. Hence our focus on considering the design and evaluation issues raised by HCI.

3.2 Educational Domains

Education is a very large topic with many different perspectives. One definition is that of John Dewey, an American philosopher, whose work has been influential in shaping contemporary understanding. Dewey proposed: We thus reach a technical definition of education: It is that reconstruction or reorganization of experience which adds to the meaning of experience, and which increases ability to direct the course of subsequent experience [38]. This process can be supported by several kinds of facilitation, each of which have somewhat differing goals. In 1956, Benjamin Bloom led a group of educational psychologists in a collaborative effort to develop a classification system of educational goals for the process of learning and educational objectives. Bloom then coordinated the classification for educators and presented the Taxonomy of Educational Objectives [10]. This identified three domains of education, the cognitive, affective and psychomotor domains. Bloom was one of the first instructional theorists to distinguish these domains of knowledge and human learning, and the taxonomy of the cognitive domain especially is widely acknowledged as the most prominent of three domains used by teachers and trainers in many of classrooms [43,142]. As well as distinguishing these three domains, the 1956 publication provided a comprehensive documentation of the cognitive domain. Later, Bloom and others provided a similarly detailed presentation about the affective domain [80]. Work exploring the psychomotor domain was later provided by others such as Harrow [61]. Although some early, and much later, work was done by others, this whole general approach is still typically known as “Bloom’s Taxonomy”.

For example, in the task of teaching teenagers to drive a car, cognitive strategy involves teaching them the rules of the road; affective strategy involves teaching them to drive safely; and psychomotor strategy involves teaching them the physical skills of driving. The original goal of Bloom's taxonomy was to motivate educators to address all three domains in their teaching, thereby giving students a more holistic education.

In this chapter, we examine four aspects of learning objectives:

- Cognitive strategies
- Affective strategies
- Multi-sensory modalities
- Quasi-sensory modalities.

3.3 Cognitive Strategies for Learning

Human cognition is distributed over a broad system of mental processes and tools. These include attention, perception, memory, learning, reading, speaking, listening, problem-solving, planning, reasoning and decision-making [117]. Nor do these cognitive processes act alone: many are interdependent, and several may be involved in a given mental activity. When you study for an exam, for example, you need to perform a number of tasks at once: attend to the material, perceive it, recognize it, read it, think about it, and try to remember it [117]. Depending on what the specific learning objective is, cognitive strategy may change from one type to another. Common cognitive strategies include: scaffolding, tracing, embodied cognition, and situated cognition. We will examine each of these in turn.

3.3.1 Scaffolding

This concept refers to the process by which instructors provide temporary support for students in the learning process. Like physical scaffolding, the support "builds up" learners to help them master the basics and accomplish specific tasks they would be unable to without

assistance. A learner's level of actual development is scaffolded to the level of their potential development. Once the scaffolding is no longer needed, it is removed: over time, the teacher or More Knowledgeable Other (MKO) increases the students' independence by shifting more responsibility for learning onto them [126].

For designers, an example might be creating activities that scaffold students in the learning task of explaining things clearly—a learned skill that they need to practice. Especially with younger students, their explanations may often be incomplete, too brief, or miss important points. We might ask a group of students to explore a series of scaffolded activities to facilitate their "learning to explain," and closely observe them during this process. Initially they need step-by-step instruction; but gradually this scaffolding can be removed, and the students can become more autonomous [131].

Belland et al. [6] developed some evidence-based guidelines to create this kind of scaffolding, with six recommendations (see Figure 3.1):

- Embed scaffolds within a system.
- Have students articulate their thoughts.
- Constrain the problem space.
- Consider motivation.
- Make scaffolds explicit for less knowledgeable students.
- Develop conceptual, strategic, and procedural scaffolds.

We might compare this approach with Csikszentmihalyi's concept of flow [31] in the design of games, which has as its goal supporting challenge without frustration.

3.3.2 Situated Cognition

The second cognitive strategy we consider is the theory that learning is "naturally tied to authentic activity, context, and culture" [15]. This theory also suggests that it is more difficult to learn from unnatural activities. For example, learning a foreign language by immersion is

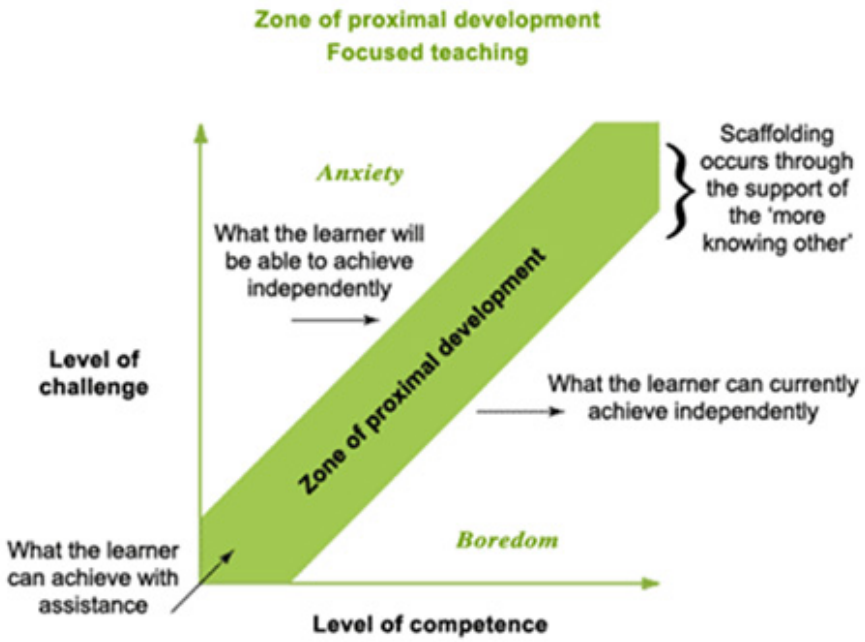


Figure 3.1: Bruner's theory builds on the work done by Vygotsky [153] (interpreting the work of Bruner).

widely held to be easier than learning from textbooks and vocabulary lists [88]. This type of cognition, which we would like our framework to support, is situational: strongly supported by social, cultural and physical contexts.

A related cognitive strategy is situated learning theory, developed by Jean Lave and Etienne Wenger [84] (and owing much to the work of John Dewey and Lev Vygotsky). They believe that students are more likely to learn by actively participating in the process, rather than merely listening to lecturers. Situated learning creates meaning from the real activities of daily living—in an informal cultural setting, through social relationships, and by connecting prior knowledge to new contexts. Lave and Wenger's theory of "communities of practice" holds that learning should not be viewed just as transmission of knowledge, but as an active process. This type of learning is stimulated by specific social and physical contexts, and within specific environments [153]

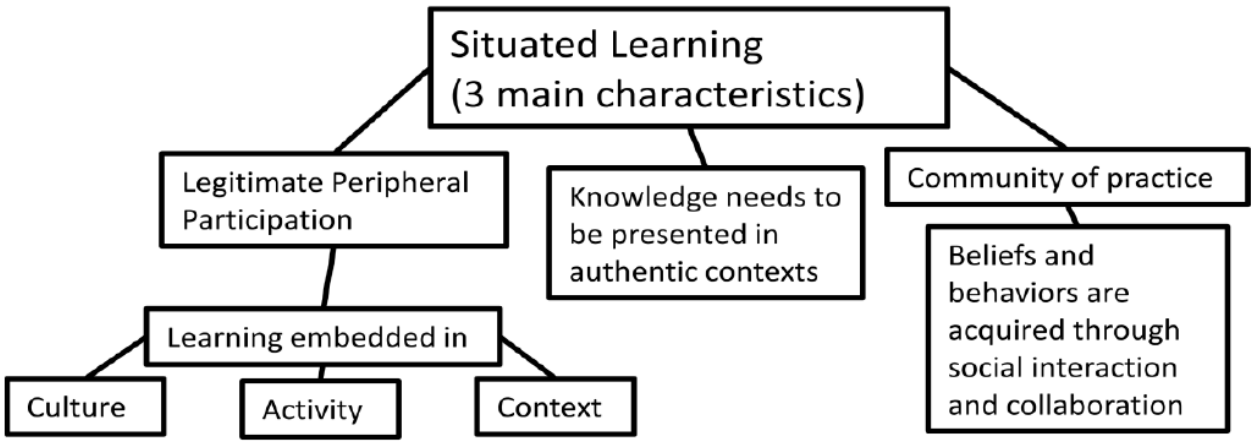


Figure 3.2: Situated learning theory [153] (interpreting the work of Bruner).

3.3.3 Cognitive Tracing

The third cognitive strategy we consider involves learners organizing information and tasks into orders or structures. In the classroom, instructors might inform students of what they have already studied; an interaction diagram can be used to highlight exercises already completed, and those that must still be studied [117].

The concept of tracing can be applied in different ways. When assessing student ability, instructors observe the students' work and apply a technique called *model tracing* to identify which procedural skills they are using. After identifying these skills, instructors employ *knowledge tracing* to assess the students' ability in each. This assessment is based on how many opportunities students have had to apply the skill, and how often they have correctly applied the skill [78].

Another aspect of a cognitive learning is *enactive learning*, in which students learn by doing, and experience the consequences of their movements and actions. This type of learning includes testing learned mental models in an environment that provides feedback based on action. It emphasizes the role of self-modeling in a structured environment, with controls and feedback [59].

Similarly, *formative feedback* is information provided to learners to modify their thinking or behaviour [138]. This type of feedback may include a variety of forms, such as verification of

response accuracy, clarification of correct answers, hints, etc. It can be provided at different times during the learning process: immediately after an answer, for example, or with some delay. It may interact with other variables, such as aspects of the learning task or the personal characteristics of the learner [137]. Formative assessment is strongly assisted by an affective teaching strategy, which creates trust between instructor and learners. As example might be a teacher walking around the students in a class and offering feedback such as: "Good work, but you're not quite there yet." This assessment provides immediate feedback to learners, to help them improve (see Figure 3.3).

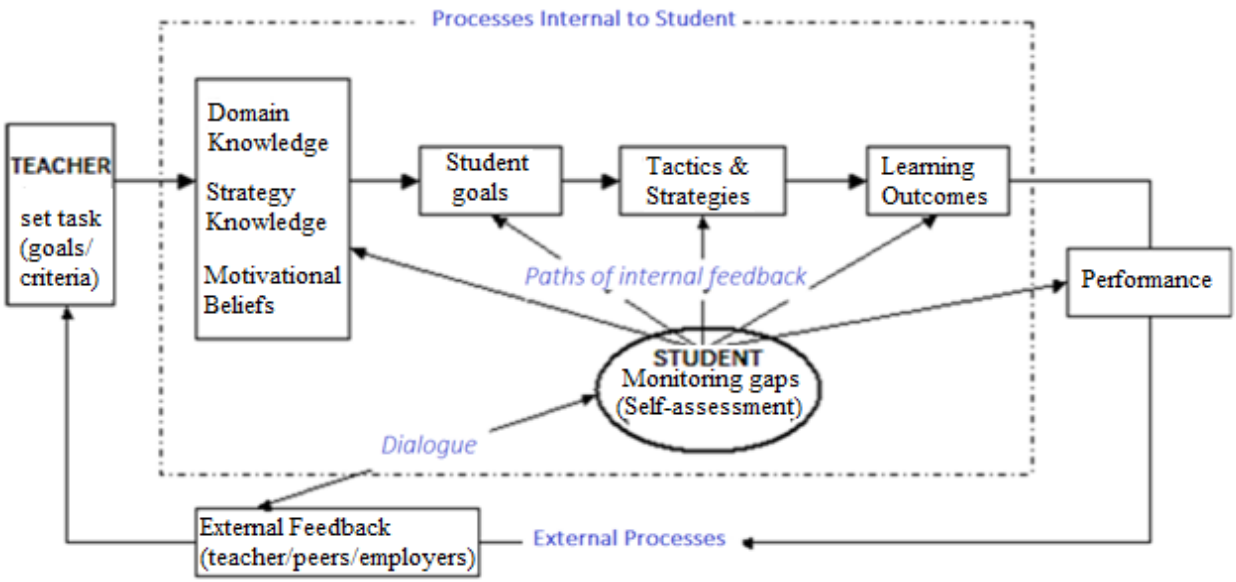


Figure 3.3: A model of the Formative Assessment and Feedback [70].

When designing educational systems, we can increase learners' mental skills by considering the aspects of cognitive strategies: embodied cognition, scaffolding, situated cognition, cognitive tracing, and formative feedback.

3.3.4 Embodied Cognition

The final cognitive strategy we consider is the concept of "embodiment"—that is, the role of the body in learning. This theory holds that a learner's mind, body, and technological environment all interact with one another to enable them to acquire or construct new knowledge [30]. That triad of influences—particularly the physical aspect—makes it possible for

students to learn in goal-directed, real-time environments that engage the senses, perceptions and prior experiences (see Figure 3.4 [76]). The theory of embodied cognition holds that people learn best by performing tasks in a social context, where they interact with objects and with other learners (both experts and novices) to construct their own meanings.

In terms of developing tools, this approach is leads to innovative interaction design. It focuses on understanding the user's needs, knowing cognitive principles, and building them into the product.

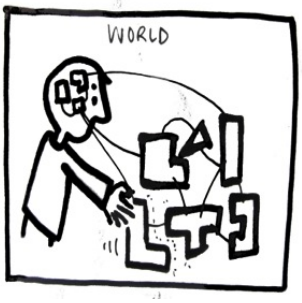
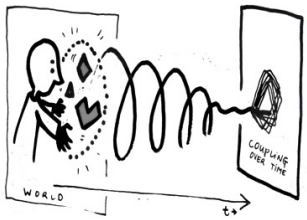
Cognition is Distributed Info-Processing	Cognition is Socially Situated	Cognition is Sensorimotor Coupling
		

Figure 3.4: Embodied Cognition. "Cognition is seen as a temporal coupling between action and perception, sustained through continuous bodily interactions with the environment. Through this process meaning is enacted" [147].

A further aspect of this is *embodied interaction*, a form of technologically supported training created by researchers interested in multimodal learning. This theory holds that our reasoning and behaviour are connected to, or influenced by, our physical and social experiences, plus our interaction with the environment [118]. Activities are designed for learners to interact mentally and physically with technology [1, 77]. Through this process, users build schematic perceptuomotor structures—that is, they create mental connections between the physical actions they perform while solving problems or responding to cues, and the automated sensory feedback the system gives them [1]. Embodied interaction contributes to the meaning-making process in learners, focusing on multiple physical modes—from body posture to gaze to physical action and manipulation.

Paul Dourish [42] has argued that *embodiment*—experiencing physical reality, both actively

(through action) and passively (through perception)—offers a way to explain how we create meaning from our interactions with the world [42]. Our experiences depend on our bodies, both physically and also through our learned and, culturally determined behaviours and experiences. Dourish emphasizes the social as well as the physical aspects of embodied cognition, pointing out that bodily movements and emotional processes are tightly linked. As discussed by Sheets-Johnstone [136], there is "a generative as well as expressive relationship between movement and emotion."

The concept of embodiment can be used in a wide variety of situations. For example, Jack Katz [7, 75] employs it to analyze anger among car drivers in Los Angeles. When we are cut off by another car, he makes the case that anger is produced as a consequence of a loss of embodiment with the car (as part of our body), the road, and the experience of motion. Katz contrasts the social situation of driving, where many cars share the road, with the lack of communication possible between drivers. He even posits that anger is a graceful way to regain embodiment after an incident. This theory offers an elegant account of how people actively produce emotion—both individually and in groups—as part of their social practices.

3.4 Affective Strategies for Learning

In this section we consider the *affective* strategies that can influence learning objectives: these include concepts such as *emotion, humour, personality, ethos* and *pathos*, as well as social and cultural aspects. According to Picard [116], who coined the term affective computing, HCI can be considerably improved when multimodal systems can react to their users' feelings. Emotion may play an important role in such systems, since the truth is that people often tend to resist learning: they get distracted; feel disconnected, bored or fearful; or otherwise fail to engage with the learning process.

3.4.1 Emotion

An important aspect of the learning process, emotion is essentially the brain's information-processing system. It evaluates the world to determine what is good or bad, safe or dangerous;

it operates much faster than cognition, and most of the time takes precedence over it. In a sense, emotion bridges the gap between mind and body. In the learning context, cognitive ability depends on a person's emotional state—a fact that can be leveraged to increase attention and improve memory capacity [52]. It is well known that students' results can be improved with the right encouragement and support.

In pursuit of this goal, the basic approach is to observe a person's behaviour patterns via sensors such as cameras, microphones or pressure sensors attached to the learner's monitor, mouse, chair, keyboard, etc. A sophisticated computer system then associates these patterns with the person's probable affective state [115]. Examples of this type of research include the 2001 work done by Kort et al. on affective learning [79]. Their approach (based on Russell's circumplex model) involved building a "learning companion" designed to keep track of a student's emotional state, and to decide what help they need. They proposed a model of the affective interplay between emotions and learning, shown in Figure 3.5. Their model consists of a four-quadrant learning spiral, in which emotions change as learners progress in a spiral motion through the quadrants.

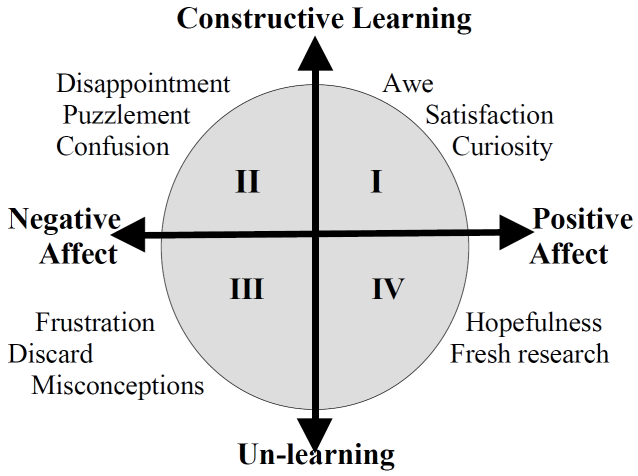


Figure 3.5: Kort et al.'s model relating phases of learning to emotions. [79].

Another application, from Picard's group, is a chair with pressure sensors that analyzes posture: the team has identified nine of these that they use to associate students' interest levels with their affective states. Both these systems assess learners' states, and decide what input they need [63, 97].

Figure 3.6 illustrates how the system works. Emotional sensors include a seat device to analyze posture: it detects the learner's movements relative to the computer screen. A conductance bracelet measures skin conductivity, which is known to correlate with emotional arousal. A camera mounted on the monitor, together with a facial expression sensor, detects movements of eyes and mouth [52], and identifies states such as interest, reflectiveness or uncertainty. A mouse measures the learner's global pressure as they manipulate it, and sensors on the chair back and seat measure blood pressure. In addition, sensor devices worn by the learner—such as finger clips and rings, a waist belt, and a neck pendant EEG—detect the physiological signs of emotions.

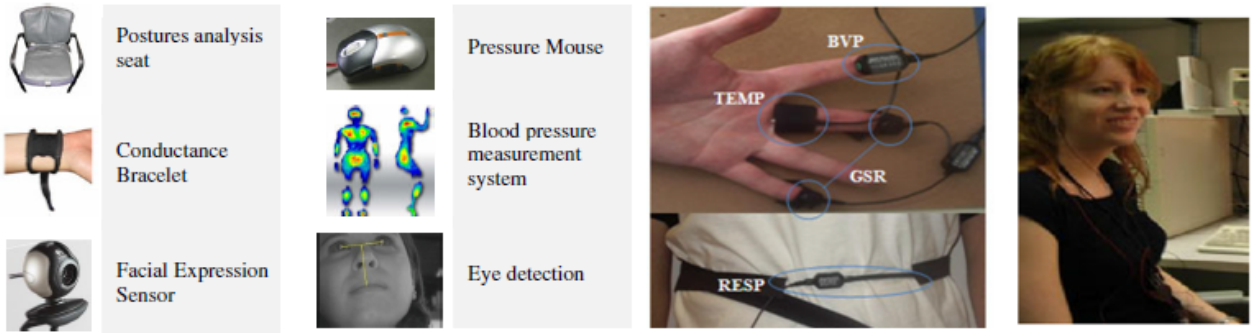


Figure 3.6: Left side: Emotional sensors, Middle: Physiological devices for emotion detection, Right side: A learner wearing Pendant EEG [52].

3.4.2 Humour

Another paradigm of an affective strategy, humour plays a very important role in human-to-human interaction (HHI). In the educational context, instructors often use humour to create a positive and energized atmosphere that's conducive to learning: it captures students' attention, engages their interest, and promotes critical thinking and creativity [122]. It can also resolve any problems or conflicts in the teacher-student relationship [96].

Humour is a social phenomenon: we rarely laugh when we are alone. And learning is a social activity, taking place through the interaction of individuals [23]. According to Vandaele [149], the three main theories to explain humour are:

- **Incongruity:** we laugh at things that are unexpected or surprising.

- **Superiority:** we laugh at the misfortunes of others, from an innate sense that we are better than them.
- **Relief:** we laugh because of a release of nervous energy.

In learning situations, the essential function of humour is to relieve stress, which in turn diminishes students' feelings of anxiety, frustration, hostility or anger. Dion [39] found that second-language teachers often tend to use humour to motivate students, help to build language skills, improve class atmosphere, and generally create a more satisfying learning experience [39]. According to Dormann and Biddle, many educators believe that humour creates a more pleasant learning environment by increasing interaction, diminishing pressure, and making the teacher appear more approachable and supportive [41]. Coleman [21] mentioned that the use of humour makes learning materials more effective.

Humour is not limited to human-human social communication: interactive learning environments can also display this attribute, most often through animated interface agents. Mishra and Hershey [96] describe pedagogical software that draws on HHI scripts by embodying human characteristics such as the use of gestures and facial expressions [96]. Other advances include the software created at the Center for Advanced Research in Technology for Education, at the University of Southern California. They developed interactive pedagogical agents that exhibit politeness, expressiveness and empathy. Figure 3.7 shows an agent known as STEVE (Soar Training Expert for Virtual Environments) that focused on issues such as identifying a student's cognitive and affective states, tracking learner-agent interaction as a social relationship, and managing interaction to improve communication effectiveness [96].

3.5 Multi-Sensory Modalities

This section discusses how learning objectives, cognitive strategies and affective strategies relate to multiple sensory modalities. The idea is that the multimodal system will support the affective and cognitive strategies, connecting those functions to physical movements.

As the field of HCI advances in sophistication, multimodal systems are increasingly using more natural "human" multi-sensory interaction techniques such as expressions, gestures,

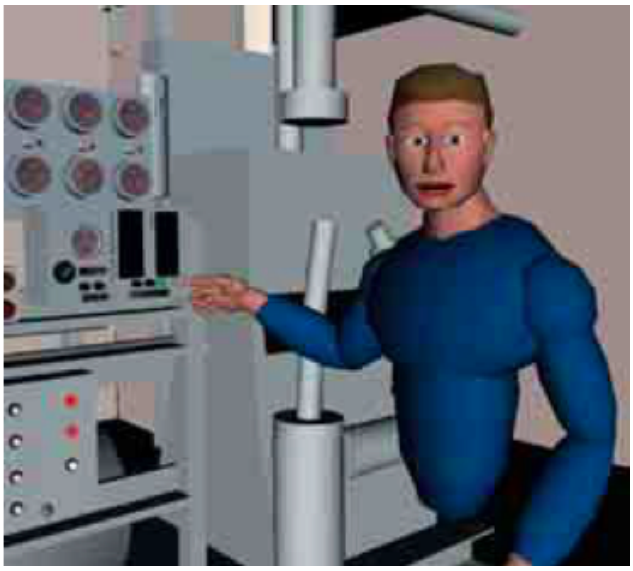


Figure 3.7: STEVE (Soar Training Expert for Virtual Environments) developed by the Center for Advanced Research in Technology for Education (CARTE) [96].

voices, and even body movements [144]. The term "natural" here is intended to refer to modalities with which humans are already familiar. These and other affective and cognitive strategies can help instructors help students to learn through more than one sense. The input modes of some computers correspond to the human senses: cameras (sight), haptic sensors and motion input (touch), microphones (hearing), olfactory (smell), and even taste. But the most common input devices—keyboard, mouse, writing tablet—correspond to no particular human senses.

These multi-sensory devices connect users to the learning process in a way that cognitively and emotionally involves them. Lin [86] has shown that the human brain serves as the ultimate centre for managing bodily, cognitive, and emotional input, and their interaction with one another in the process of task performance. Figure 3.8, shows Lin's model [86], illustrating the relationship between cognitive, affective and body sensory elements.

The goal of affective feedback is to guide learners toward emotional states that promote motivation and are conducive to learning. In cases where learners are off-task, confused, unmotivated, or have a low estimation of their ability, cognitive supports may be useful. Such supports involve techniques—which may be both overt and covert—for prompting learner to take desirable actions. Techniques can take several forms, such as conversations between

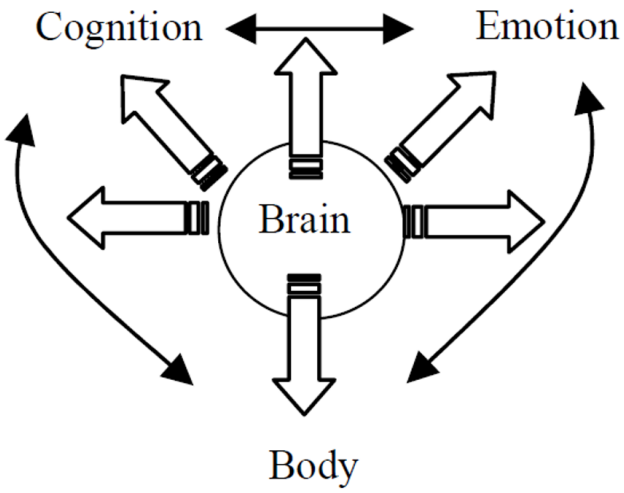


Figure 3.8: Brain, cognitive, body, affective, and their conceptual relationship [86].

instructor and learner while the latter is analyzed by tools such as a facial expression sensor, pressure mouse, blood pressure measurement system, eye detection device, posture-analysis seat, etc. For example, Picard [115] has designed a system to automatically recognize a child's interest level in natural learning situations by using chair-pressure information and Tekscan pressure arrays.

3.6 Quasi-Sensory Modalities

Quasi-sensory modalities differ from multi-sensory modalities in that are not strictly based on the physical senses. In this section, we discuss two modalities: narrative-based HCI and persuasive technology.

3.6.1 Narrative-Based HCI

Narrative, or storytelling, is one of the oldest human activities. Even (perhaps especially) primitive peoples and children use stories to relive past events, and to imagine new possibilities. The latter is a particularly effective way to creatively visualize good ideas and better outcomes, and to share them with others. Like humour, narrative is an intrinsically social activity.

In terms of affective learning, a narrative approach can help to develop an educational application that motivates students to engage with learning tasks, and keeps them immersed in the process. Harnessing the power of storytelling allows educational materials to be more attractive to learners. Narrative-based learning might include elements such as empathy and role-playing, exploration, reflection, and sharing ideas. In terms of interaction design to affect behaviour, a narrative structure—combined with the use of multimodal system tools—is perfectly suited for the task of motivating learners. Interaction design narratives can be as simple as showing some slides and creating a story from them; or they may be as attractive and compelling as movie storyboards [26].

Good narratives build trust between learner and software, and facilitate user stories based on positive experiences [139]. For example, Rowe et al. [126] developed the concept of "narrative scaffolding" when discussing cognitive and affective support in the Crystal Island game (as shown in Figure 3.9). Built on Valve Software's Source engine, this 3-D game platform for Half-Life 2 is designed to make learning engaging through a series of tailored challenges. The authors conclude that the narrative-centred learning environment—designed as a supplement to classroom instruction—is effective and engaging, and offers promising results [127].

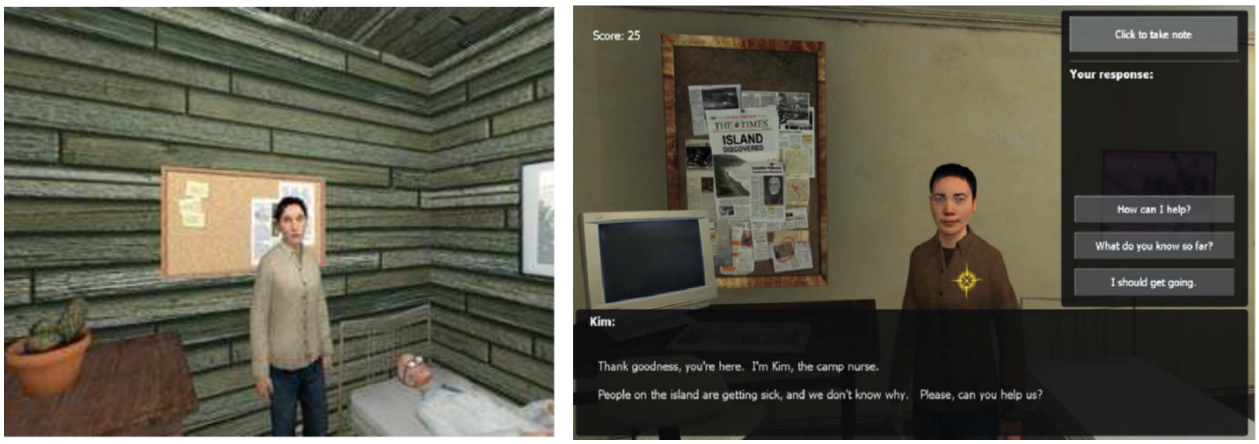


Figure 3.9: Conversation with Crystal Island's camp nurse [127].

Another example of narrative-based learning is a mobile storytelling application called StoryKit (see Figure 3.10). The app integrates multimodal media-capture tools to allow children to create, edit and share personal multimedia stories [12].

Figure 3.10: StoryKit extends the mobile reading application from the International Children's Digital Library to integrate story creation, editing and sharing features [13].

3.6.2 Persuasive Technology

The fine art of persuasion—that is, a gentle, non-coercive attempt to influence a person's behaviour, feelings or thoughts—is one that has been discussed for thousands of years, ever since classical times. In the ancient Greek schools of philosophy, Aristotle taught that rhetoric is the skill of knowing which means of persuasion to use. He described three main forms of persuasion: *logos*, *ethos* and *pathos*, illustrated in Figure 3.11.

- **Logos** is the logical component of an appeal or argument, a way of persuading an audience by reasoning, facts, figures, statistics, charts, etc.
- **Ethos** is an appeal to ethics—convincing an audience that the proposed course of action is the right thing to do; and also convincing them of the moral character, authority or professional credibility of the persuader.
- **Pathos** is an appeal to an audience's emotions, convincing them of the rightness of an argument by creating an emotional response (such as telling a tearjerker story to sway

people's judgment).

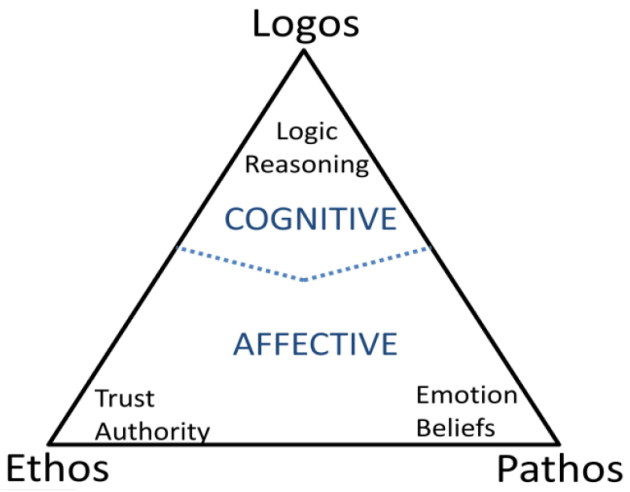


Figure 3.11: Modes of persuasion proposed model.

Like narrative-based HCI, persuasive technology is another quasi-sensory modality that can be used to motivate learners., In computer terms, persuasion is much the same thing as in human terms: any interactive system that's designed to (gently) influence users' attitudes or behaviours [49]. In the context of multimodal system tools for affective education, it may include wearable sensors and mobile interfaces [99]. Others might be text-message reminders to learners, to keep them engaged with the classroom environment; or applications designed to offer feedback during the teaching process. The goal of all these technologies is to persuade participants to proactively engage with their learning, rather than relying only on lecture notes.

(Apart from the educational sphere, much research has been done on persuasive applications in the field of health. These include using mobile phones to monitor the health status of patients wearing physiological sensors; or using game applications to motivate people to engage in healthy activities [24, 134]. An example of the latter is MoviPill, a game-style application designed by Oliveira [36] to encourage users to take their medication at the right time (see Figure 3.12). The closer a person's log-in is to the prescription time set by their doctor, the more points they get.) The concept of persuasion always implies an intent to change attitudes or behaviours. Fogg (1998) [50] defines three types of intent: *endogenous* (from within), *exogenous* (from external factors) and *autogenous* (self-produced). According

Figure 3.12: Example of points and emoticon assignments for a patient that takes two doses per day of a certain medication (8am/pm) [36].

to Fogg, a computer technology exhibits:

- **endogenous** intent: when a designer creates software with the intent of persuading users to some course of feeling or action—for example, a game designed to coax learners into understanding course material.

- **exogenous** intent: when one person (an instructor) gives the technology to another (a student) in the hope of changing that person's attitude or behaviour—such as becoming more engaged with course material.

- **autogenous** intent: when a person chooses to adopt a technology in order to change their *own* attitudes or behaviour: they may decide to use a persuasive technology device to help change their learning style [50].

Fogg (1998) [50] also identifies three forms of persuasive technologies, following the standard classification of the three ways that people view computer technology: as a *tool*, a *medium*,

or a *social actor* [50]. In our context, multimodal software and systems have the following attributes.

- As **tools**, they give users new abilities and powers, allowing them to do things they could not do before (or at least accomplish tasks more easily).
- As **media**, they can share content that's either symbolic (text, data graphs, icons, etc.) or sensory (video, games, virtual environments, etc.).
- As **social actors**, they can help users to respond better to computers through the use of interactive technologies with animated characteristics. These might include computers that have simulated physical features, emotions or speech; ones that play roles, such as coach, assistant, opponent or pet; or ones that follow human social conventions, such as giving greetings or apologies, or offering to take turns [50]. (This aspect relates to the use of agents, as discussed in Section 3.4.)

3.7 Summary

In this chapter we reviewed the theoretical background for multimodal affective education. We began with Bloom's Taxonomy, and its identification of cognitive and affective elements. We then described several affective and cognitive educational strategies, as well as the multi-sensory and quasi-sensory modalities that help to implement learning objectives. In the next chapter, we describe our framework for design and evaluation, based on this background.

Chapter 4

MADE Framework: Multimodal Affect Design and Evaluation

4.1 Introduction

In this chapter we develop a framework, *Multimodal Affect for Design and Evaluation*, that we call *MADE*, for multimodal affective education systems. We begin by recalling the important elements from the previous chapter on cognition and affect in education, and how they can inform the understanding of multimodal systems. We then describe a process that highlights how the elements are connected, supporting design and evaluation, as well as explaining the stakeholders involved in the process, and the roles they play. Finally, we conclude with a discussion of the potential benefits of our framework and outline of the work that follows in the later chapters.

4.2 Learning Objectives and Strategies

Much thinking on education emphasizes the primary importance of the educational objective. For example, Benjamin Bloom, in the introduction to his taxonomy, says the following:

> "By educational objectives, we mean explicit formulations of the ways in which students are expected to be changed by the educative process. That is, the ways in which they will change in their thinking, their feelings, and their actions." [10]

A more straightforward definition is: "A learning objective is a statement that describes what the learner will be able to do upon completion of the learning experience." [93] Bloom's Taxonomy was an effort to identify and organize many aspects involved in learning objectives. As we discussed in Chapter 3, this included both cognitive and affective aspects.

In practical terms, Bloom's taxonomy helps teachers articulate specific objectives for courses, lessons, textbook chapters, or other kinds of teaching materials.

In Chapter 3, after discussion of Bloom's Taxonomy, we showed how both the cognitive and affective domains can be supported by specific teaching strategies. The factors discussed in the previous chapter that influence our framework are illustrated in Figure 4.1.

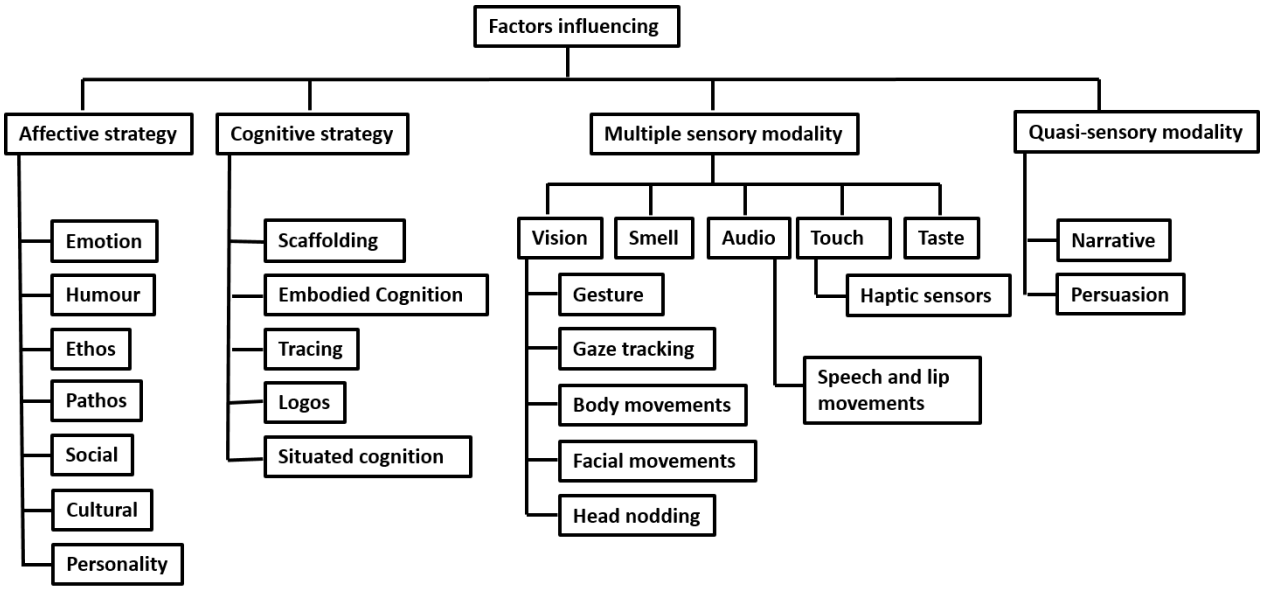


Figure 4.1: Influencing factors in the proposed framework.

While the objectives identify *what* is intended, the strategies identify *how* these objectives might be achieved.

Figure 4.2 illustrates the relationship between learning objective with cognitive strategy. It identifies four types of cognitive strategy discussed earlier: scaffolding, logos, tracing, embodied cognition, and situated cognition, as we explained in Section 3.3.

The next diagram (Figure 4.3) provides the relationship between learning objective with affective strategies including emotion, humor, ethos, pathos, social, cultural and personality. We again explained each strategy earlier in Section 3.4.

Our interest is in multimodality and affect to support education. As discussed in the previous chapter, the proposition is that multimodality can support both cognitive and affective strategies. This was illustrated with several examples in Section 2.6, and discussed in Section

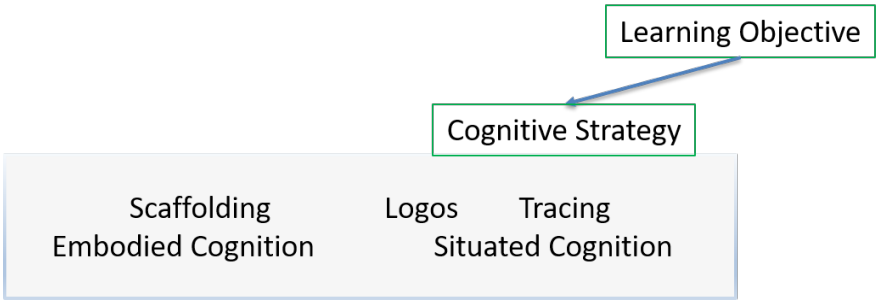


Figure 4.2: Learning objective and cognitive strategies.

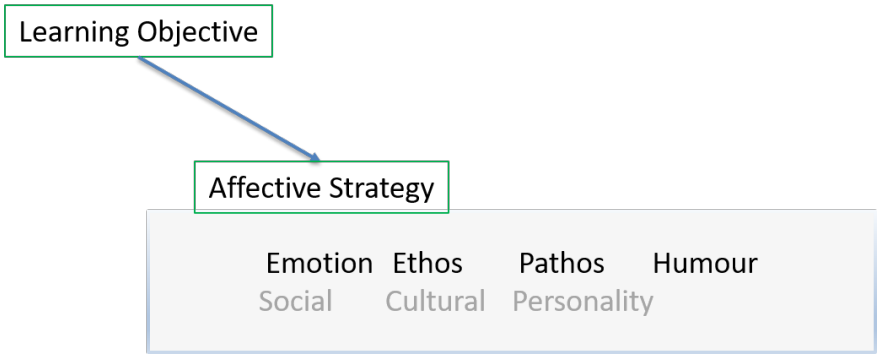


Figure 4.3: Learning objective and affective strategy.

3.5.

In education, affective and cognitive strategies work together. For instance, consider software to help students learn about the effects of environmental pollution. Rather than merely presenting facts and figures, educational material has more of an impact if it can shown visually, with images and animation, which is a cognitive strategy. This will highlight how only a small amount of contaminant in a large body of water to cause all kinds of harmful effects to the people who live in nearby towns, and the animals that live around it.

Children may take more of a personal interest in such a problem if an affective strategy is also used. For example, they might be encouraged to identify with the people affected—whether the water pollution takes place in California or in Ottawa. Another affective element that can help children engage with the learning task is role-playing: they may act the part of a pollution control officer, or a police investigator.

There is a connection between the body and both cognition and affect, as proposed by

Lin [86]: "...the brain serves as an ultimate center to manage the body, cognition, emotion, and their interaction." We identify the role of multimodality as augmenting cognitive and affective strategies, as shown in Figure 4.4.

We also include in this the quasi-sensory modalities that we discussed in Section 3.6, such as narrative or persuasion. Similar to multi-sensory modalities, these also influence both cognitive and affective strategies.

Together with the elements discussed above, our initial framework is therefore as illustrated in Figure 4.5. We choose this as the initial element in our framework. For educational

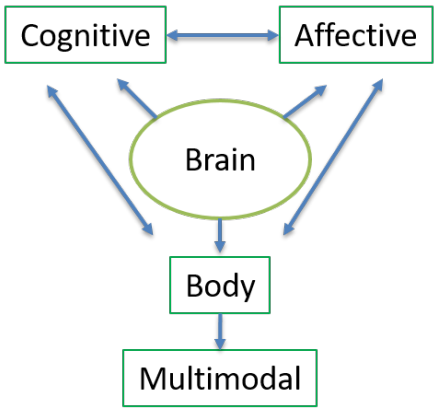


Figure 4.4: Brain, cognitive, body, affective, multiple sensory modalities and their conceptual relationship.

systems, the same principles apply as for other teaching materials. We should therefore expect articulation of the learning objective for which the educational system is created. Moreover, we should then expect selection of cognitive and affective strategies to support that objective.

4.3 Stakeholders and Roles

The principal stakeholders in education are students and teachers. Teachers aim to help students learn, and to do this they identify learning objectives and strategies: cognitive, affective, and multimodal. Any educational computer system should support this process.

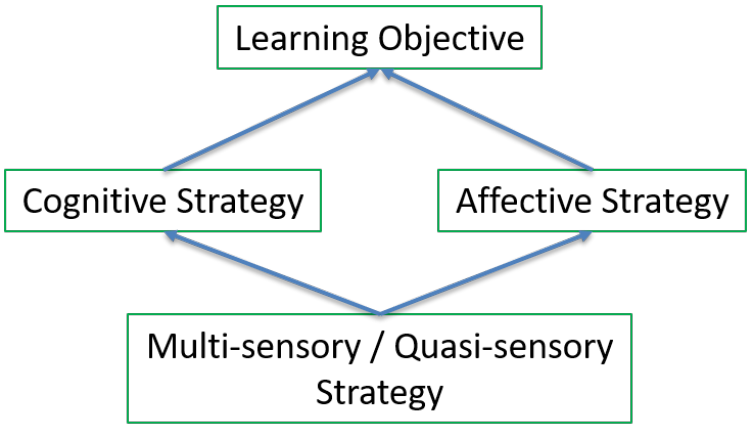


Figure 4.5: Learning objective, cognitive, affective strategies and multiple sensory modalities.

But should the computer system be designed to *replace* the teacher, or to *help* the teacher? And what is the role of the system designer?

There are several aspects to a teacher's work. For example, the first aspect is typically identifying the specific learning objective for the students. Then the teacher must select both a cognitive strategy (e.g. how to explain the study material) and an affective strategy (e.g. how to make students enthusiastic about learning). Next, teachers must create a lesson plan of activities for the students to undertake. Teachers then facilitate these activities, monitor students, and provide formative and summative feedback. Later, teachers may reflect back on the learning experience; and review the course to consider what might be improved.

A single teacher might perform all these roles, from choosing a learning objective all the way through to summative feedback and reflecting on the process. More commonly, however, an individual teacher will take on some roles and rely on other teachers for other aspects of the process. For example, some teachers might focus on determining learning objectives, others on preparing activities, and other still on facilitation and feedback. Where this happens, however, it is important that all roles understand the joint aim. So, for example, where an individual teacher has a focus primarily on facilitation and feedback, they must also understand the learning objective, strategies, and activities.

Our primary interest is in supporting affective education. One way for a system to do this by using sensor input and techniques from Artificial Intelligence. This might address or *replace*

the teacher's role by automatically monitoring students to provide facilitation and feedback. This approach is shown by Kapoor and Picard [73], who emphasize how a system can adapt automatically to learner emotion, using sensor input and artificial intelligence. They describe the architecture of a proposed system, shown in Figure 4.6. They focus on a scenario where children try to solve puzzles on a computer. The system first senses the children's faces, gestures and postures. It then extracts these features and passes on all the information to a multimodal pattern analyzer, which combines it to identify the users' current affective state.

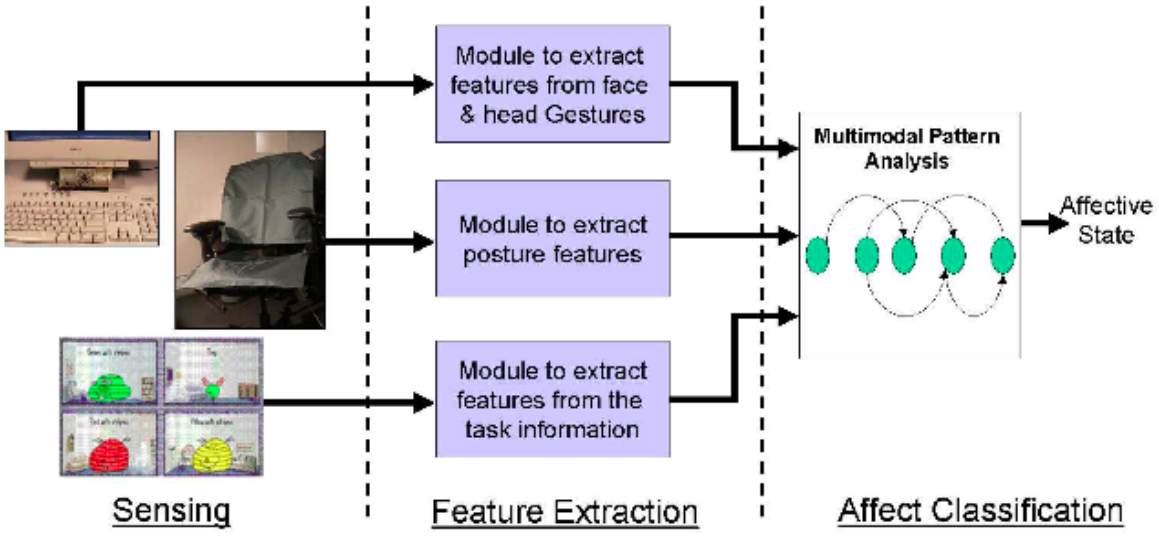


Figure 4.6: Architecture to predict affective state [73].

Our approach involves user interactive design and techniques from HCI. We focus on how the interface and interaction should be designed to support and *help* a teacher's affective and multimodal strategies, rather than replace the teacher. In taking this approach, we most explicitly support the teacher role of designing activities, which in our case will be supported by a system. However, we also implicitly support other teacher roles, as long as there is understanding of the inter-dependence involved.

Our framework therefore involves the system designer working to create a system that supports the teacher's affective strategy, to further their learning objective for the student.

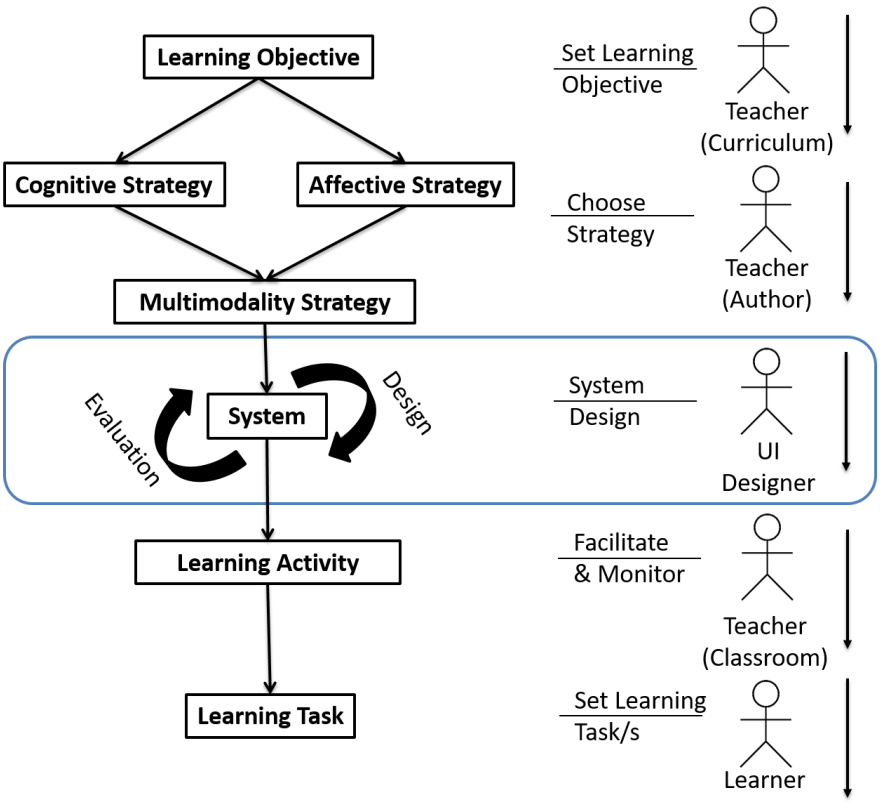


Figure 4.7: The MADE Framework showing the role of the system in relationship to the teacher.

4.4 System Design

The system design goal is always to support a sensory modality that includes an affective strategy as well as a cognitive strategy—the latter informed by what the specific task is, since some strategies are best suited to different kinds of learning objectives. Our initial focus is on the affective side.

It is well-established that interaction design is an iterative process involves evaluation and refinement. This is documented in ISO Standard 9241 [69], and illustrated in Figure 5.3. Accordingly, we study both HCI techniques for design, in Chapter 5, and also for evaluation in Chapter 6.

To adapt HCI interaction design techniques for affective education, we need to understand how processes involved in affective education. We are especially influenced by the work

of Kort et al. [79], discussed in Chapter 3. The process starts with positive affect, becomes challenging with negative affect, and then becomes positive again as challenges are overcome. We suggest this sequence will be helpful in designing or evaluating affective educational software.

For our own proposed model of the affective and cognitive dynamics of the learning process, shown in Figure 4.8, we used Kort et al.'s model as a basis. We also consider the circumplex model of emotion, first proposed by James Russell in 1980 [129] and based on the two dimensions of arousal and valence.

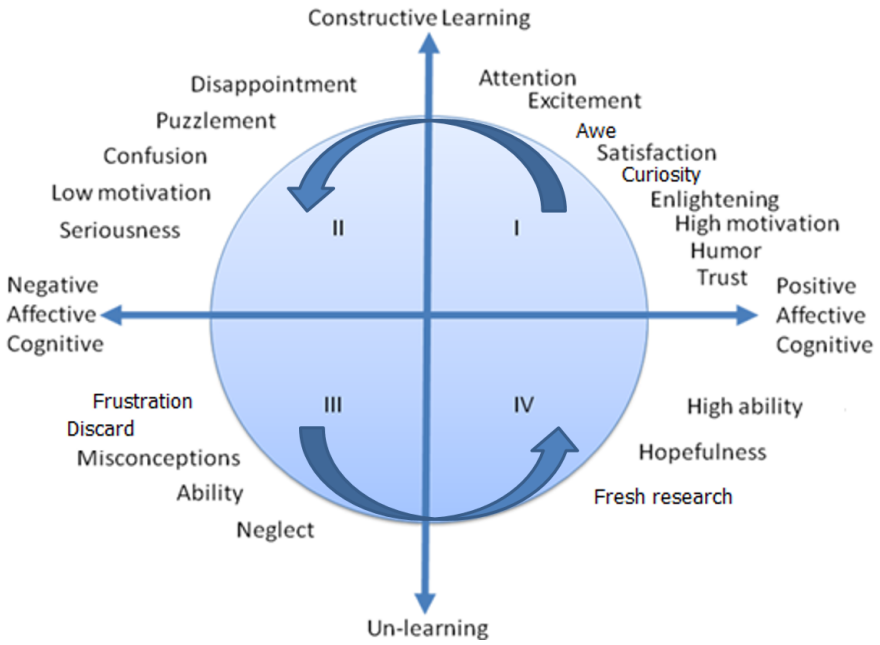


Figure 4.8: Proposed affective cognitive model relating phases of learning.

4.5 The MADE Framework

Our framework is based on the three elements, as described above. First, we consider the learning objective, and the cognitive, affective, and multimodal strategies support it. Second, we identify the stakeholder roles involved. Third, we address interaction design and evaluation. This structure is illustrated in Figure 4.7.

The real-world application of the MADE framework is to support educational system designers to create and evaluate systems to support teachers and learners. Such systems should enjoy the benefits of multimodality and affect: reduction in learners' cognitive load, improve communication by the use of richer interfaces, and model human sensory perception and communication patterns. The system's affective feedback should support the teacher's efforts—by, for example, proving activities and feedback based on the interplay between affect and cognition in learning. As well, the system might provide that information to teachers instead, prompting them to come and check on the student. In an educational environment without technology, teachers might not even notice if a student needs help; their problem might go unnoticed. Our multimodal system might help teachers to be more alert to struggling students.

4.6 Summary

In this chapter we introduced the Multimodal Affect for Design and Evaluation (MADE) framework—our proposed framework structure for designing and evaluating affective multimodal education systems. We described its usefulness for the task by considering affective and cognitive strategies, as well as the multi-sensory and quasi-sensory modalities that help to implement learning objectives. Our goal is for the interactions between these elements to guide design and development of a e-learning environment — one that promotes affective engagement, and multimodal strategies for learning.

In particular, we propose that the learning objective and cognitive and affective strategies are within the education domain. They should inform the software design, which should then support the role of the teacher in facilitating activities. The main benefit is to clarify the role of the designer identifying their purpose and what they need to begin work. We suggest that the nature of their work is interaction design.

In the following chapters, we explore how conventional interaction design techniques might be adapted for this work. Chapter 5 addresses design methodologies, and Chapter 6 addresses evaluation techniques. The later chapters present case studies of our adapted methodologies and techniques.

Chapter 5

Using the Framework for Design

The main idea in this chapter is to adapt well-known user interaction design methods to the MADE framework. For example, system designers have to work with teachers regarding the learning objectives, and appreciate what affective and cognitive strategies they might use to get those learning objectives. (This relates to a teacher who selects strategy, and not necessary to teachers who later facilitate and monitor students.) Our goal is then to address what multimedia systems might support those affective strategies and explore how they could be build into a system. We will explain in detail in the next sections, discussing two of the most influential design methodologies including: *User-Centered Systems Design*, *Scenario-Based Design*, *Goal-Directed Design* and *Usage-Centered Design*.

We illustrate the nature of the challenge with the *MADE strategy diagram* in Figure 5.1. The teacher, the student and the developer are associated with the use cases. The teacher has to come up with and decide the learning objective, the cognitive strategy and the affective strategy. The teacher has a monitoring role with the system as well. The developer's job is to take this model into the system. In the figure, the cloud is kind of a conceptual model for the actual system. Therefore, the top layer is a kind of a cloud that has a general learning objective, affective strategy, cognitive strategy and the multiple sensory modalities involved.

Figure 5.1 is the general version of the diagram, but we need to inform design of a specific system. We describe a specific version of the strategy diagram considering the work of Abrahamson and Trninic [1]; a mathematical learning system, (Figure 5.2). They explained an embodied-interaction design framework for mathematical concepts, and demonstrated a mathematical imagery trainer. They utilize new kinds of multimodal interaction for the concept of proportionality. In this activity the learners who are grade four, five, and six, pupils, have to make and keep the screen green by moving their right and left hands with

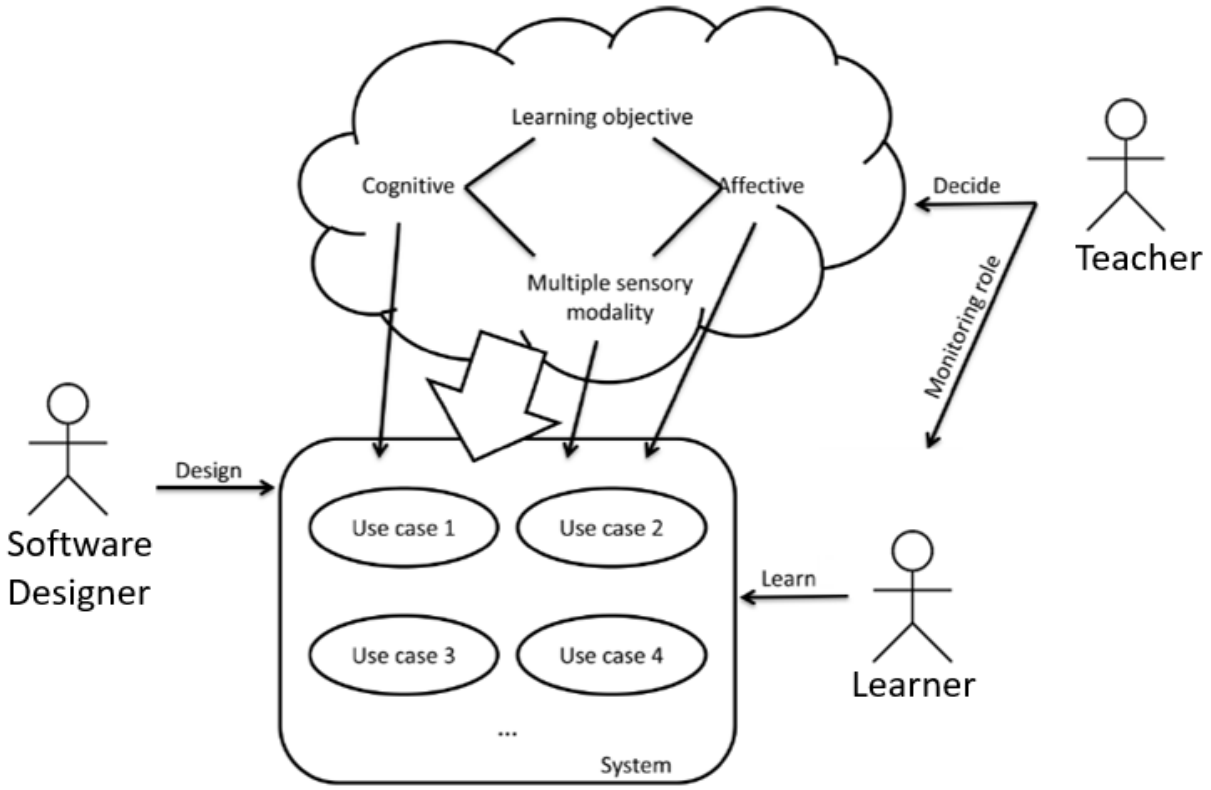


Figure 5.1: MADE Use Case Diagram.

regard to a specific ratio and distance between their hands. We have taken this concept and relate it to our framework.

The first and second use cases are "Move hand to correct ratio" and "learner move hands, keeping bigger the distance" (Cognitive strategies). The affective strategy and the multiple sensory modality is "Play happy music when correct ratio is determined" and "Using haptic device to keep the correct distance".

5.1 Survey of Design Methodologies

There are many general HCI design methodologies. In this section we introduce some of them that have been influential.

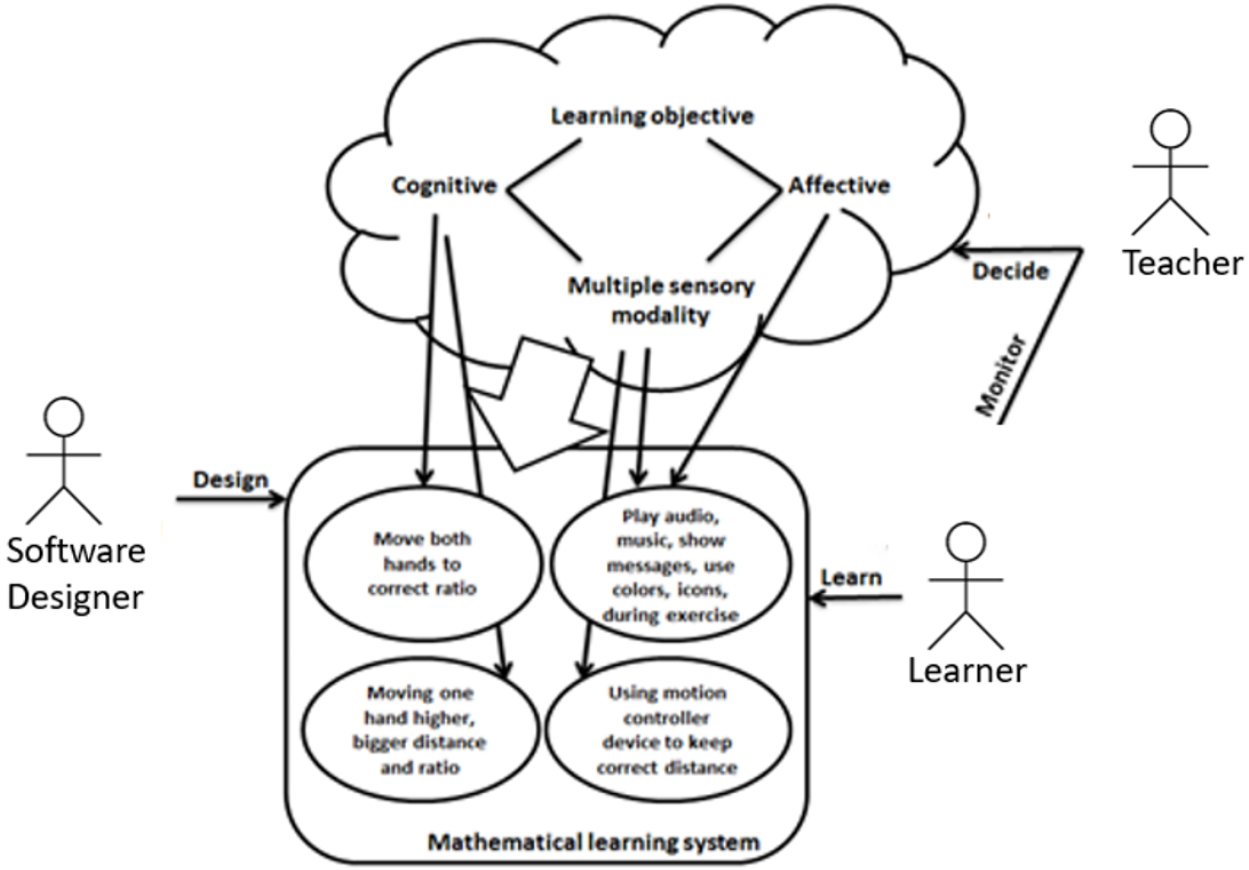


Figure 5.2: MADE Use Case Diagram specific to the mathematical learning system.

User-Centered Systems Design: One of the most commonly used design methodologies is *User-Centered Systems Design* (UCSD or UCD), which is the basis of ISO Standard 9241: Human-Centred Design Processes for Interactive Systems. UCD is a process focusing on usability throughout the entire development process and further throughout the system life cycle. It is a methodology for designing usable applications and an interactive software that meets the needs of its users throughout the iterative process (see Figure 5.3).

Neuwirth et al. [101] outlined in detail the twelve UCD key principles, which include: user focus, active user involvement, evaluating systems development, simple design representations, prototyping, evaluate use in a context, explicit and conscious design activities, a professional attitude, usability champion, holistic design, process customization, and a user-centered attitude. UCD methods are commonly applied to not what functions will be delivered but to *how* functions are delivered, understanding how people think to develop technologies designed to be easily usable [58, 112].

Scenario-Based Design: Another long-standing design methodology is Scenario-Based Design (SBD). Carroll and Rosson's use of SBD focuses on describing how users accomplish tasks [123]. Scenarios serve as a central representation throughout the development cycle, first describing the goals and concerns of current use, and then being successively transformed and refined through an iterative design and evaluation process (See Figure 5.4).

SBD elaborates the design and then proposes design solutions for the problems. Scenarios have characteristic elements, such as actors, actions, task goals, user evaluation and events. Actors are the users that are interacting with the system. Actions are the observable user behavior. Task goals are effects on the situation that motivate actions carried out by the actors. User Evaluation is the mental activity directed at interpreting features of the situation. Events are the external actions or reactions produced by the computer, or other features of the settings; some of these may be hidden to the actors, but important to the scenario [123].

We start with a problem scenario step, which is writing a problem scenario about the current situation and analyzing the requirements. The next step is writing activity scenarios, what and how an activity is going to address a problem is the focus of activity scenarios, which design team presents concrete ideas of new functionality and new ways of thinking

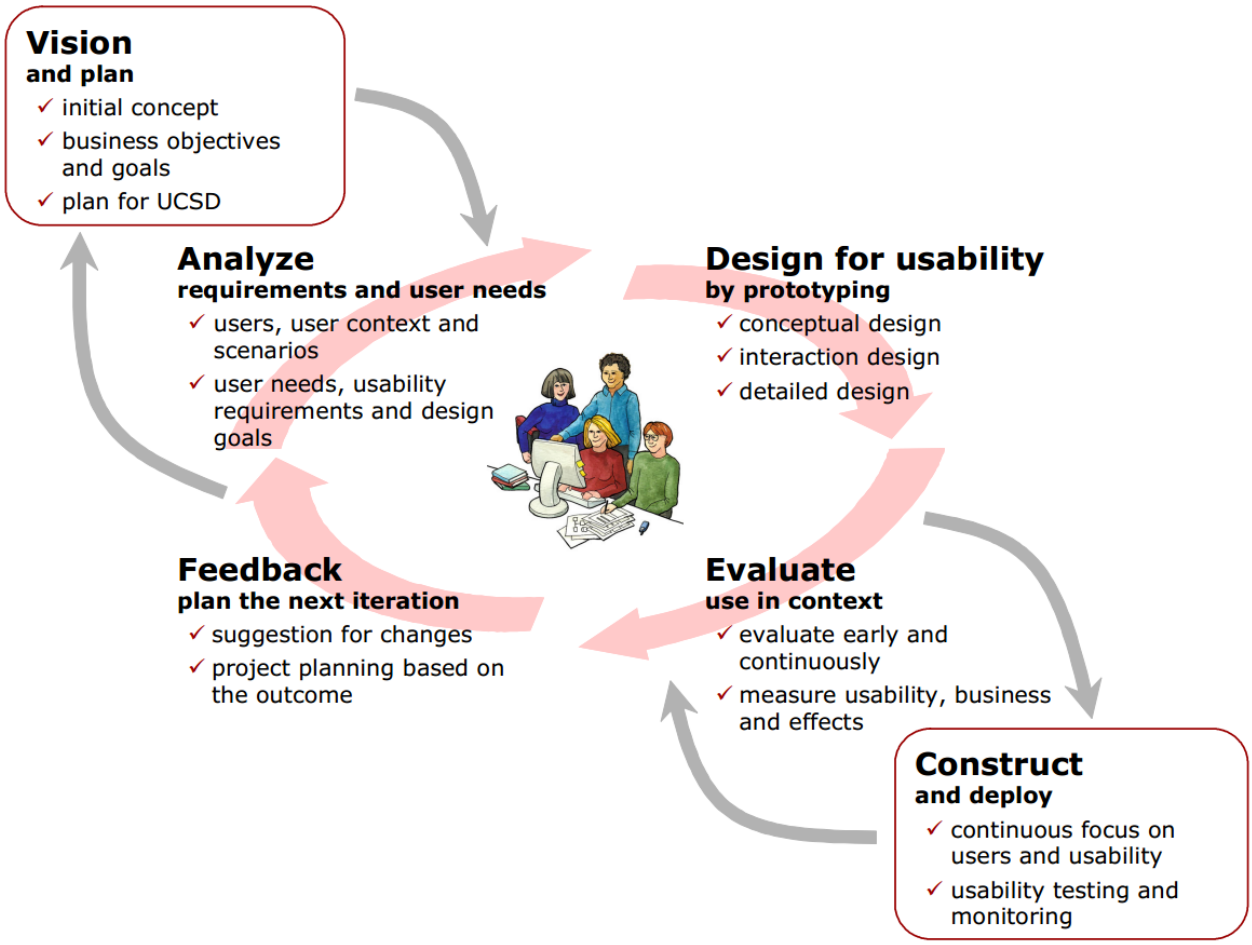


Figure 5.3: Key principles for system design [57].

about users' needs, and next is the information and interaction design scenarios specifying representations of a task's objects and actions that will help make users perceive, interpret and make sense of what is happening [107].

Goal-Directed Design: Cooper et al. [26] created an interaction design methodology known as *"Goal-Directed Design"* (GDD), the one that includes *"Personas"* as practical interaction design tools to create high-tech products (see Figure 5.6). Personas represent the people who will use the system. Cooper's GDD process results in a solid user model and a comprehensive user plan. It is a powerful tool for answering questions such as: who are the users, what are they trying to accomplish, or how users interact with the system and how the system should behave and deal with problems the users encounter. To create a system

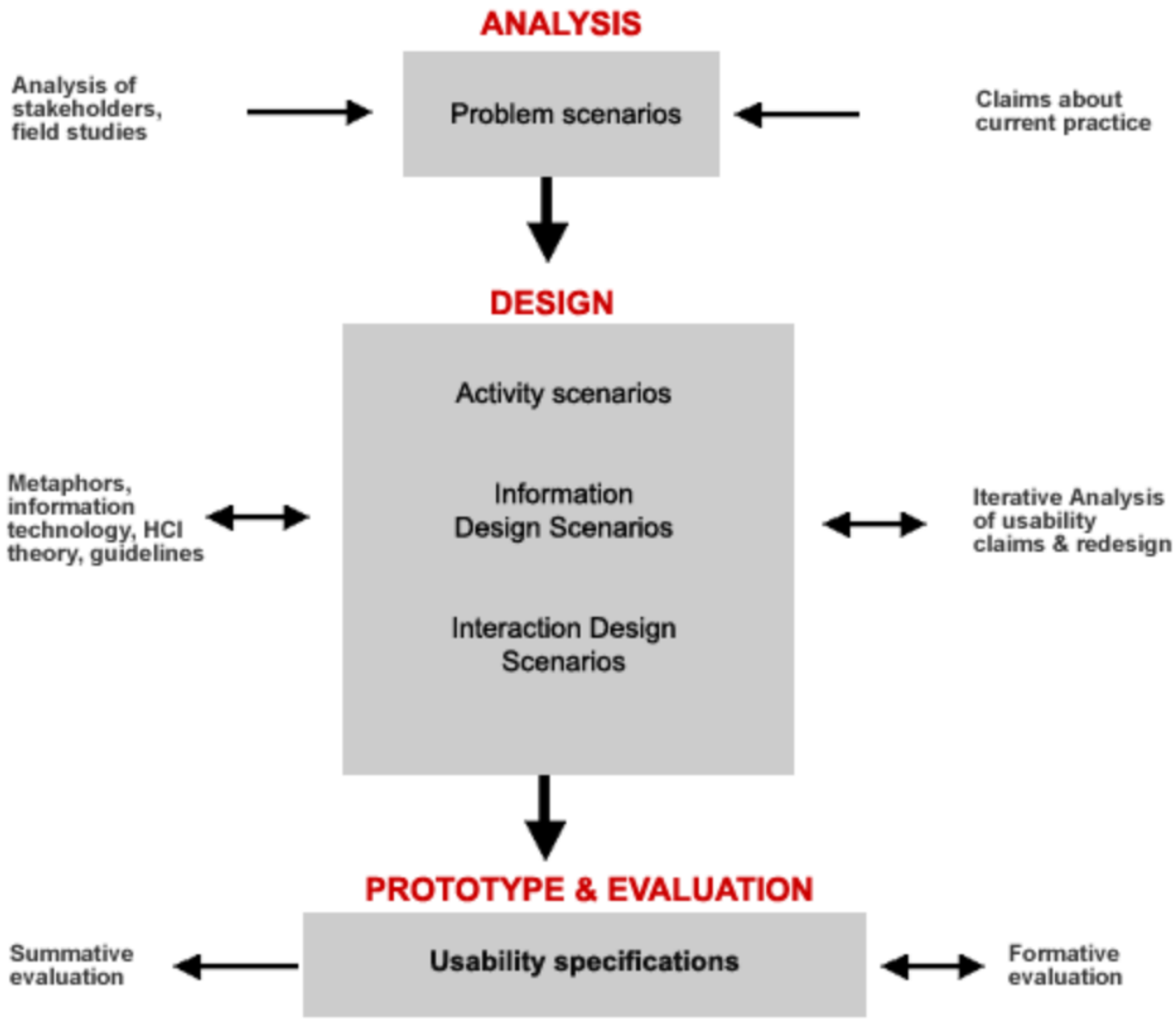


Figure 5.4: An overview of the Scenario-Based Design framework [124].

that must satisfy a variety of users by using personas and designing for specific types of individuals with specific needs. Personas are user models that represent a class or type of user of a specific interactive system. We suggest personas to show concrete representations of the different users of the multimodal system.

In persona-based scenarios we have concise narrative descriptions of one or more personas using a system to achieve specific goals. Here we start the design from a story describing an ideal experience from the persona's perspective, focusing on users and how they think and behave, and not the technology [26]. We will discuss it in more detail in the next section.

Usage-Centered Design: In *Usage-Centered Design (UsageCD)* the focus is on the interaction tasks as a basis for significantly more usable tools. It is not primarily users who must be understood, but *usage*, which is how and for what ends software tools will be employed. UsageCD focuses on the work that users are trying to achieve and on what the system will need to supply via the user interface to help them accomplish it [25].

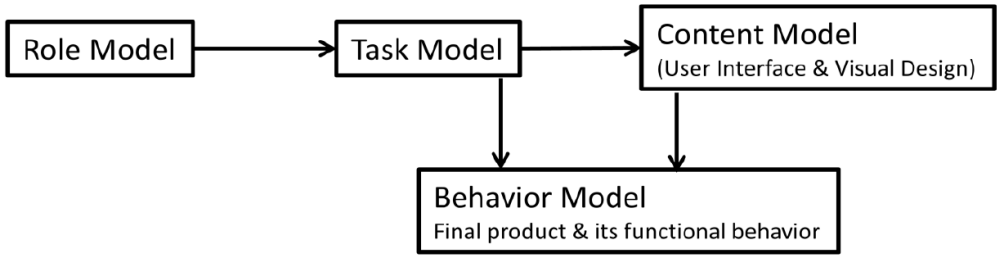


Figure 5.5: Model based Usage-Centered Design – Relationships between models [25].

Figure 5.5 shows a visual representation of the relationships between models in UsageCD [25]. It has three simple models of what one needs to apply to find solutions for user interface design problems: role model, task model and content model. The role model collects all information about the users, their background and relationship with the system, etc. The task model is a list of all the tasks and use cases that the user will do when interacting with the system. The content model deals with the basic presentation of the user interface and where, how, and what tools will the user need to perform a particular task. The task model and content model will help us in getting to the behavior model, which in a way reflects the face of the final product that we will design [25]. We will discuss it in more detail in section 5.4.

Our Approach: To explore interaction design methods for our framework, we selected personas from GDD, and essential use cases from UsageCD. As well as being influential techniques, these two methods are specially relevant to our needs. Personas promote consideration of personal characteristics of important categories of users, and this commonly includes issues related to affect. Essential Use Cases (EUCs), distinct from ordinary use cases, allow design ideas to be described conceptually, rather than concretely. This should allow us to specify strategies to drive design.

5.2 Detail Exposition of Goal-Directed Design

The components of GDD are persona, scenario and end goal. Personas are rich descriptions of typical users of the system under development that the designers can focus on and design the product for. They do not describe real people, but are *realistic* and not idealized. Each persona has a number of goals relating to the particular system under development. It also includes a description of the user's skills, attitudes, motivations, main points, tasks and environment. These personal precise credible details will help designers to see the personas as real potential users to design for [117].

A persona is an archetypal model that communicates research patterns about a type of user in the present. A persona is depicted as a specific person, but is not a real individual; they are synthesized from observations of many people.

A scenario defines when, where, and how the story of the persona takes place. The scenario is the narrative that describes how the persona behaves as a sequence of events. The scenario is a narrative that describes how a persona would interact with software in a particular context to achieve their end goal(s). Scenarios are written from the persona's perspective, at a high level, and articulate use cases that will likely happen in the future. The goal defines what the persona wants or needs to fulfill. The goal is the motivation of why the persona is taking action. When that goal is reached, the scenario ends. End goal is an objective that a persona wants or needs to fulfill by using the system. The software would aid the persona to complete their end goal(s) by enabling them to accomplish their tasks via certain features [56]. According to Cooper et al. [27], GDD is a process that follows six phases: research, modeling, requirements definition, framework definition, refinement, and support as shown in Figure 5.6.

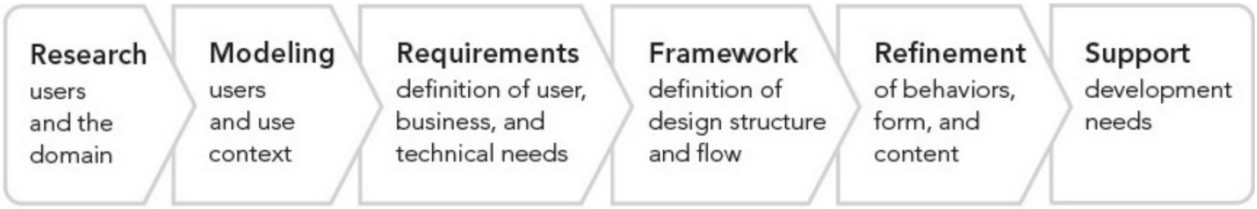


Figure 5.6: Goal-Directed Design process [27].

The research phase employs observation and contextual interviews providing qualitative data about users of the system. In the modeling phase, behavior and workflow patterns discovered through analysis of the field research and interviews are synthesized into domain models (information flow and workflow diagrams) and user models (personas). In the requirements definition phase, design methods are employed by teams. They provide the connections between the user and other models and the framework of the design. During the framework definition phase, designers make the overall system concept. The refinement phase proceeds similarly to the framework definition phase; however, increasing the focus on detail and implementation. The development support answers developers' questions during the construction process [27], [46].

5.3 Proposed Adaption of Personas from Goal-Directed Design

We propose taking the established GDD methodology and adapting it to our purposes by adding learning objectives, affective and cognitive strategies and multiple sensory modalities. Our framework suggests that system design needs to accommodate ideas for cognitive, affective and multimodal strategies. By modifying the persona technique, we hope to provide a method for doing this.

Here we focus on how to adapt personas in an affective multimodal system. Cooper et al. [26] have provided examples of personas. Personas can be useful in developing a better view of the user, are part of GDD, and are useful in developing a better view of the users.

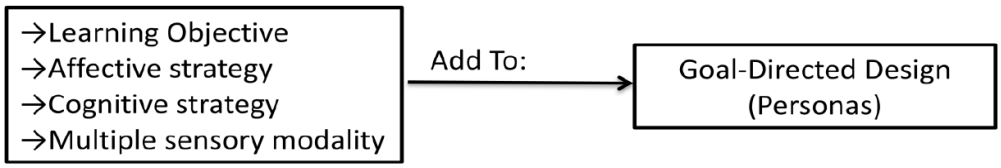


Figure 5.7: Adding education, affective, cognitive and multiple sensory modality to GDD.

The personas in the educational multimodal system, the instructor and the learner, want to be able to do things to: use the system, achieve their goals, and undertake meaningful activities using the multimodal system. Our overall suggestion is that MADE should support design of multimodal system. We create the personas so that we can envisage whom we are

designing for and to allow and imagine ourselves to see or experience something from someone else's point of view and from a user's position. Therefore, it is important to create several different personas for our new multimodal system, because it will be used by different types of people. Although we are going to focus on affective outcomes, we can never completely ignore the cognitive ones, because if the cognitive ones in fact do not work, then the affective strategies are pointless. For example, in a primary school arithmetic class, teachers might provide a lot of affective strategies such as little stories, cartoon characters. Those are very nice to help of multiple sensory and quasi-sensory modalities, but that does not help the cognitive strategies such as "how to add numbers, columns, etc". If teachers don't have a cognitive strategy for teaching arithmetic, no affective strategy will help them to succeed (see Figure 5.8).

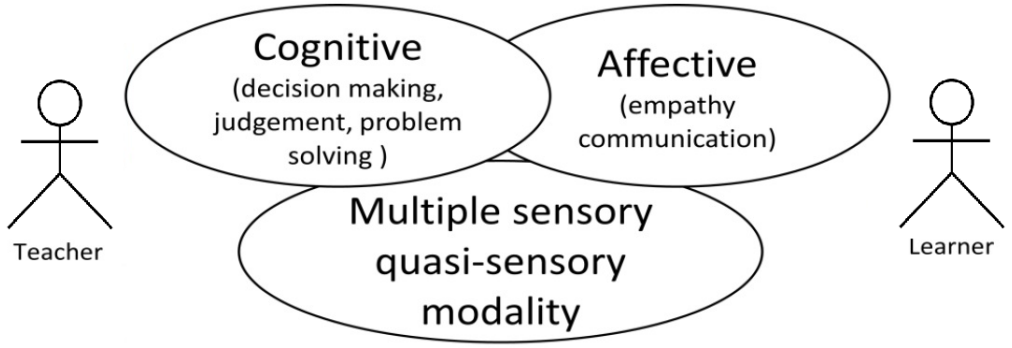


Figure 5.8: Teacher provides affective and cognitive strategies through sensory and quasi-sensory modalities.

We are identifying whether there are, e.g., genius students who always want to be challenged with extra materials, or very hesitant, insecure students who need to have confidence, since all of these will effect affective strategies. However, the really critical thing is that those personas should have elements that support affective learning. Then a system designer will know how it might support them. With personas we consider: behavior patterns, goals, skills, attitudes, environment, and some fictional personal details to bring the persona to life.

For a multimodal educational system, personas are developed to explore the various needs of different people. In our example the personas are fictional. To ensure comprehensive coverage of our framework we defined a better learning environment according to the affective and cognitive strategies with the help of the multiple sensory modality, we use three personas:

one unconcerned student, one student lacking confidence, whose parents would like him to become an engineer, and finally, one more inclined student who has a reading disability (dyslexia). The first scenario introduces the reader to Sarah, an unconcerned student who likes to listen to music when studying. Next, we assume the role of Sarah's classmate Oliver. His parents want him to become an engineer and therefore he feels pressure on him and he worries about it. Lastly, Mike has a learning disability. Therefore, they for example would need more engagement and persuasion to learn by taking into consideration the affective and cognitive strategies. These personas are shown in Figures 5.9, 5.10, and 5.11, following the graphic approach advocated by Cooper et al. [26].

Figure 5.9: 1st learner persona for MADE educational scenario.

Figure 5.9 shows first learner persona (Sarah). She prefers to listen to music, dance and go clubbing, and does not have the passion to study and concentrate in the classroom. The instructor has to encourage her with affective strategies such as humor, emotion, and use

theories of cognition such as embodied interaction, with using multiple sensory modalities such as funny sound effects, pleasing interface to bring more engagement, attention and pleasure. Figure 5.10 shows a second learner persona (Oliver). He worries about the future to be an engineer, because of his parents. He likes to be relaxed and have more fun, and is not very ambitious. He is afraid of not being successful. The instructor has to encourage him with affective and cognitive strategies like humor, emotion and use persuasive learning techniques and multiple sensory modalities such as encouraging videos and haptic devices to bring more motivation. Of course in real applications we should follow the methods, such as contextual inquiry and qualitative analysis, of Goal-Directed Design to determine the persona. Figure

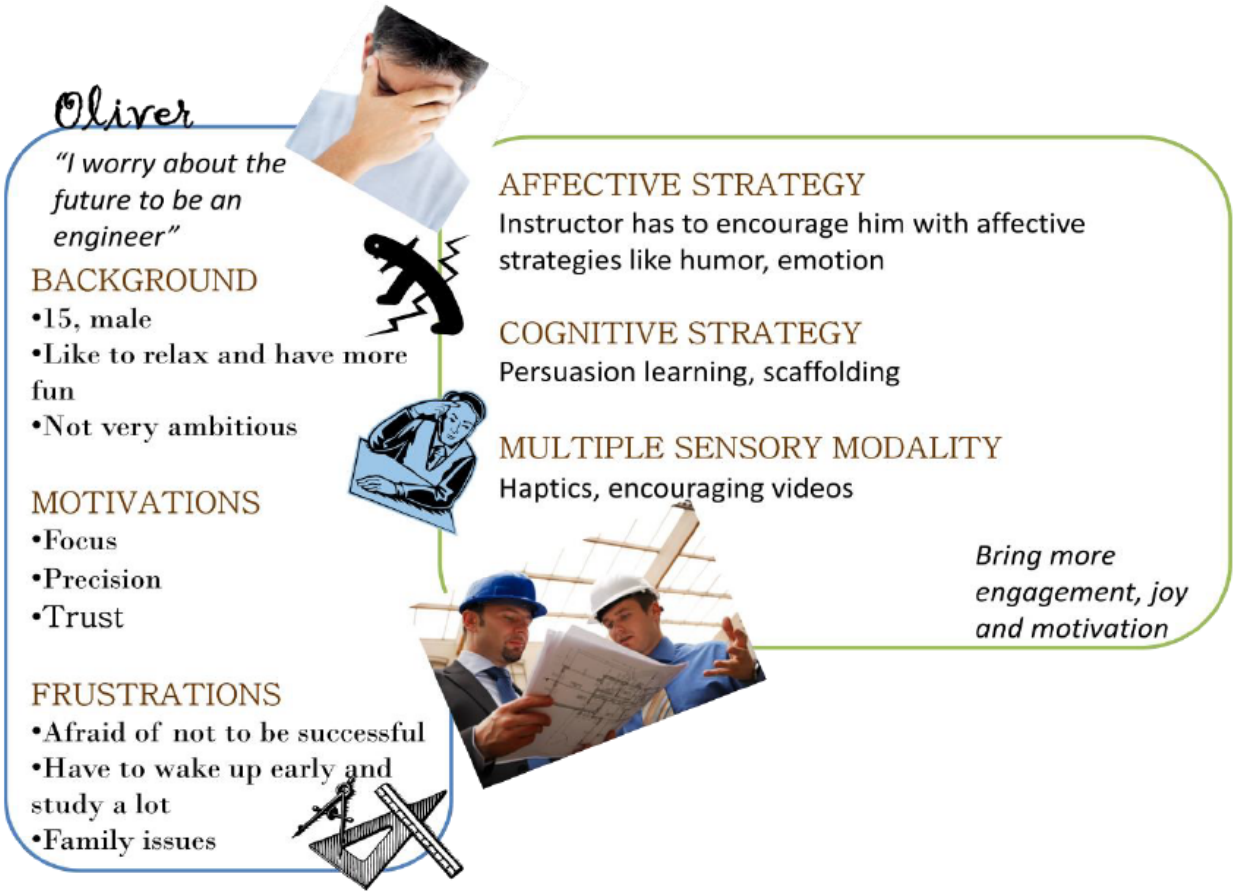


Figure 5.10: 2nd learner persona for MADE educational scenario.

5.11 shows the third learner persona (Mike). Mike has a disability called Dyslexia (related to difficulty reading) that can affect his learning. He does not feel confident and fears making mistakes. He worries other students make fun of him. Instructors have to encourage him with

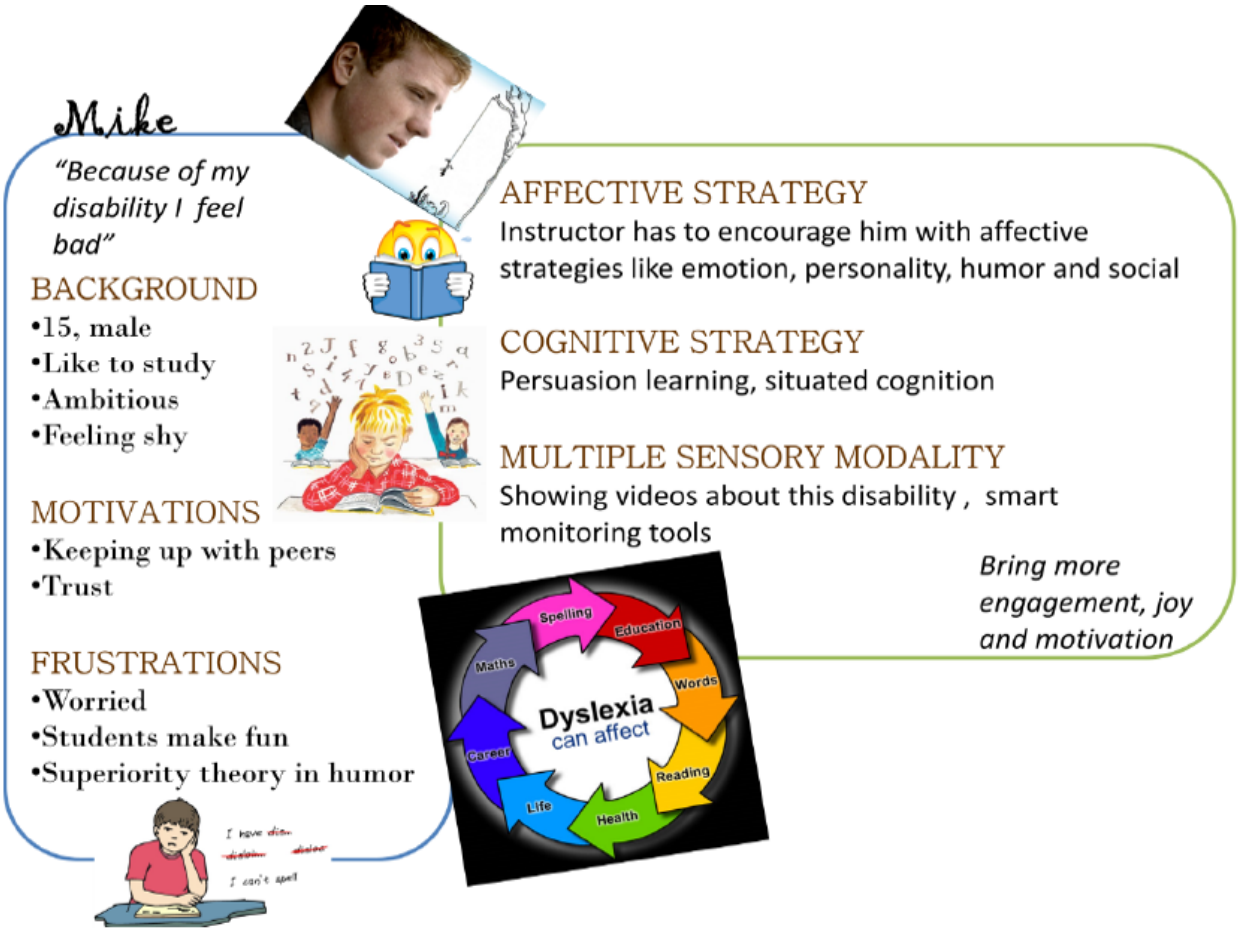


Figure 5.11: 3rd learner persona for MADE educational scenario.

affective and cognitive strategies like humor, emotion, persuasion learning, and use multiple sensory modalities such as haptics and encouraging videos to show in the classroom so he feels better and other students cheer up Mike when he fails to read or interpret letters and other symbols.

We will explore the effectiveness of our modification to personas in a case study presented in Chapter 7.

5.4 Detail Exposition of Usage-Centred Design

We can also adapt UsageCD with learning objectives, affective and cognitive strategies and multiple sensory modalities. UsageCD was based on user intentions and usage patterns for user interface design. It analyzes users' roles that they play in relation to systems and employs *abstract (essential) use cases* for task analysis (see Figure 5.12).

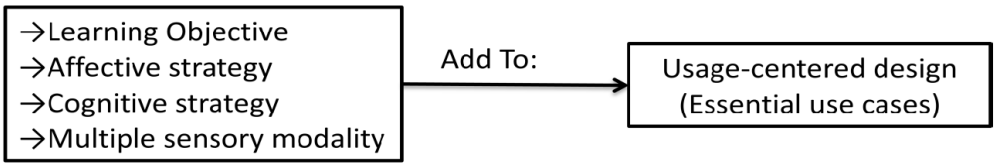


Figure 5.12: Adding education, affective, cognitive and multiple sensory modality to UsageCD.

Use cases are created to focus on user goals. They were introduced in the book Object-Oriented Software Engineering by Jacobson et al. [66], and work by focusing on and documenting the interaction between the user (actor) and a software system. Use cases can be described graphically as well (see Figure 5.1 and 5.2). Actors can be linked to more than one use case and each use case also can be linked to more than one actor. To develop a use case, at first we have to identify the actors, that is the people or other systems which will be interacting with the system under development. Then we look at these actors and identify their goal or goals in using the system and the interaction that will support these. Each of these will be a use case. They are precise and we can implement directly from them. Use cases are beneficial when we are ready to build the software. Therefore, however, traditional use cases contain certain assumptions including the technology, user interface interaction design [117]. They therefore assume some design decisions have already been made.

Essential use cases (EUCs), also called task cases, were developed by Constantine and Lockwood [25] to compensate the limitations of concrete scenarios and use cases. They are called *essential* because they represent the *essence* of the use case. The EUC is often dramatically shorter and simpler than a use case for the same interactions. The key feature of an EUC is that it is of a higher level compared to traditional use cases. It does not say exactly "how" to build a system. It says "what" we need to accomplish in our design. They represent a

more general case than a scenario embodies, and tries to avoid assumptions of traditional use cases. A EUC has three component parts: a short, fully descriptive name; a name that expresses the overall user purpose or intention, plus a two-part narrative comprising the user intention model and the system responsibility model. The division between user intention and system responsibility can be helpful in conceptual design considering task allocation and user responsibility, and system scope and what it can do. Essential models are in the form

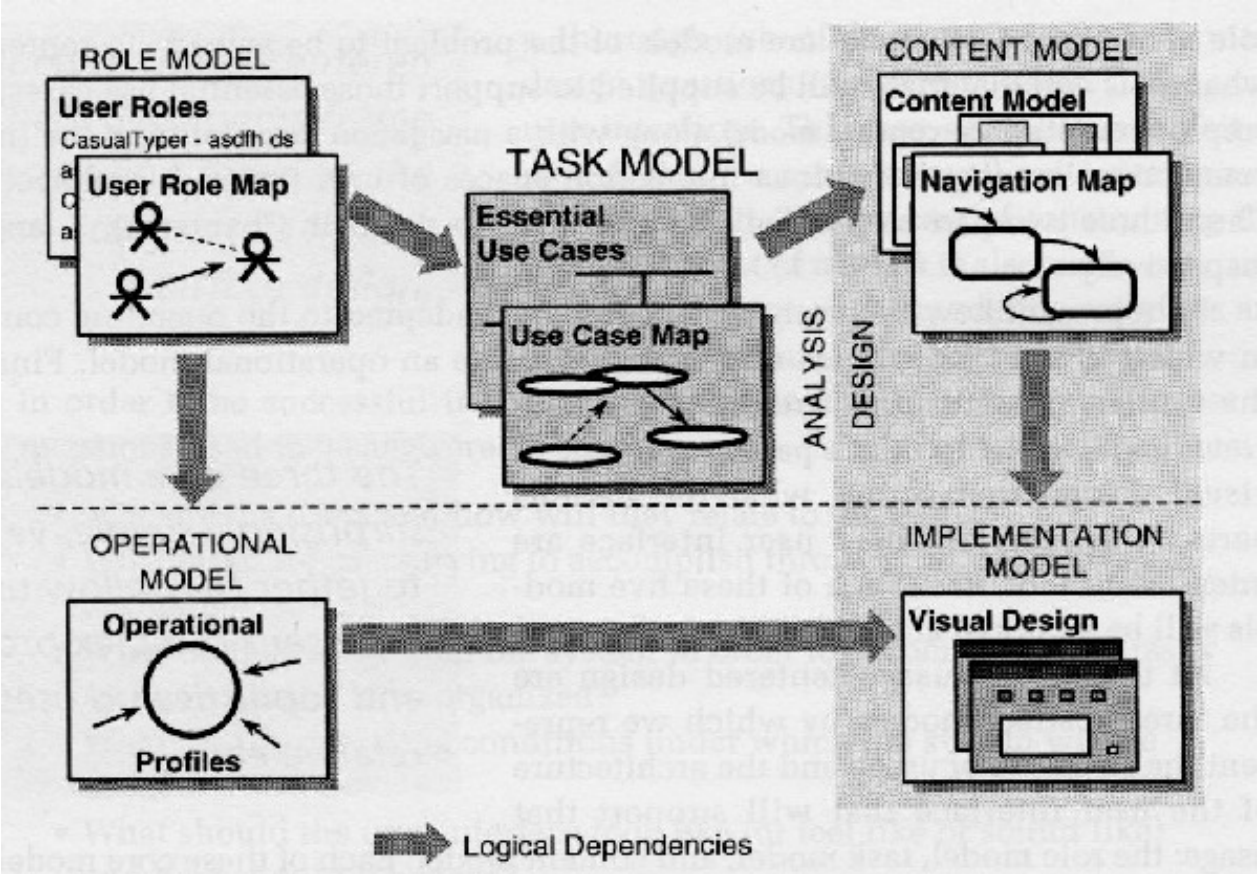


Figure 5.13: Relationship among essential models in Usage-Centered Design [25].

of a collection of descriptions and a map of interrelationships (see Figure 5.13). The role model captures relationship between users and the system (user roles and user role map). The task model shows structure of tasks users need to do (essential use cases and use case map). The content model tools are supplied by the interface (interface context model and navigation map). The operational model specifies the conditions and constraints under which

the system will operate (operational constraints). The implementation model shows the visual design of the model. Figure 5.14 presents an overview of the UsageCD process showing the key activities and some of the relationships among the tasks it comprises. In Figure 5.15

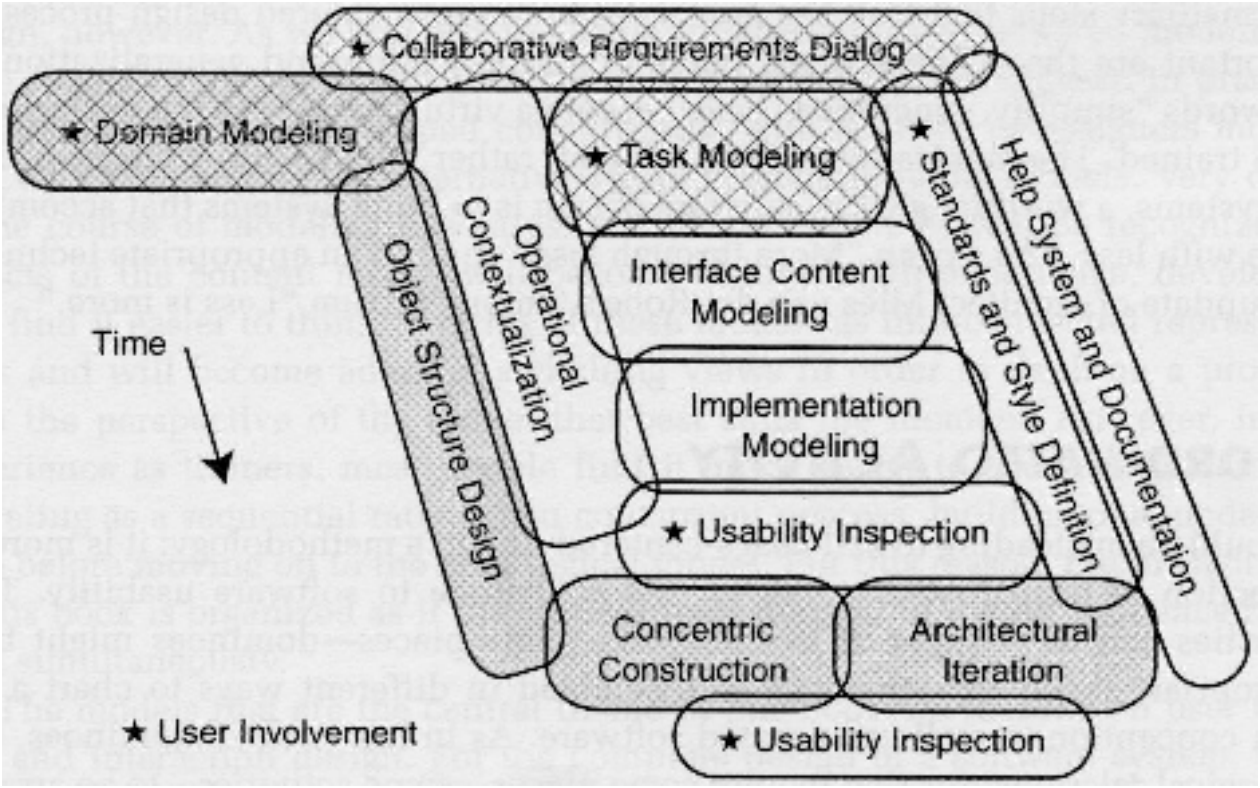


Figure 5.14: Constantine and Lockwood's usage-centered design activities [25].

we illustrate an example of the cash withdrawal use case expressed as a partitioned narrative [74]. In conventional (concrete) use cases there are the user *action* model and the system *response* model, while in essential use cases we have the user *intention* model and the system *responsibility* model. This EUC is dramatically shorter and simpler than the conventional use cases, because it contains only the steps that are essential and of intrinsic interest to the user [25].

5.5 Proposed Adaptation of Usage-Centered Design for Affective Education

We now consider adaptation of the essential use case, the key part of the UsageCD process. As we mentioned before, it is called an essential use case because it captures the essence of

gettingCash

User Action	*System Response*
insert card	
	read magnetic stripe
	request PIN
enter PIN	
	verify PIN
	display transaction menu
press key	
	display account menu
press key	
	prompt for amount
enter amount	
	display amount
press key	
	return card
take card	
	dispense cash
take cash	

gettingCash

User Intention	*System Responsibility*
identify self	
	verify identity
	offer choices
choose	
	dispense cash
take cash	

Figure 5.15: Left side: a conventional use case for getting cash from an automatic teller system, Right side: an essential use case for getting cash from an automatic teller system (From Constantine and Lockwood) [9].

the use case. That means it is supposed to document what the result is meant to be, and not how one may expect to accomplish it. The advantage of it is that then we can document the "what" and leave it to the designer to come up with "how" (see Figure 5.16). The whole idea of EUC is to focus on what the outcomes of the use case are meant to be. After, we have to see the different ways of implementing it. Therefore, the outcome of that use case basically should involve for example from the teacher, cognitive outcomes and affective outcomes.

In adapting the EUC we might identify the affective goal of the use case to support consideration of how we might achieve it. We need to consider what multi sensory modality methods might have desirable affect. Again, our framework requires that affective and multimodal strategies need to inform the system design. Our proposal for personas addresses the need to understand these strategies in relation to the learner. By similarly adapting the EUC, we can identify an interactive concept for affect and multimodality in the system design.

For this use case there would also be a cognitive goal. In our earlier example, the student must understand how to add two one digit numbers or understanding the ratio, and an

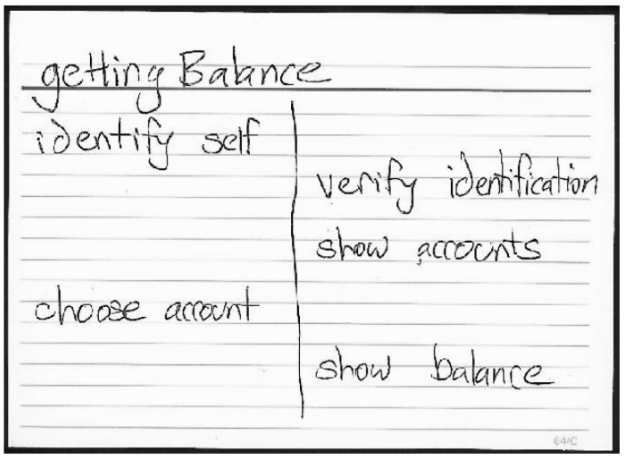


Figure 5.16: An Essential Use Case Card [9].

affective goal might involve receiving affective affirmation when the task is complete such as a congratulation text, change of screen color, or a happy audio. We know that using sound has some beneficial effects; happy encouraging sounds and animation on the screen can come to our design to have emotional influence on the user.

EUCs always have to start with goals of what is going to be accomplished at the end of use case. For instance, in the use case of depositing money in the bank, goals will be that at the end of transaction user's balance has gone up by that amount and the log records the transaction. We are adding the affective objective so designers can decide how to achieve that affective objective. What we are doing here is different than with the personas idea. The personas do not explain what we are trying to do. They just address the emotional characteristics of the user. Each use case will have an affective objective.

We add a part called the "Affective sidebar" to the essential use case card in Figure 5.17, and have a column in the right side. Our focus is on education and not on banking, but here we extend the example of a banking situation. We want the users to feel good by suggesting the bank is taking care of their money and it is doing good work for them, which transmits a kind of affective messages, as well as letting them take the cash out. The bank's strategy is to be friendly, trustworthy and professional, and their intent is that we come up with a good feeling. With multimodal strategy we can play different kinds of music or video clips to further support the goals.

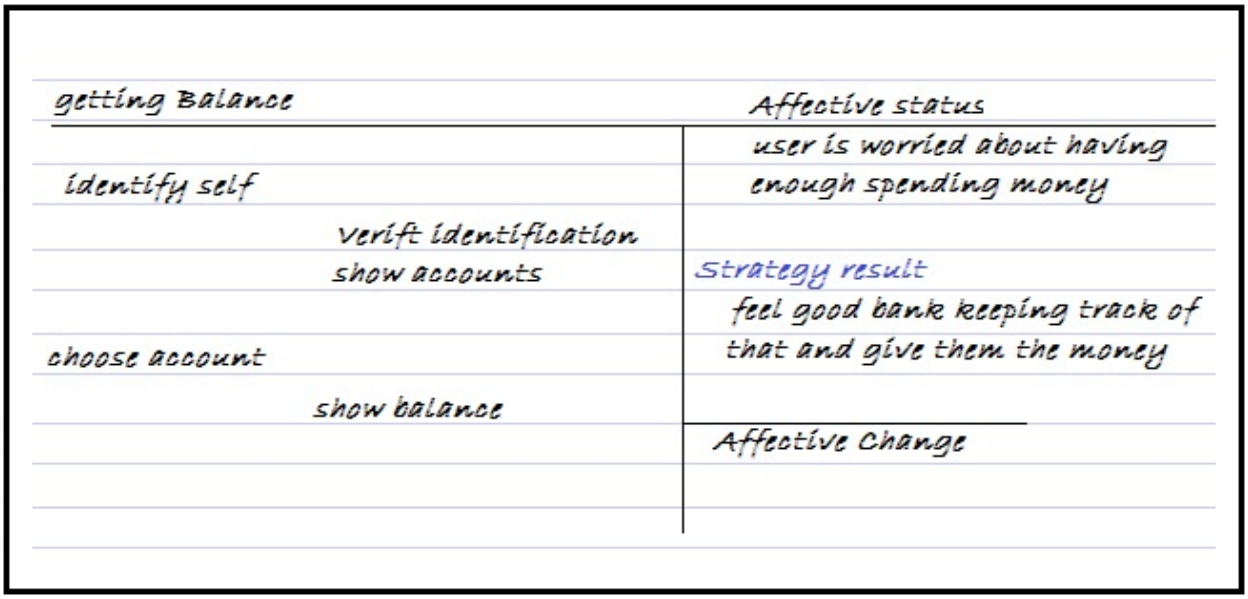


Figure 5.17: An Essential Use Case Card using the Affective Strategy.

The whole idea of an EUC is that we only specify the essence of what is going to work, and then we try different alternatives to come up with that essence. Therefore, we could be talking about it in computer terms and the essence of enjoyment and fun. One thing that we know about the enjoyment and fun in multimodal systems is that we could be using funny sound effects, e.g. happy sounds. Therefore, now we have one possibility for implementing the EUC. There are other things we could be doing as well. We could have pictures with cartoon characters and we could personalize it. There could be a joke or similar things, which are alternatives that can be explored.

To assess our proposal for the modified EUC, we utilized this approach in the case study presented in Chapter 7.

5.6 Summary

In this chapter we took our MADE framework and applied it to design. This requires ideas for learning strategies to feed into the system design process. We first reviewed several UI design methodologies. We then took one of the methods; the GDD, and applied our

framework to this method, and showed an adaptation of personas from GDD. We then took another method; the UsageCD, and applied our framework to this method, and showed an adaptation of Essential Use Case from UsageCD. These adaptations are explored in the case study in Chapter 7.

Chapter 6

Using the Framework for Evaluation

This chapter focuses on evaluation techniques for affective multimodal software. We explore how multimodal educational systems can benefit from affect, and give proposals for how to evaluate these systems. Our goal is to come up with HCI evaluation techniques for multimodal systems to support *affective* education. We are proposing two usability inspection methods. These are demonstrations of how the proposed MADE framework can be applied to well-known evaluation techniques.

In the area of multimodal system design, user studies have been conducted to explore how users combine modalities when interacting with a system. The main idea of this chapter is to adapt some popular evaluation methods to the *MADE* framework and answer the question: does this framework support evaluation of affect in a multimodal interface. The *heuristic evaluation*, the *cognitive walkthrough* and the *pluralistic walkthrough* belong to a family of techniques called *Usability Inspection Methods*, which do not need real users. We demonstrate the MADE framework using the evaluation methodologies and propose an *Affective Walkthrough* and an *Affective Heuristic Evaluation* (see Figure 6.1). The proposed affective walkthrough is an adaptation of the cognitive walkthrough and pluralistic walkthrough, and we are adapting them to develop new sets of heuristics for evaluating affective multimodal educational systems trying to better explore the new design space of educational multimodal system design, and to help us with our designs.

In this chapter, we wish to evaluate a HCI system and we need to answer the questions: how might the MADE framework be leveraged for evaluation, and does the software support affect in education. We proposed the following usability inspection methods: affective adaptation of the cognitive walkthrough, having a proposed affective walkthrough system and adding a modification of the pluralistic walkthrough to it and the proposed affective

heuristic evaluation.

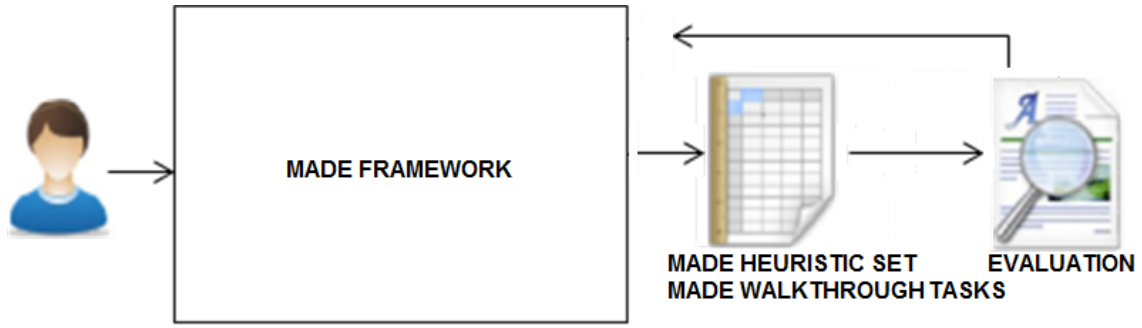


Figure 6.1: The MADE framework and the evaluation.

6.1 Review of the MADE Framework

We have proposed the MADE framework based on principles of multimodal design and theories of multimodal interaction. We considered both *affective* and *cognitive strategies* of learners while interacting with a multimodal system, which has *multiple sensory modalities* (visual, auditory or tactile) and *quasi-sensory modalities* (e.g. narrative or persuasion) domains with *learning objectives* for education and learning. For learning objectives there are a number of affective strategies and cognitive strategies.

The learning objective controls the metrics, affective and cognitive strategies and the linkages. These strategies will inform the instructor, student and educational technologies. In the next sections we explain cognitive walkthrough, pluralistic walkthrough, heuristic evaluation and the proposed usability inspection methods. All will be adapted for the framework.

6.2 Evaluation Methodologies

In this section, we focus on cognitive walkthroughs, pluralistic walkthroughs and heuristic evaluation. Walkthroughs are an alternative approach to heuristic evaluations for predicting users' problems without doing user testing. They involve walking through a task in a system and noting problematic usability features. Cognitive walkthroughs involve simulating a user's

problem-solving process at each step in the human-computer dialog, checking to see how users progress from step to step in these interactions and determining if the user's goals and memory for actions can be assumed to lead to the next correct action [89]. A key feature of cognitive walkthroughs are that they focus on evaluating designs for ease of learning [117]. The pluralistic walkthrough method involves following a scenario and the discussion of potential usability issues (see Table 6.1). Heuristic evaluations are low cost, fast and efficient methods of being able to identify many usability issues that may occur with the system. Most people who use heuristic evaluations would perform them based on their intuition and common sense [104]. A heuristic evaluation finds a list of problems as well as indications of how to solve each problem. It is one of the most cost-effective methods to discover usability issues in the design process [82]. The next evaluation technique is the user testing which actually tests the system with real users. It's a technique in user-centered interaction design to evaluate a system by testing the users, e.g. applying the heuristics.

Table 6.1: Summary of software usability testing methods - inspection methods.

Method Name	**Brief Description**	**Who Is Involved**
Cognitive Walk-throughs	Simulate users' problem-solving processes. This test evaluates whether the simulated user's goals lead from one action to the next correctly.	Usability specialists, e.g. HCI experts
Heuristic Evaluation	An informal way to determine whether the interface conforms to established usability principles	Usability specialists, e.g. HCI experts
Pluralistic Walk-through	This method involves following a scenario [e.g. a possible software use], and the discussion of potential usability issues.	Representative users, Developers, and HCI experts

6.3 Walkthroughs

Now we give a detailed exposition of cognitive walkthrough and pluralistic walkthrough, and after the proposed affective walkthrough; the adaptation of those two walkthroughs.

6.3.1 Existing Walkthrough Sources

Cognitive Walkthrough

The cognitive walkthrough method was developed in the early nineties by Wharton et al. [150]. This method required asking four questions that the evaluator asks for each action, and produces extensive documentation of the analysis (see Table 6.2).

Table 6.2: The cognitive walkthrough questions [150].

The Cognitive Walkthrough Usability Inspection Method	
1	Will the user try to achieve the right effect that the subtask has?
2	Will the user notice that the correct action is available?
3	Will the user understand that the wanted subtask can be achieved by the action?
4	If the correct action is performed, will the user see that progress is being made toward solution of the task and get feedback?

Cognitive Walkthrough Procedure: The cognitive walkthrough is a usability inspection method that links the system walkthrough to a cognitive model. It is based on one idea that the learnability of the system helps the usability, a standard approach in the usability literature for identifying "pain points" that cause users to fail at completing given tasks. These walkthroughs show that users of all skill levels are likely to encounter problems [53]. Thus, it is all about how easy it is for the user to learn how to use the system. The evaluator will use the interface (in our case, a multimodal system) to perform tasks that a typical system will need to support. The actions and reactions of the system are evaluated according to the user's goals and knowledge through reactions to questions related to the method's cognitive model, the differences between the user's expectations and the reality. The cognitive walkthrough focuses on the basic principles of usability and it focuses on the cognitive activities of users especially on their goals and knowledge while performing a specific task [90].

Pluralistic Walkthrough

The pluralistic walkthrough [8] is used to identify usability issues in a piece of software in an effort to create a maximally usable HCI, where each participant takes the role of a user walk through in the design. This method increases empathy and focuses on using a group of representative potential users, developers and usability and human factors professionals.

Pluralistic Walkthrough Procedure: The group is asked to put themselves in the shoes of the users, to step through a task scenario, discussing usability issues and problems involved in the scenario steps and assumed the role of typical users in the testing, rising developers' sensitivity to users' concerns about the software design [8].

It begins with a brief overview followed by all participants going through the system. They represent with hard copy scenarios and write down their actions. The participants will have a semi-formalized discussion after each scenario. A session administrator moderates and facilitates the session to keep users motivated, comment on their actions, and not allow developers to influence users [2].

Table 6.3: Five defining characteristics of the pluralistic walkthrough [8].

Five Characteristics of the Pluralistic Walkthrough	
1	The method includes three types of participants in the same walkthrough session: users, usability experts and system designers
2	The system is presented with hardcopy panels and these panels are presented in the same order as they would appear in the system
3	All participants taking the role of a user
4	The participants write down the actions they would take to perform the given tasks
5	The group discusses the solutions to which they have reached. The administrator first presents a correct answer. Then the users describe their solutions, and only after that, do the designers and usability experts offer their opinions

In the next subsections we develop some walkthroughs and heuristic.

6.3.2 Affective Walkthrough

The Affective Adaptation of the Cognitive Walkthrough

The whole basis of the cognitive walkthrough involves understanding of the interface. We can adapt the cognitive walkthrough for multimodal cognition and affect. We explained in the previous subsection how cognitive walkthroughs and pluralistic walkthroughs work. Below we describe how we change the cognitive walkthrough to add affective and multimodal factors, by including some aspects of the pluralistic walkthrough (see Table 6.3).

In this part we bring an affective dimension into the cognitive walkthrough and create a blended cognitive and affective walkthrough. When we modify the cognitive walkthrough, we must not only address the learnability of the user interface, but the learning objective that the instructor has determined; thus, focusing on the cognitive learning strategy of the instructor. We separate the learnability issue in the education from the learnability issue in the user interface. And then we consider the affective strategy that the instructor has come up with, and consider the affective cognitive multimodal walkthrough.

Therefore, it is all about the learner achieving the instructor's objective and learning the objective of the instructor, which is about the cognitive strategies and the affective strategies. The proposed affective walkthrough is based on Wharton [150] and the paper of Dormann and Biddle [40], which studied the affective elements of computer games. The idea of the affective walkthroughs is that we walk through and seriously think about the affective manipulation that is going on and then reflect on how the design was successful and better understand how affective learning can be supported in the system. In multimodal systems specifically, the nature of goals and interaction typically involves both cognitive and affective elements. In these systems with interacting characters, affective strategy is naturally involved as a way of motivating learners and it influences the trust perceptions. Dormann and Biddle [40] created an affective walkthrough for games. They considered emotions in users, the avatar and other characters; the relationships depicted between characters; the aesthetic issues that relate to these issues, and ask the following questions while taking into consideration the overall question of: how would the affective experiences persuade the experience and goals in gameplay (see Table 6.4):

Table 6.4: The affective walkthrough questions for computer games [40].

The Affective Walkthrough for Computer Games	
1	What might the player, their avatar, or other characters, experience emotionally?
2	What relationships are depicted or might emerge between characters?
3	What effect will the setting have on the player, their avatar, or other characters?

We propose the affective walkthrough by considering the learnability and the learning strategy (affective and cognitive strategies that the instructor has come up with) to understand and analyze affective and cognitive strategies in education multimodal systems. Our proposal is based on Wharton et al. [150] and Dormann and Biddle [40]. As with those methods, the evaluator steps through the system, and at each point, considers relevant questions. We adapted the questions from the method of Wharton et al., adding consideration specific to affect and multimodality. This approach was similar to that of Dormann and Biddle, but their questions where specific to computer games, where as ours are specific to education. Our questions are shown in Table 6.5

Table 6.5: The affective walkthrough.

The Affective Walkthrough	
1	Will the learner be able to use the multiple sensory modality correctly (learnability in the user interface) and taking into consideration the learning objective that the instructor has set up and the cognitive and affective learning strategies of the instructor that might help the learner to achieve the right effect that the subtask needs (learnability in the education)? E.g. what cognitive and affective strategies might help the learner achieve the instructor's cognitive and affective objectives? Does the learner understand that a subtask like specific hand gesture with correct distance is needed to reach the learner's goal? What might the learner experience emotionally?
2	Will the multiple sensory modality helpful and does the learner notice that the correct action is available? E.g. is the multimodal system sound clear and visible?
3	Will the learner understand and associate the correct action by using the multiple sensory modality with the effect trying to be achieved? (learnability of the user interface) E.g. the multimodal system sound is clear and visible but the learner does not understand the text and sound, and will therefore not respond to it.
4	When the correct action is performed, will the learner know that they have done the right thing, receive affective strategies, and see that progress is being made toward solution of the task? E.g. does the learner get haptic feedback, sounds or written message when the correct action is performed?

As we mentioned in the previous subsection, we created the modified version of the pluralistic walkthrough for the MADE framework to consider in the affective walkthrough. Here, we have different personas and we add the instructor to it (see Figure 6.2).

Proposed Modification to the Pluralistic Walkthrough for MADE

We consider the pluralistic walkthrough because we would like to consider the different perspectives. We will have participants playing the role of the personas in order to do the evaluation, as well as adding the instructor. Each different persona would say what they think about the scenario, e.g. if the one that is too confident, says: "I think it is great", that

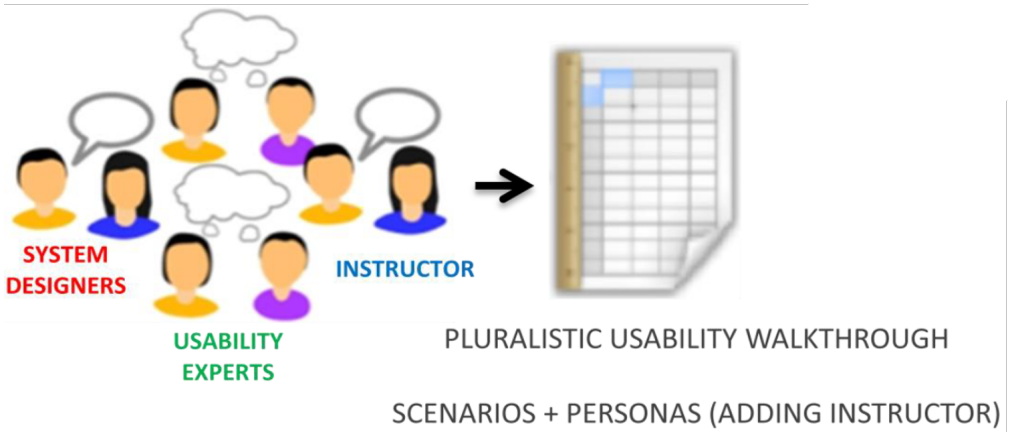


Figure 6.2: The modified pluralistic walkthrough.

would benefit the one who needs encouragement. We are trying to support each of those personas.

Pluralistic Walkthrough Procedure: The pluralistic walkthrough consists of asking participants with a different perspective on the system to engage in a task scenario. It implies a procedure like the standard usability walkthrough however users, developers, usability experts, and instructors step through the design and together discuss usability issues that they discover during the walkthrough process. In the process, there are five defining characteristics (see Table 6.3) [8]. First, there are three types of participants in the same walkthrough: representative users, product developers and human factors professionals. We are basically representing the users through the personas. Therefore, we have participants (evaluators) playing the role of the personas. In the second part, a scenario is defined. Third, participants all assume the role of the user. The developers and the usability professionals will try to put themselves in the place of the users and make written responses. Fourth, participants will write down the actions they perform to complete the tasks in as much detail as possible, before any discussion, e.g. instead of "I would choose the fourth item on the list", he/she will write "Press the down arrow key three times, then press *Enter*", which produces some quantitative data on user actions. Fifth and finally, the discussion begins after participants have written the actions they would take to complete the tasks. The representative users speak first when discussing each panel. Usability experts and the product developers say their opinions if representative user's comments are exhausted [102].

This is a group activity that participants are presented with instructions, task descriptions and scenario packages. After, a system expert (usually a designer) will offer a brief overview of key system concepts and features. Next, participants will write the actions they would perform for specific task. When they are done with writing their independent responses, the "right" answer will be announced, and the representative subjects express their responses and discuss potential usability problems. During this time the system experts remain quiet and the human factors professionals simply facilitate the discussion among the users.

6.4 Heuristic Evaluation

In this part we first explain some heuristics including Nielsen's ten general principles of interaction design heuristics [104] and Sankey's fifteen multimodal design heuristics [132], which are used for education. Next, we explain Kort et al.'s affective model [79] and Norman's emotional design model [105], considering them in our heuristics. Finally, we use these two heuristics and affective models, and will come up with new kinds of heuristics.

6.4.1 Existing Heuristic Evaluation Sources

Nielsen's Heuristic Evaluation

Heuristic evaluation is one of the most applied of the usability inspection methods and is essentially based on the identification of usability issues associated with a list of quality criteria. It involves examining the system independently using the recognized usability principles (the heuristics) and the usability issues [102].

The heuristic evaluation process is based on the usability principles developed by Nielsen and Molich [104] and Nielsen [102] to test the usability aspects of a system and to identify and classify usability problems. Each heuristic is displayed in Table 6.6.

Table 6.6: Nielsen's ten heuristics for user interface design [89].

Heuristic Number	Heuristic Description
H1	Visibility of system status
H2	Match between system and the real world
H3	User control and freedom
H4	Consistency and standards
H5	Error prevention
H6	Recognition rather than recall
H7	Flexibility and efficiency of use
H8	Aesthetic and minimalist design
H9	Help users recognize, diagnose, and recover from errors
H10	Help and documentation

Heuristic Evaluation Procedure: In heuristic evaluation, each individual evaluator first inspects the system alone. Once all evaluations are done, the evaluators then can communicate and have their findings collected. This procedure is done to make sure independent and unbiased evaluations come from each evaluator [102]. We believe that a custom set of principles are needed so that heuristic evaluation can be used to find usability problems in multimodal affective educational systems.

Sankey's Heuristics

In Table 6.7 we show the 15 multimodal design heuristics proposed by Sankey [132]. Sankey has tried to demonstrate distinct advantages for students in providing course resources designed to suit a range of different learning modalities (multimodal).

Table 6.7: Fifteen multimodal heuristics proposed by Sankey [132]

Heuristic Number	Heuristic Description
H1	"Less is more". Lean, precise text gets the point across better than lengthy, elaborated text (Clark &Mayer, 2003; Doolittle et al., 2005)

Table 6.7: Fifteen multimodal heuristics proposed by Sankey [132]

Heuristic Number	**Heuristic Description**
H2	Incorporate, where possible, images that tell a story, giving the learner a reference point or anchor for the information being transmitted. Pedagogical benefit must be present for their use (Clark & Mayer, 2003)
H3	Avoid including additional music or sounds, unless are essential for learning interaction (Sweller, 1999)
H4	Provide the learner with some control over the learning environment, ensuring that the instructional strategy is made clear (Ainsworth & Van Labeke, 2002)
H5	When creating animation, use image and spoken text (Clark & Mayer, 2003)
H6	When using animation, allow access to an alternative version of the material (Mayer, 2001)
H7	Build knowledge gradually with stepwise segments of information (sequentially), not in one long presentation (Kalyuga, Chandler, & Sweller, 2001)
H8	Ensure background image/colour not interfere with clarity of information presented in foreground (Doolittle, 2002)
H9	Use simple graphics initially where possible, then add to complexity as the learning sequence progresses - scaffold visual learning (Kalyuga et al., 2001)
H10	Prevent the need for visual search (Mayer, 2001)
H11	Presence of additional multimedia enhancements made explicit by markers/easily recognisable icons - Clear instruction (Sankey, 2006)
H12	Video/audio may preferred for a lecture style presentation. Visual material reinforce concepts allow students to concentrate which complements the voice (Doolittle et al., 2005)

Table 6.7: Fifteen multimodal heuristics proposed by Sankey [132]

Heuristic Number	Heuristic Description
H13	Printable resource materials is important, for students with a read/write learning preference - make easy to access, clear guidance (Sankey, 2005)
H14	If formative quizzes are used, these should be incorporated and contextualised into the environment, rather than requiring students to link to the Internet, taking them away from the learning environment. Programmed feedback provided with formative quizzes (Neal, 2007)
H15	If audio/video components used, text transcripts be available, if subsequent content not replicate audio content (Sankey, 2005)

Other Sources

In this part, first we explain Kort et al.'s affective model [79], then Norman's [105] three levels of emotional design model, and finally the proposed *Affective Heuristic Evaluation*.

Kort et al.'s Affective Model: Kort proposed an affective cognitive model relating phases of learning process. In Kort et al.'s spiral model the general guide is basically always taking the learner to the next step. We have a loop with an affective strategy at each time. First, the instructor challenges the learner and then he/she lets them explore. Next, he/she helps them to overcome the challenges and finally he/she affirms. It has a four quadrant learning spiral model in which emotions change when the learner moves through the quadrants and up the spiral. When we are at the challenge area, we are not learning because we have to try out new things (see Figure 6.3). In each quadrant we have:

I. The instructor must challenge the learner first.

II. Instructor lets the learner explore and try various things.

III. Learner overcomes the challenges.

IV. Instructor provides the learner emotional support and affirmation.

Therefore, it is a loop, which has an affective strategy at each step and the proposed heuristics that we explain next provides support. Affective design can support each of these four quadrants; a natural cycle having a natural affect appropriate to the task. To support challenge, the software should be stating the challenge, and encouraging the learner to overcome it (encouraging involves affect). For exploring, basically the instructor has to help learners to avoid errors, and supporting and encouraging them to try new things, e.g. in every stage an instructor applies an affective strategy to help this, to start off, we had challenge and encourage. Overcoming is when learners are blocked, and we encourage them to persist to overcome. This is when they have the right idea, but they just need to work harder connecting things. This is where they have to try out new things. Overcoming the wrong ideas from before is called un-learning. After, they get closer to the solution and finally when affirmation happens, we praise them and make them feel good that they have accomplished it; hence, we are giving affective support at each quadrant.

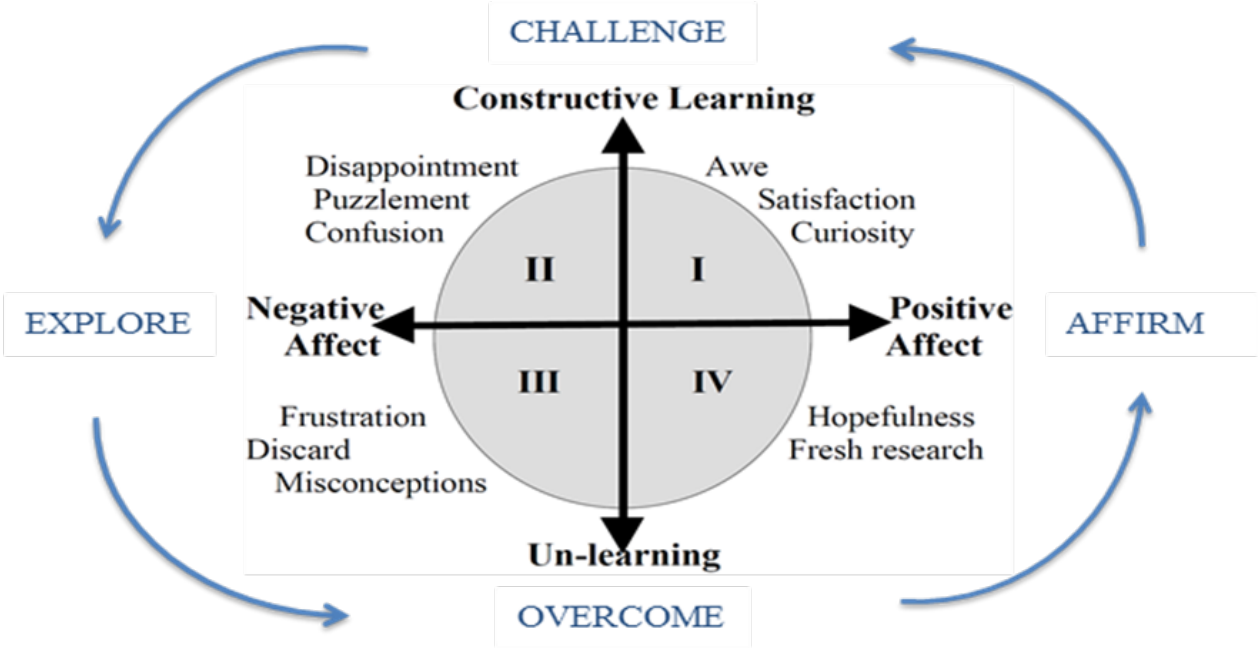


Figure 6.3: Kort et al.'s affective model [79].

In brief, when you learn something, there is a negative affect when doing something wrong and challenging. When we eventually do it right, we experience positive affect because we have done it right. The point is that our software has to work with the learner, and the instructor's job is to harness the negative affect by encouraging the learner, and harness the positive affect by praising the learner, saying that they have done the task right.

Norman's Emotional Design Model: We introduce Norman's model [105] before talking about the proposed heuristics. Norman explained three levels of emotional design including: visceral, behavioral, and reflective. Visceral design refers primarily to the nature of the user interaction (UI) and initial impact of its appearance (attractiveness, aesthetics, beauty, and visual appeal). Behavioral design is about look, feel and total experience of the use of the UI (pleasure and effectiveness of use, efficiency, errors, functional and usable). Reflective design is about the emotional reactions referring to the meaning of the UI; thoughts afterwards, how it makes us feel, the image it portrays, the message it tells others about the owner's taste and the value a UI brings to our self-image (self-image, personal satisfaction, memories, quality or relevance of the information, high in prestige, and rationalization and intellectualization of a UI) [105].

In short, visceral is about how system looks, sounds or feels, for example, image, speech or gesture. Behavioral is about interaction, for example if a learner does something right, the instructor will praise him/her, if he/she does it wrong, he/she will be supportive. Reflective relates to the person's role in the world. For example, what the understanding of the learner is from the system (in a large sense). Visceral, behavioral, and reflective dimensions are interwoven through any design considering both emotion and cognition.

According to Norman [105] the visceral level is quick at making rapid judgments of what is good or bad, sending appropriate signals to the motor system that is the muscles and alerting the rest of the brain, which is the start of the affective processing. The behavioral level is the site of most human behavior. Actions can be enhanced or inhibited by the reflective layer and in turn can enhance or inhibit the visceral layer. The top layer is the reflective thought, which does not have direct access to either sensory input or to the control of behavior. It watches over, reflects upon and tries to bias the behavioral level (see Figure 6.4) [105] . Therefore,

the sensory aspect is directly affecting the behavioral and the visceral and connects us to the reflective. The motor aspect is about how we behave, and the reflective is more how we reflect and our background. Each of these levels should be addressed in new heuristics.

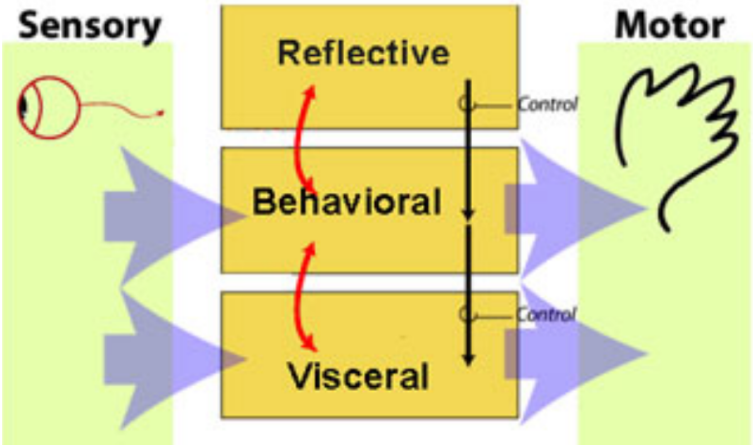


Figure 6.4: Three levels of processing: visceral, behavioral, and reflective [105].

6.4.2 Affective Heuristic Evaluation

In the Affective heuristics we are doing an extension or a comparison to Nielsen's original 10 heuristics [104] by considering cognitive, affective learning and multimodality, all playing a role. There is some overlap with Nielsen's original list, which are the fundamental usability principles.

The affective heuristics are derived from Nielsen's problem categories we had identified and their descriptions of how these problems associated with each heuristic can be avoided and are presented below.

One of our main objectives was to create heuristics that could support learners using multiple sensory modalities, which are used in the educational environment. Sankey [132] proposed 15 multimodal design heuristics. Sankey has tried to demonstrate distinct advantages for students in providing course resources designed to suit a range of different learning modalities (multimodal). The heuristics describe principles with the intention to create usable affective multimodal educational systems. For this purpose, we reviewed and modified Nielsen's usability heuristics, as well as Sankey's multimodal heuristics [132], considering Kort et al.'s

affective learning cycle model [79] and Norman's three affective design categories [105], and we proposed a new set of heuristics.

In Kort et al.'s affective model, the general idea is always to guide and take the learner to the next step, which is more challenging to overcome. Therefore, the system status has a special meaning for an educational system and supports the learner with the learning. When we are following Kort et al.'s model and considering heuristics, we focus where the heuristics are providing affective support and encouragement to affirm the learner. We gave each affective heuristic a short name and a longer definition. Table 6.8 shows the proposed set of affective heuristics for the usability evaluation of multimodal software for affective education, which is leveraging from Kort et al., Nielsen, Norman and Sankey's work.

The first two heuristics leverage Norman's visceral model; the next six heuristics leverage Norman's behavioral model, and the last two heuristics leverage Norman's reflective model. The heuristics are based on Nielsen's model [104], but our particular re-interpretation is enforced by Norman, Kort et al. and Sankey's work.

Table 6.8: The proposed affective heuristic evaluation.

Affective Design	**Heuristics**
Visceral	*Heuristic 8* **Aesthetic and minimalist design:** Learning should not contain information which is irrelevant or rarely needed. Every extra unit of information in learning competes with the relevant units of information and diminishes their learning ability. Including additional music or sounds if that is an essential component of the learning interaction. The use of video may be preferred for a lecture style presentation (supports Sankey's H6, H7, H12, H13, H15 and H16, and Kort et al.'s explore step).

Table 6.8: The proposed affective heuristic evaluation.

Affective Design	**Heuristics**
	Heuristic 10 **Help and documentation:** Even though it is better if the multimodal software can be used without documentation, it may be necessary to provide help and documentation. Any such information should be easy to search, focused on the student's task, list concrete steps to be carried out, and not be too large (supports Sankey's H14 and Kort et al.'s explore step).
Behavioral	*Heuristic 1* **Visibility of system status:** The multimodal software should always keep learners informed about what is going on, through appropriate feedback within reasonable time, so the learner knows how he is doing, and what to do to get better (support Kort et al.'s challenge and encourage step).
	Heuristic 3 **Learner control and freedom:** Provide the student with some control over the learning environment considering the emotional aspects, ensuring that the instructional strategy is made clear. Learners often choose system functions by mistake and will need a clearly marked "emergency exit" to leave the unwanted state without having to go through an extended dialogue. Support undo and redo. The learner has to have the freedom to try new things (supports Sankey's H4 and Kort et al.'s overcome and challenge step).

Table 6.8: The proposed affective heuristic evaluation.

Affective Design	Heuristics
	Heuristic 5 **Error prevention in learning:** Even better than good error messages is a careful design and learning objective providing cognitive strategies that prevent a problem from occurring in the first place. Either eliminate error-prone conditions or check for them and present learners with a confirmation option that the multiple sensory modalities provide before they commit to the action (support Kort et al.'s explore step).
	Heuristic 6 **Recognition rather than recall:** Minimize the learner's memory load by using the cognitive and affective strategies making objects, actions, and options visible. The learner should not have to remember information from one part of the dialogue to another. Instructions for use of the multimodal software should be visible or easily retrievable whenever appropriate (support Kort et al.'s explore step).
	Heuristic 7 **Flexibility and efficiency of use:** Acceleration strategy-unseen by the novice learner-may often speed up the interaction for the expert learner such that the multimodal software can cater to both inexperienced and experienced learners. Allow learners to tailor frequent actions (supports Sankey's H6, H7, H12, H13 and H15, and Kort et al.'s explore step).

Table 6.8: The proposed affective heuristic evaluation.

Affective Design	**Heuristics**
	Heuristic 9 **Help users recognize, diagnose, and recover from errors:** Teacher should monitor the students with the multiple sensory modalities and if students face errors, motivate and persuade students with the affective strategies. Error messages should be expressed with text or sound, precisely indicating the problem, and constructively suggest a solution, e.g. providing a video clip (support Kort et al.'s explore, overcome and affirm step).
Reflective	*Heuristic 2* **Match between multimodal software and the educational environment:** The multimodal software and the teacher should speak the learner's language using emotional and social factors, with words, texts, images and concepts familiar to the learner, rather than system-oriented terms. Follow real-world conventions, making information appear in a natural and logical order (supports Sankey's H2, and Kort et al.'s explore step).
	Heuristic 4 **Consistency and standards: Consistent and intuitive mapping** Users should not have to wonder whether different words, situations, or actions mean the same thing. Follow platform conventions (supports Kort et al.'s explore step).

6.5 Summary

In this chapter we focused on evaluation techniques and have proposed usability inspection methods. Our main claim is that issues of affect have not been addressed for multimodal software evaluation (i.e. heuristics and walkthroughs) in education. The challenge is how HCI evaluation inspection techniques can inform multimodal affective evaluation in education. We adapted the well-known evaluation techniques to the MADE framework, and we proposed the affective walkthrough and the affective heuristic evaluation to evaluate affective educational multimodal software. This chapter makes significant contributions in UI evaluation and usability inspection methods for affective education, and identifying usability issues in interactive educational multimodal software.

We conducted case studies on these inspection techniques which are described in Chapters 8 and 9. These case studies evaluate affective and cognitive aspects of the software and enable us to validate and refine the evaluation techniques. It is hoped that the techniques introduced in this chapter may identify software that leads to more student engagement with the learning materials.

Chapter 7

Case Study 1 – MADE Ratio: Affective Multimodal Software for Mathematical Concepts

7.1 Introduction

In this chapter we focus on interaction design and evaluation for multimodal software in affective education, and provide a case study of our MADE (Multimodal Affect for Design and Evaluation) framework presented in Chapter 4. In particular, we are utilizing the two design methods presented in Chapter 5, based on well-known methods from Goal-Directed Design (GDD) and Usage-Centered Design (UsageCD).

This chapter presents a case study of a tool to support mathematical learning. We applied the techniques that we described in Chapter 5. We have designed technology for providing students with an affective mathematical learning system understanding different ratios, including understanding equivalent proportions (e.g., the sequence of equivalent proportions 1:3, 2:6, 3:9, etc.). We then implement our design using hardware to detect hand positions. Finally, we recruited participants to test the usability of the system. Our findings were that the design methods were effective, but we also gained insight about how our system could be improved.

Students will control and maintain a vertical distance between the levels of their two hands to indicate a proportion. Figure 7.1 shows the prototype affective mathematical learning tool in use by a student.

As we mentioned in Chapter 4, the MADE framework is based on educators identifying a learning objective, and strategies for achieving it. The strategies will include cognitive, affective and multimodal elements.

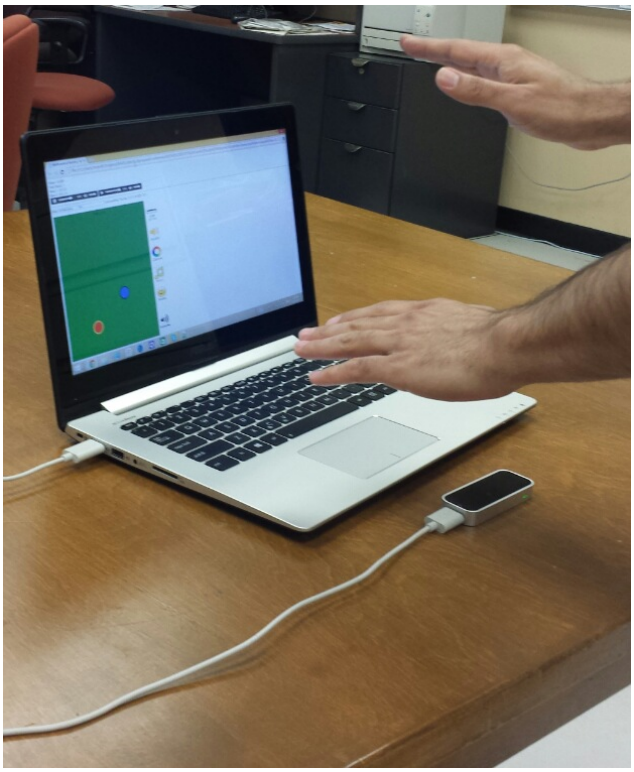


Figure 7.1: The affective mathematical learning system in use by a student.

For instance, a learning objective in mathematics can have a cognitive and an affective strategy. Also, multiple sensory modalities can support both a cognitive strategy and an affective strategy. The cognitive strategy is informing what the learning is, and affective strategy is how to make it fun and help to bring enthusiasm, as a learner is emotionally involved in the learning process. Multiple sensory modalities can help both. While interacting with educational software, a student can employ their sensory modalities (e.g. vision or audio); they may also need to employ some quasi-sensory modalities (e.g. embodiment, narrative or persuasion) while interacting with a supportive technology.

For our case study, we adopted the learning objective and multimodal strategy from the work of Howison, Trninic, Reinholz and Abrahamson [64]. Their learning objective was for students to understand the concept of proportionality. They created a simple Mathematical Imagery Trainer, to help students understand the proportional equivalence concept with a multimodal system to detect the positions of their hands using a Nintendo Wii controller and hand-held infrared emitters [1, 64]. We have taken this concept and used it as a starting

point. We applied our framework, and our design techniques to create an *affective* multimodal education system.

7.2 Design Process

We are using affective strategies because of students' cognitive difficulties in understanding mathematical proportions. We next illustrate the MADE strategy diagram for affective education and then show how we adapt the two design methodologies for affective multimodality education. Figure 7.2 shows a specific version of the MADE strategy diagram for our case study: the mathematical learning system. This is the example diagram we gave in Chapter 5 The teacher actor, the student actor and the developer actor are associated with the diagram. The first and second use cases in the left side are "Move both hands to correct ratio" and "Moving one hand higher, bigger distance and ratio" (cognitive strategies). For this case study, we chose several simple affective strategies: play praising audios, happy music, show encouraging messages, use different colors, and icons during exercise; and a similar multimodal strategy: use of motion sensor device to indicate correct distance. Each actor represents a role. As before, we assume that the teacher has to decide the learning objective, the cognitive strategy and the affective strategy, as well as having a monitoring role with the system. The software designer has to take this model into the affective learning system.

In this study the learners have to make and keep the screen in different colors by moving their left and right hands in regard to a specific ratio and distance between their hands. Now we describe the proposed adaptation of personas from Goal-Directed Design (GDD) for our mathematical learning system and the proposed adaptation of Usage-Centered Design (UsageCD) for affective education.

7.2.1 Goal-Directed Design

The first interaction design methodology we considered was GDD (explained in Chapter 5), created by Cooper et al. [26]. His GDD process results in a solid user model and a comprehensive user plan. It is a strong tool for answering questions like who are the users

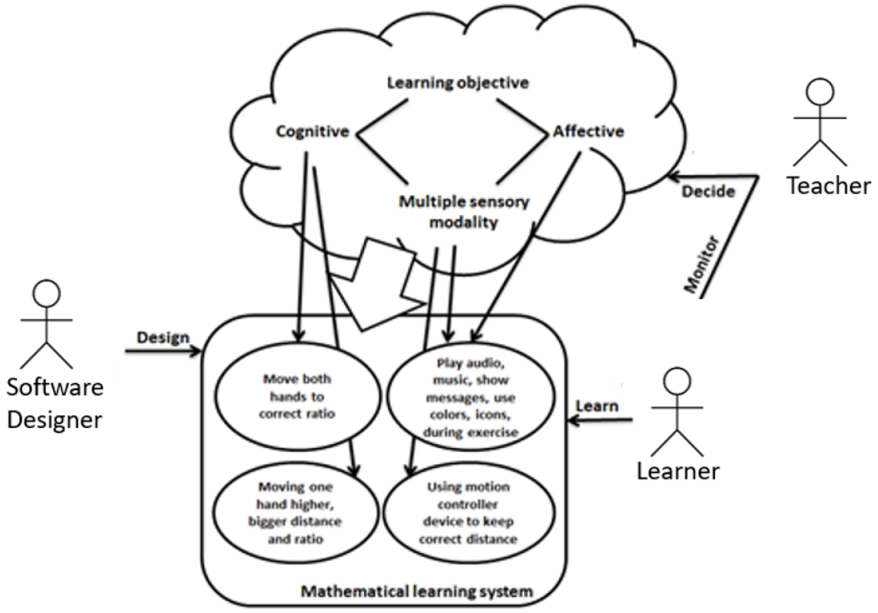


Figure 7.2: The MADE strategy diagram specific to the mathematical learning system.

and what they are trying to accomplish, or how users interact with the system and how the system should behave and deal with problems they may encounter. It includes personas, which are user models that represent a class or type of user of a specific practical interactive design tools to create an outside view and help give new perspective to high-tech applications. The personas do not describe real people, but are realistic and not idealized. They are rich descriptions of typical users of the system under development, which we as designers are focusing on.

Affective Personas: To create a system to satisfy a variety of users we are using affective personas as described in Section 5.3. For our multimodal affective educational case study, we use three personas, showing the emotional characteristics of the user; and considering affective and cognitive strategies with the help of the multiple sensory modalities to have a better learning environment. The personas are: one student whose parents would like him to become an engineer, one unconcerned student, and one more inclined student who has a reading disability (dyslexia).

1st Learner Persona: It is hard for Oliver to learn mathematics. He has family issues and worries about the future, as his parents' desire is he be an engineer. Therefore, he is afraid of not being successful. By using a 3D motion sensor device in math course materials that provides embodiment, Oliver can see how to interact with mathematical software and learn it using his hands. By bringing mathematical principles closer to the user, it offers the opportunity to help bring science and mathematical learning to life, and tries helping Oliver in understating better mathematics, and reducing his fear and stress. The teacher can encourage Oliver with affective strategies as well, such as humor and emotion, and using encouraging and praising audio, messages, icons and colors when correct ratios happen with the cognitive strategies. The art of persuasion and scaffolding in learning maths can make more engagement and brings focus, precision, motivation and encouragement (see Figure 7.3).

2nd Learner Persona: Badly designed interfaces are frustrating for Sarah in learning mathematics, who likes to listen to music, text messaging, dance and go clubbing. By having course materials with embodied interaction using motion sensor device for embodiment, and having happy audio, icons and colors results in learning that is more interesting with less distraction, and persuades the learner to change behavior attitudes. Affective strategies can be humor and emotion using praising audio, happy sound effects, different colors, happy icons and a pleasing interface when correct ratios happen. The cognitive strategy is embodied cognition. Sensory modalities are audio, music, icon, color and the motion sensor device. Quasi-sensory modalities are embodiment and persuasion. By considering these strategies the teacher can bring more engagement, attention, and pleasure to the learner (see Figure 7.4).

3rd Learner Persona: Mike has Dyslexia and has difficulty to read or interpret words, letters, and other symbols. With this motion sensor Mike has the possibility to control some course material just with a wave of his hands using this 3D modeling software. He likes to study; he is ambitious, but he feels shy in the social life, which is in this context a classroom environment. He has frustrations over expressing course materials in the classroom and he is not comfortable. He worries that students may make fun of him, and laugh

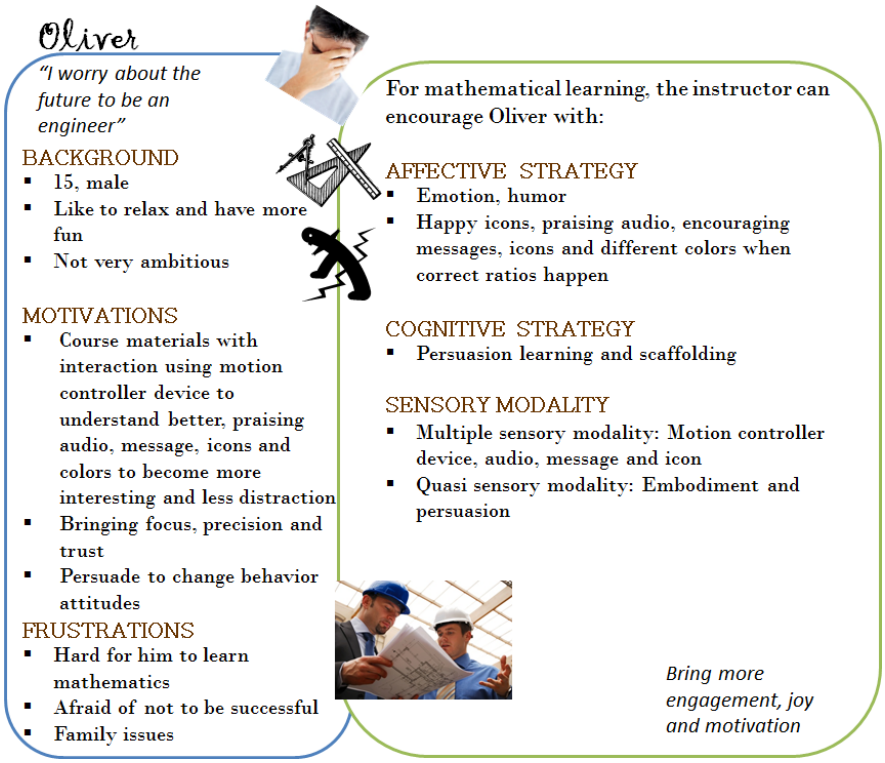


Figure 7.3: 1st learner persona in mathematical learning system.

about his misfortunes (superiority theory in humor [41]). The teacher can encourage Mike with affective strategies like emotion, humor, personality and social, encouraging audio and messages, and cognitive strategies, for example persuasion learning and situated cognition. By using multiple sensory modalities including motion sensing, audio, messages and icons, and also quasi-sensory modalities such as embodiment and persuasion, the teacher can bring more comfort and less stress to Mike (see Figure 7.5).

All these personas identify key affective issues relating to users. For our system design, the next step is to focus on interaction, where we use our adaptation of essential use cases from Usage Centered Design.

7.2.2 Usage-Centered Design

Usage-Centered Design (UsageCD) was introduced and developed by Constantine and Lockwood [25]. It is based on user intentions and usage patterns for user interface design; usage

Figure 7.4: 2nd learner persona in mathematical learning system.

is how and for what ends software tools will be employed. It analyzes users' roles in relation to systems, employing abstract (essential) use cases for task analysis [25].

Affective Essential Use Case: Constantine and Lockwood developed essential use cases (EUCs) [25] (explained in Chapter 5). These are simplified, abstract, generalized use cases, to compensate for the limitations of both scenarios and use cases. They capture the essence of the use case, which means you focus on what the result is meant to be and not how you achieve it. Therefore, the advantage of it is that we document "WHAT" the outcome of the use case is meant to be and leave it to the designer to come up with "HOW". Later, we have to see the different ways to implement it [9, 25]. We now employ the adaptation of essential use cases that we described in Section 5.5.

Each use case has an affective objective. We add a part called the "Affective sidebar" to the EUC card, and have a column in the right side called "Affective status" (see Figure 7.6). Thus, the outcome of the use case basically should involve from the teacher, cognitive outcomes as well as affective outcomes. Here our focus is education. With affective multimodal strategies

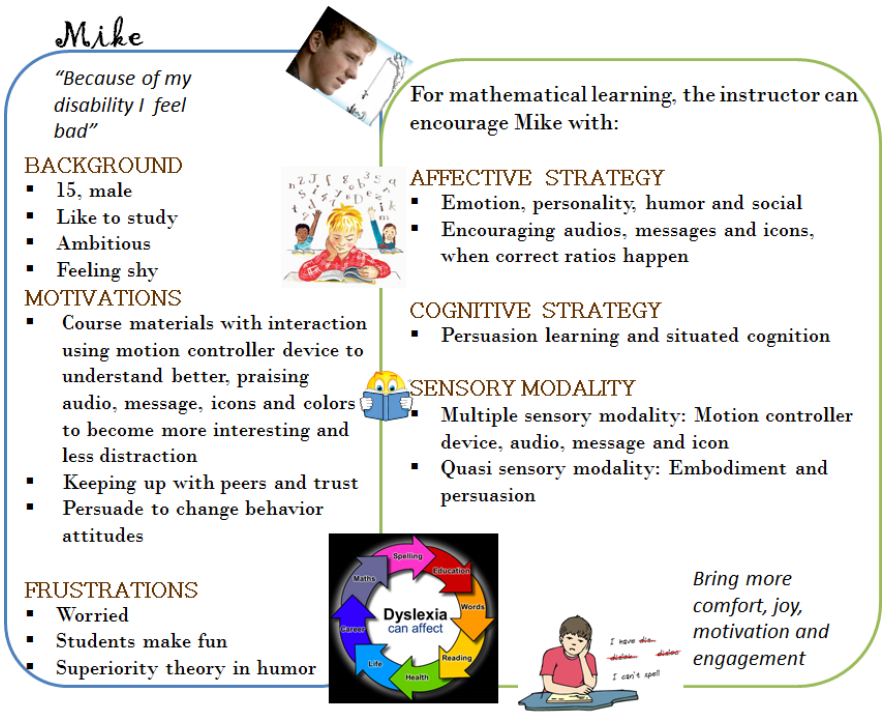


Figure 7.5: 3rd learner persona in mathematical learning system.

that the mathematical learning system provides and transmits, the teacher can use different kinds of strategies such as affective messages, audio, and colors for the learning objective. The system's strategy is to persuade, motivate, and be encouraging in learning math, and brings a good feeling to the students to make them feel it is fun.

Our case study was to design a system to help learn the concept of a mathematical ratio. The modality involves embodied cognition to support pedagogy, and is based on the work of Howison, Trninic, and Abrahamson [64].

This design should increase the proprioceptive, kinesthetic and episodic memory experience, as the controller that we use requires physical motion of the arms and the hands. The proprioceptive memory involved will aid remembering how to perform intended physical actions in the future. The learning takes place by the students carrying out a physical activity by moving their hands and body movements to interact with the UI. By using affective strategies we can increase the episodic memory, which actually happens to learners; the learner will think about the time back when they did this with their hands. The system

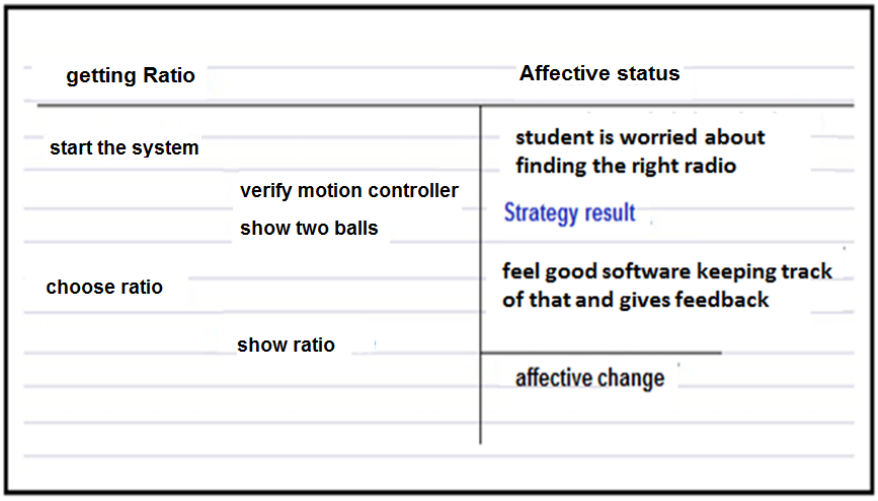


Figure 7.6: An EUC card using the affective strategy.

supports discovery learning and exploration using sensory modalities.

In our system, one of the tasks is that learners have to make a requested ratio. Feedback is provided by showing the screen in different colors and giving persuasive feedback i.e. praising messages, audio, icons or images. The learners move their left and right hands in regard to a specific ratio and distance between their hands.

We used two design methodologies, *Affective Personas* and *Affective Essential Use Cases*. Affective Personas identify key aspects of the emotional state of users, and Affective Essential Use Cases describe goals for emotional aspects of user experience to be supported by the system. We described these design processes in detail above.

7.3 Low Fidelity Prototype

For designing the mathematical learning interface, we used the Balsamic Mockups software (https://balsamiq.com/). Figure 7.7 shows the prototype layout, which is explained in the following:

- Top bar: provides us with the number of hands, shows position of each hand, ratio, affective feedback (audio, message and color), and if it is left, right or both hands detected above the motion sensor.

- Main window: in the main part of the screen there is a gray box with two circles representing each hand. And the circles move, as the student moves his/her hands up and down.

- Right side bar: on this side there are controls for the motion sensor, audio, icons, colors, music, and volume being able to select colors, audio, etc.

- Bottom bar: shows the area we interact with the system; the motion controller and keyboard.

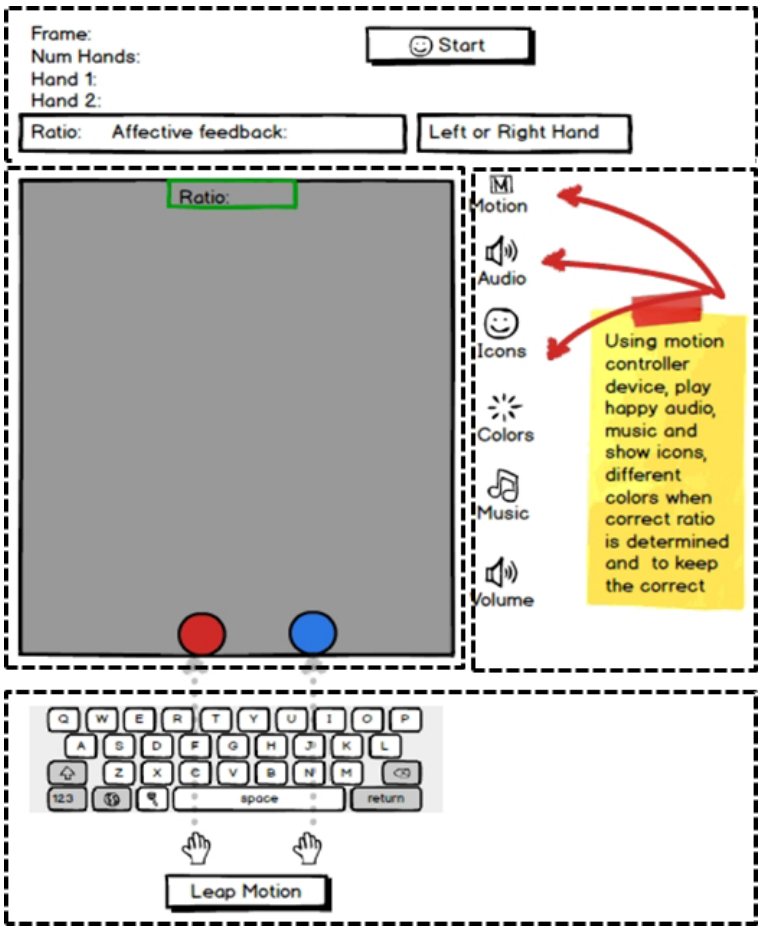


Figure 7.7: Sketch-style wireframes for the mathematical learning system at the starting point.

Figure 7.8 shows when the system is being used, and it matches the ratio, there are affective feedbacks: the color turns into yellow, audio plays an encouragement sound by praising the user, and a message shows "Well Done!" above the screen box.

There are settings on the right, so the learner can customize the various affective features such as sound, music, text and images. We would also expect that teachers would be able to select these settings on the basis of their knowledge of the learners.

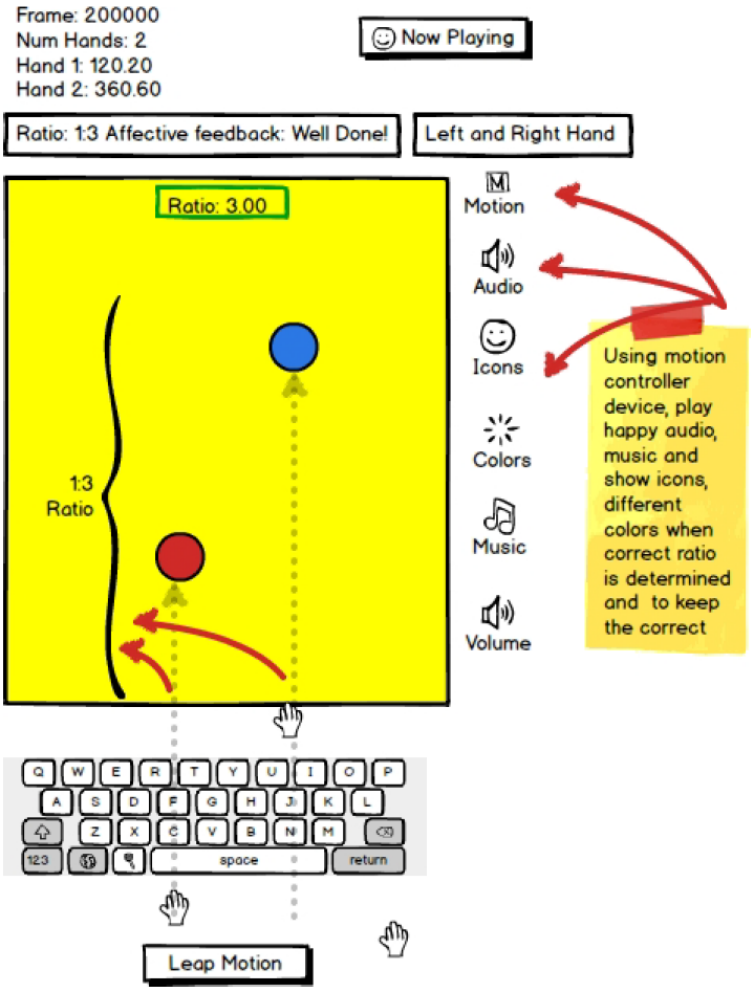


Figure 7.8: Sketch-style wireframes for the mathematical learning system in use with ratio 1:3.

7.4 Implementation

In this case study, we are using an ease of use, low-cost hand-motion tracking controller: Leap Motion sensor (https://www.leapmotion.com/) to design mathematical learning software using the MADE framework (see Figure 7.9). The Leap Motion device includes two cameras and infrared LEDs under a black glass. It enables the software to track two hands

finger movements when they are moved over the device. Several HCI research studies and applications have used this 3D controller as an input device for tracking accurate hand movements and to implement gesture-based and tangible user interfaces [19]. We are extending and considering a new application area: mathematics education.

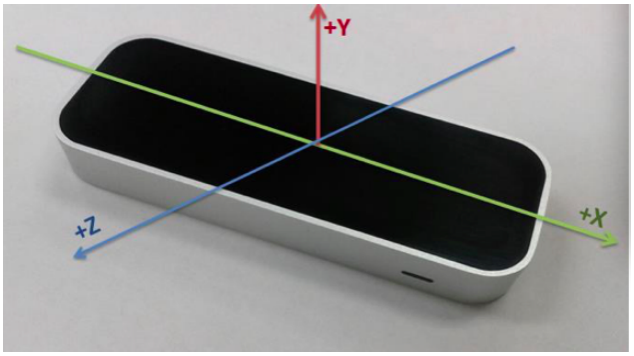


Figure 7.9: The Leap Motion Controller.

Figures 7.10 and 7.11 show the initial version of our mathematical learning environment. At the start (see Figure 7.10), we show how a user is interacting with a Leap Motion sensor device using his/her hands moving them to fix 1:3 proportional progression having audio and background color to indicate their accuracy, and to have an emotional influence on the learner. In Figure 7.11 a student who has moved both hands up and down along the screen (does the sequence of equivalent proportions 1:3, 2:6) has achieved the ratio 2.96, and gets affective feedback: an audio and a message saying "Well Done!" and the background screen changes to yellow.

Figure 7.12 shows a later version of our system at the point when the ratio task been completed. It shows a student has completed a task (here a specific proportion of 2:6) with customized affective feedback, which include: change of the background screen to green, and on the left side of the screen, some happy icons, encouraging stickers, affirmation message, and a clapping sound. Although the design is improved, a need for graphic art skill is still obvious, but is beyond the scope of our study.

For developing the software we used JavaScript and the Raphael graphics library (`http://raphaeljs.com/`) as well as PHP and MySQL.

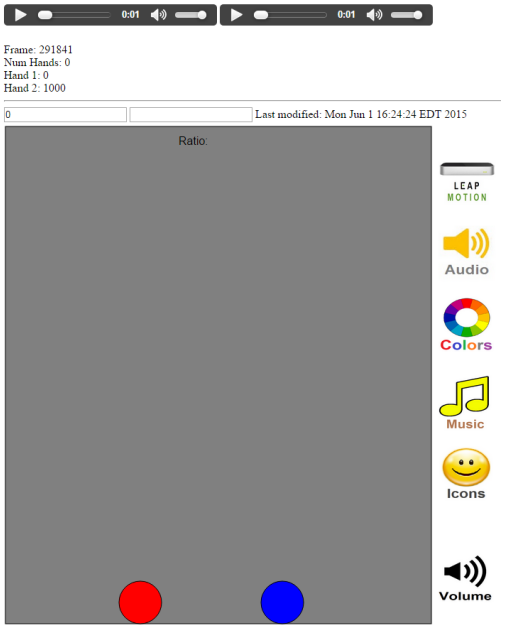


Figure 7.10: At the starting point, the circles representing the user hands are at the bottom, and the background color is gray.

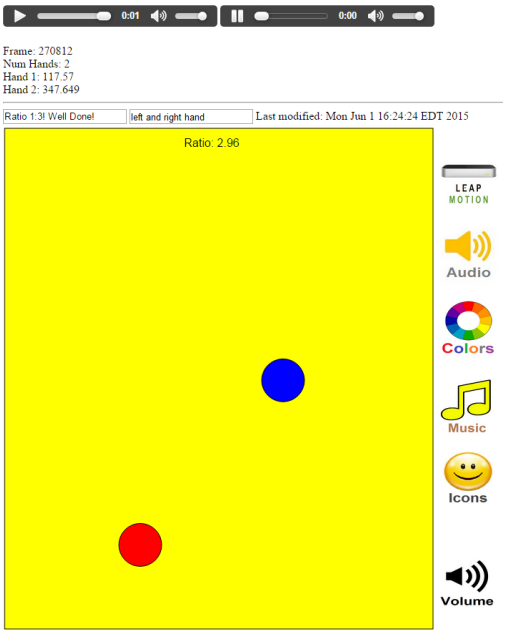


Figure 7.11: By moving the right hand higher searching for equivalent proportions 1:3 and 2:6; a ratio of 2.96 is reached, resulting in affective feedback of audio, a message, and screen color change.

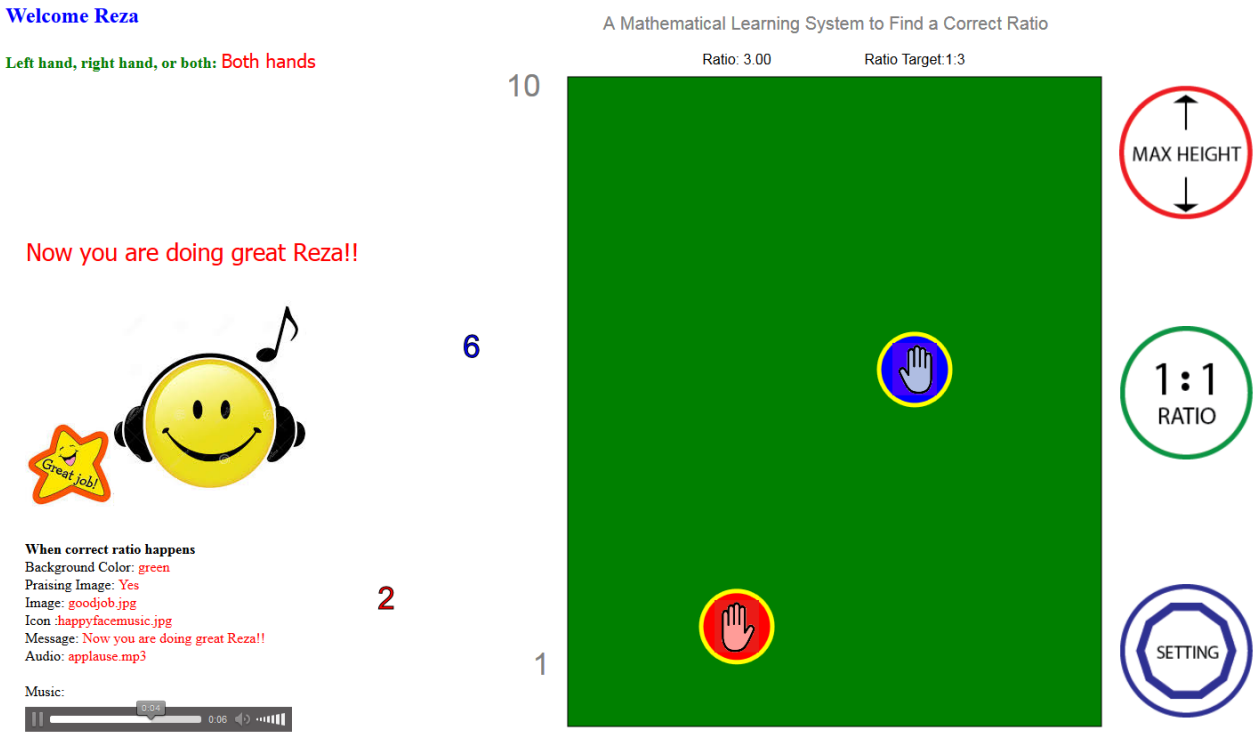


Figure 7.12: A student sitting at a desk is moving the right hand higher from the desk searching for equivalent proportions 1 to 3 or 2 to 6; a ratio of 1:3 is reached, maintaining the goal proportion and resulting in affective feedback of audio, a message, icon, image and screen color change.

7.5 Study

After development, we conducted a usability study of our system. Below we explain the study, talk about the expectations, our participants, the equipment, the procedure, and the analysis methods. The study was approved by our university research ethics board (see Appendix B). We were not doing formal hypothesis testing, but simply doing usability testing and comparing the results with our expectations. Our expectations were as follows:

1. The Leap Motion Controller is easy and engaging
2. Fun results from the physical interaction
3. The affective design supports learning

Participants

Eleven undergraduate and graduate students (3 females and 8 males) from two universities in Canada volunteered to participate in this study. Some of the participants had experience in teaching mathematics. They ranged in age from 21 to 60, with a mean age of 33 (S.D. = 12). All participants were able to use their hands and arms freely and had ability to work in English. All participants reported using computers daily, and most participants (85%) reported using them for 11-20 years. All participants stated they had used pencil and paper, or mouse and keyboard, when they learned about ratios while in school. Seven participants were right-handed, three were left-handed, and one was ambidextrous.

Equipment

A Leap Motion Controller, a Dell desktop computer with a 22-in high-resolution LCD monitor, and two speakers were used. We used an audio recorder to capture the participant's voice comments.

Procedure

The study took thirty minutes for each participant. We taught the participants how to interact with the system using their hands before the study, for five minutes. We used a think aloud protocol, and did audio recording. The main part of the procedure had four steps:

1. Training phase: We had a five minute testing phase, providing a short training exercise for the user before starting with the ratio task, asking them to have their both hands at position 5 (on a scale from 1 to 10), and then space left-hand and right-hand four units apart to get familiar with the system.

2. Testing phase: After the training phase we had a testing phase of fifteen minutes. First they had to do the default tasks and see the affective feedback we provided them,

and we monitored their reactions. Next they would customize their settings, and we monitored their reactions and saw their preferences.

3. A Usability questionnaire that we explain in the “Quantitative Results" section, below.

4. An Open-ended questionnaire that we explain in the “Qualitative Results" section, below.

We collected both qualitative and quantitative data at the same stage of the study. We used thematic analysis for the qualitative aspects of the study. We collected the data from comments and answers to open ended questions, went through it all, and identified themes.

7.6 Quantitative Results

First we address our stated expectations:

1. The Leap Motion Controller is easy and engaging: Almost all participants (82%) agreed that there is no fatigue involved while interacting with the UI.

2. Fun in the physical interaction: Regarding the pleasantness and how pleasant is to interact with the UI, 91% of participants agreed that using Leap Motion Controller would be more engaging than using mouse and keyboard.

3. Affective design support of learning: Regarding the affective strategies 72% of participants agreed that it would help in motivating students in learning ratio.

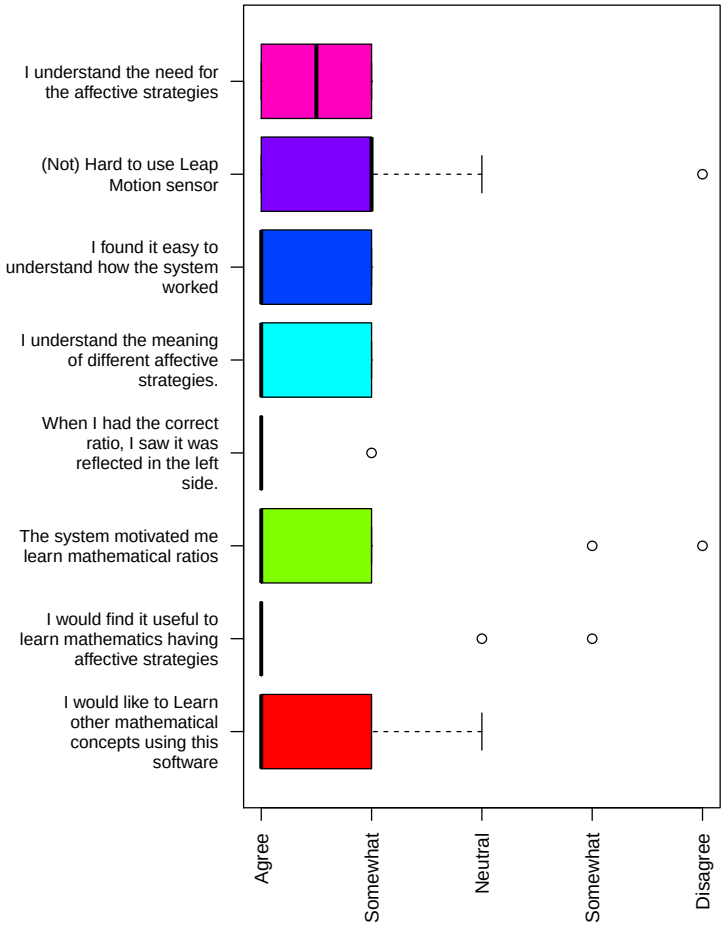


Figure 7.13: Usability questionnaire about the MADE Ratio system.

Figure 7.13 shows the usability questionnaire results. On the left side there are eight statements. The participants were given a questionnaire to evaluate and rate their experience and subjective impressions of the system. For each question, subjects were given a Likert scale between agree and disagree. For analysis we used *R* (`https://www.r-project.org/`). The results of this analysis are shown in Figure 7.13. The responses are shown as boxplots, where the inner quartiles are shown as colored boxes, the outer quartiles with whiskers, and outliers as circles.

The first statement is about understanding the need for affective strategies in this learning system; all the participants agreed. For the second statement (Not hard to use Leap Motion

Controller) mostly everyone agreed. By providing a training phase at the beginning of the study, everyone understood how the system worked. The forth statement was understanding the meaning of different affective strategies, which everyone understood. The fifth statement was about task affirmation and reflection on the UI, and everyone but one agreed. The sixth statement explore whether the system motivated the participant in learning about mathematical ratios, but not all agreed. The seventh statement was whether it is useful to use affective strategies in learning mathematics; only two of the participants were neutral but the rest agreed. The last statement was if the participant would like to learn other mathematical concepts using this software; almost everyone agreed except one participant somewhat agreeing, and another was neutral.

Overall, the rating indicated approval of the system design; however, some differences were observed among participants.

7.7 Qualitative Results

At first we did a pilot study with one participant to capture firsthand behaviors and interactions that might have not been noticed. We found some problems with the system, and fixed them for the real study. One issue was that the Leap Motion Controller was far from the participant, which meant the participant was not feeling comfortable and had some fatigue with his hand. Having it closer made it easier to hold up his hands, giving a finer control over the height making it easier to use with no fatigue. Another concern was the participant said that the text wording could be improved in a lot of places, and we fixed all those instances. Another issue was brought up was the audio. He said: *"I'm not a fan of audio in general, but the applause seemed over the top, and the music was in the way because we were talking."*, so we provided controls for the audio.

7.7.1 Multiple Sensory Modalities

Fun and Engaging: Almost all the participants agreed this system would be fun and engaging for children. They said the Leap Motion was easy to use and quite fun to try for

the first time. One participant said: *"I personally have cousins who are currently learning ratios, and believe this would be an interesting way for them to learn, that they would enjoy."* Another participant said: *"I think having Leap Motion is an effective and innovative approach for kids to be engaged in learning mathematics. The device is small, light and easy to use and having a friendly interface would be successful in my opinion for learning."* Another participant said this system is great for allowing students to respond in the modality that they have the most comfort with. He said the Leap Motion is good for visualization representations for how a ratio scales. He had a suggestion that there be option to interact in a horizontal way as well, suggesting that would more directly correspond to hand positions.

Easy to Use But Could be Improved: One participant said the Leap Motion is mostly easy to use but it is a bit finicky. Some participants mentioned there should be a physical reference such as a ruler displayed vertically. Another said it would be nice to have a ruler (scale) guide on the side as reference points to help user more accurately move up and down.

Learnability and Memorability: One participant said that the Leap Motion would help with learnability and memory, because you've experienced it yourself. Another participant said it was very smooth and there was no lag-time when interacting with the Leap Motion.

Control of Audio: One participant said the audio output would be suitable as long as it is controllable. Another participant mentioned: *"While performing a task, if the music is playing, it might be distracting for some people, but the praise (e.g. clapping) when the job was complete can be a good thing."*

7.7.2 Affective Strategies

Effect on Learning: One participant said this multimodal system had benefits by providing multiple interaction modes including audio, visual and using an embodied interaction approach; the Leap Motion can make UI learnable and memorable. One participant said: *"I feel proud of myself when I hear applause."* He said that the happy face is encouraging,

audio is motivating, but he did not like the specific music in the background, and he said it depends on type of music. Another participant said: *"The affective strategies are great for kids, and it's very encouraging."* Another participant said: *"The affective strategies like music are helpful in learning, since it challenges the students to try, and in the meantime it presents a mathematical problem like a computer game. This interactive approach is helpful since it draws students attention more than traditional learning techniques."* Most of the participants said this system could be very good for children, e.g. music, audio, color, text and using humor and funny things can be engaging for kids. The participants explored the features of the program such as customizing the settings and exploring the effects. One participant said: *"I like the idea of giving different types of feedback and customizing that feedback."* She said that affective strategies present in this system are appropriate for kids but not for adults.

Challenges: Some participants would have liked if the system provided the learner with a score. This possible addition could be a mechanism to make the learning seem more like a game that would keep the students' attention and challenge them. One participant suggested that there be continuity to the music during the task (e.g. each part of the task for 30 seconds, and eventually music finishes when the person completes the whole task and it is done. Another participant suggested that when a user wants to start using the system, it can ask the user their favorite color for the background choice. Some participants had concerns for the location of content on the screen. Some liked the affective feedback's location on the screen, but they said it should not go on top of the learning objective. Some participants suggest the ratio can move closer to the hand circles on the main screen as maybe a better idea.

Virtual Characters: A participant mentioned adding virtual characters; an avatar of the learner. For example, there is an appeal to emotional investments of virtual characters like Pokemon, and we could provide for the student to choose an avatar that reacts to their performance, and lets them earn token currency as a reward for success – which could be spend to customize the avatar – and allow custom themes for success/fail sounds, a custom look and feel theme to the interface as personalization.

7.7.3 Facial Recognition

We suggested that our system might use analysis of facial expression to detect emotional state, and asked for their comments. After this case study, we explored this idea further, as we describe in Chapter 10.

Useful for teaching: The participants agreed with having a facial recognition with the system, and they said it might be a great way for adapting difficulty to success/failure e.g. frustrations vs. excitement. One participant said: *"I think that facial recognition should be used, as it would allow the teacher to be able to monitor individual performance of students, to know when to modify the difficulty or focus on a skill that doesn't seem to be understood."* This approach might be useful to help better understand the students' moods and feelings and how they feel about solving mathematical problems. Teachers might then adapt the exercise based on the learner's frustration with difficulty. If a student feels that they are behind, facial expressions would help teacher to adjust/modify level of difficulty. It might be interesting data to collect for a teacher to reflect back on and use in their assessment of teaching methods, and assessing their student's performance.

Privacy issues: At least two participants said: "Showing the student's face in the screen would be distracting." One of these participants mentioned people may not like capturing their face (privacy) and added: *"Probably not? It would have to be done very well, so that it's not really telling you how you feel. [There is a] danger of being creepy monitoring. How would it improve my experience? and is it going to be worth the costs?"* Another participant said that she is not sure how useful it would be. Almost all the participants agreed that it would be very interesting for the teacher to use facial expression recognition. Only one of the participants said: *"Would there not be some privacy concerns here? I personally would not like it. Some people do not like to express or have their emotions show/being monitored. E.g. with frustration/anger/sad. Learning environment should be more visual/interactive rather than computer monitoring; the human touch is needed here"*. Therefore, they felt it could be useful for the teacher to see the video, but it would be more distracting if there was a real time video on the same page, and some people might not like having their image

recorded.

7.7.4 General Advantages

As one participant mentioned, this system has a lot of potential for some people who have problems focusing on written material (e.g. ADHD). He added that this system provides a visual, auditory and physical interaction which is the environment they learn best in. He said: *"I find it very effective for my age and personality: I don't like voice; I'm more visual."* He would like the system to somehow show the movements of his hands on the screen. Other participants agreed on the following general advantages:

1. This system is fun and is more engaging compared to the traditional way of learning ratios.
2. Customization in this system is an advantage.
3. The system provides a hands-on physical interaction.
4. The system would raise the curiosity of children in learning.

7.7.5 General Disadvantages

"The User has to focus on each the left and right side of the screen. It took his attention and made he look at different directions while looking at numbers and make calculations." This was said by one participant, who suggested that more coherence on the main screen. Another participant said: *"If you could combine the points of attention; you have three-points of attention: our hands, hands shown on the screen and ratios on the left, which causes three focal points. Combine them into one."* This issue was perhaps the most prominent disadvantage; the entire list was:

1. Potential problem of different places to focus.
2. More visual cues about what is going on.

3. Need for challenges and game-like structure.

4. Potential problem with motor skills hand-eye coordination.

7.8 Discussion

We have described our rationale for, and implementation of, the MADE Ratio system, including our design-based research approach and reporting on qualitative and quantitative studies with eleven undergraduate and graduate students who engaged in problem-solving tasks with the MADE Ratio. Our study is consistent with the work of Howison, Trninic, and Abrahamson [64], while they have not emphasized the multimodality aspect and affective strategies. Our tasks were performed by the participants who had prior experience with ratios and proportions, but not in a multimodal environment. They had learned mathematical concepts with pencil and paper, or using mouse and keyboard.

The math topics we considered in this system were ratios and proportions. We tried to implement a multimodal system with a super-simple affective experience to increase engagement and attention of the students. This system is customizable, meaning the teacher can select different learning objective, affective and cognitive strategies in the database for different personas. The teacher monitors the students understanding of the material and provides different learning strategies to each student. These strategies employ sensory and quasi-sensory modalities, which we explained earlier.

Reflecting on our experience, we now suggest that the system might also have elements of *persuasive technology tools*, as described by B.J. Fogg [49]. This approach is consistent with the cognitive and affective strategies discussed in Chapter 3, but in our case study we had not considered them. Specifically, the system might be seen to be using *tailoring*, which is persuasion through customization (the teacher is able to customize the database according the each student's interest), as well as *reduction* approach to make the complex ratio task simpler (by using the Leap Motion Controller) to increase a student's motivation to engage in the task more frequently. Fogg has proposed that tailored information is more effective than generic information in changing attitudes and behaviors; the information provided by

a system will be more persuasive if it is tailored to the individual's needs, interests, or other factors relevant to the individual. We also discussed affect and so we might be seen as employing Pathos, from Aristotle's modes of rhetoric. Pathos represents an appeal to the emotions of the students, and elicits feelings that already reside in them.

The reflections of our participants also leads us to consider *gamification in education* as a potential new approach we might explore in refining the system.

7.9 Summary

This chapter focused on a case study on UI design and implementation for multimodal software in affective education. In particular, the purpose of this chapter was to describe a design approach for an emotional multimodal education case study considering the affective and cognitive strategies. The ability to communicate emotionally and cognitively plays an essential role in HCI and education. Our main claim is that issues of affect in multimodal software have not been addressed for software design for education. The challenge is how theoretical models of HCI can inform multimodal affective design in learning. For the design guidelines, we adapted well-known design methodologies explained in Chapter 5 to the MADE framework, and we proposed the adaptation of personas from GDD and propose adaptation of UsageCD and essential use cases for multimodal affective education.

We applied the proposed MADE design methodologies to mathematical learning case study that investigates the use of affective learning in multimodal software, trying to increase the engagement in students with learning materials.

The ability to communicate emotionally and cognitively plays an essential role in HCI and education. This chapter focused on UI design and implementation for a multimodal system in affective education, and conducted usability testing. We presented the rationale, design, implementation, and early results from our study of a novel educational technology. The goal of this study was to determine if this system can motivate and engage students in learning ratios and proportions. The affective strategies and the Leap Motion Controller were well suited for this environment.

We validated our three expectations: The Leap Motion Controller was easy and engaging, the system did support fun through the physical interaction, and the affective design support will increase the motivation in learning and the students did better understand some tasks in mathematics.

Our next steps will consider the following potential improvements. First, the interface should show a list of the tasks for the students to attempt, rather than relying on a teacher. Second, the system should show the student their learning progress. Third, and perhaps the most important, the system could allow the teacher to monitor and see current student status and how he/she is doing, showing the affective status, and allowing changes to be made to the task list and the affective feedback strategies. We might also explore detection of emotion using analysis of facial expressions. All these issues are addressed in Chapter 10.

Chapter 8

Case Study 2 – Museum Explorer: Evaluating Software for Affective Education

8.1 Introduction

This chapter addresses the use of proposed heuristic evaluation and walkthrough methods proposed in Chapter 6 to evaluate affect in the user interface of an existing educational multimodal software system.

A case study is presented where these methods are used to evaluate a Museum Explorer system. This system was designed to support an educational objective of making the users more engaged and reflective, by supporting the use narrative to help the users tell a visual story. We recruited participants to apply the evaluation methods to the Museum Explorer system.

Our methods were adapted from usability *inspection* methods. These rely on comparing system features with established principles, rather than the results of empirical testing. This is therefore less accurate in specific cases, but much easier and cheaper to apply.

Our findings were that the evaluation methods were effective, but we observed a number of effects that suggested necessary improvements. For example, the emphasis in the methods had to be changed to avoid a focus on simple usability issues. Moreover, we found that the evaluators had to be clear on the affective teaching objectives and on how the software design was intended to support them. We conclude with revisions.

8.2 Multimodality, Affective Education with Narrative

Interactive narratives are a form of digital entertainment that allow users of software to interact and change stories according to their own desires and participate in a collaborative experience [35]. In a multimodal learning environment, educational elements can be presented in more than one sensory mode such as audio, music, text, picture, etc. to overcome the limitations of each modality alone. The combination of narrative, audio, picture and text supports proficient communicative behavior and leverages natural human capabilities to communicate and bring pattern recognition and classification methods [144]. Many multimodal learning environments have been developed, but there has been little work on evaluating their design [82]. In Chapter 6 we proposed a heuristic evaluation and walkthrough technique for evaluating multimodal educational software by considering emotional aspects of the learners. Our evaluation techniques are based on our model (Chapter 4) that was inspired by *Bloom's taxonomy* [10]. We adapted the three domains (cognitive, affective and psychomotor), we considered the *multiple sensory* and *quasi-sensory modalities*, and our particular focus is on the affective domain. In this chapter we examine our proposed evaluation techniques with a case study of software for exploring a museum, and then we refine the technique.

8.3 Overview of Museum Explorer

The idea is to evaluate the effects of the modality that is being used in the Museum Explorer, narrative visualization. We are particularly interested in evaluating how effective the modality is in achieving the affective educational objectives.

Our case study is of software system designed to apply narrative to help students visiting a museum. This software, Museum Explorer, was developed by Jesse Gerroir as part of a Master's thesis [3]. This system was designed for making visits to museums better by supporting narrative visualizations including dramatic, categorical, sequential and slideshow approaches. Narrative should support continuity, storytelling and excitement [55]. The software supports an objective, common in relationship to field trips, to help students engage more with the subject matter by planning and reflection. This is a typical objective for

teachers, because while students enjoy field trips, they are often unprepared and later they completely forget detail. As a teacher, before taking the students and going to the museum, one helps them learn about the museum and plan the visit, which means when they are at the museum, they think more deeply about what is there. Therefore, one plans an itinerary and adds exhibits to the itinerary, as seen in Figure 8.1. When one comes back to the school, they do various activities based on what they learned. The Museum Explorer is a system to support this process. The software designer's perspective on this issue was to build a system to help achieve these objectives using visual stories.

Figure 8.1 shows the Museum Explorer. On the top is a map of the visited areas and on the bottom is the visualization selection, including: the *Slideshow*, which is meant to display pictures in a way that is commonly used as a presentation, the *Categorical,* where information is arranged by the topic of the locations, the *Sequential,* which focused around giving a sense of time and place, and the *Dramatic*, to give a sense of the user's personal experience. The Museum Explorer depicts exhibit areas from the Canadian Museum of Nature in Ottawa, Canada. The design explores different visualizations to support narrative storytelling and comprehension.

This figure also shows the visualization selections with the choices shown along the bottom. The left visualization is the Slideshow, which is meant to display pictures in a way that is commonly used as a presentation. The next visualization is the Categorical, where information is arranged by the topic of the locations. The one after is the Sequential, which focused around giving a sense of time and place. The last visualization is the Dramatic, to give a sense of the user's personal experience. Figure 8.2 shows the first part of the visualization: view your visit. It includes the map and a timeline of exhibits visited. The map is static and simply displays the exhibits the user visited in a colour coded fashion. And the timeline shows the time user spent in a particular exhibit.

Gerroir evaluated the system and found the following results: the dramatic visualization ranked very highly for the majority of the ranked questions and seemed to be the one most liked overall. The categorical visualization ranked positively overall and was highest for learning. The sequential narrative visualization ranked neutral, being neither the highest nor lowest ranked visualization. And the slideshow ranked last in the qualitative satisfaction

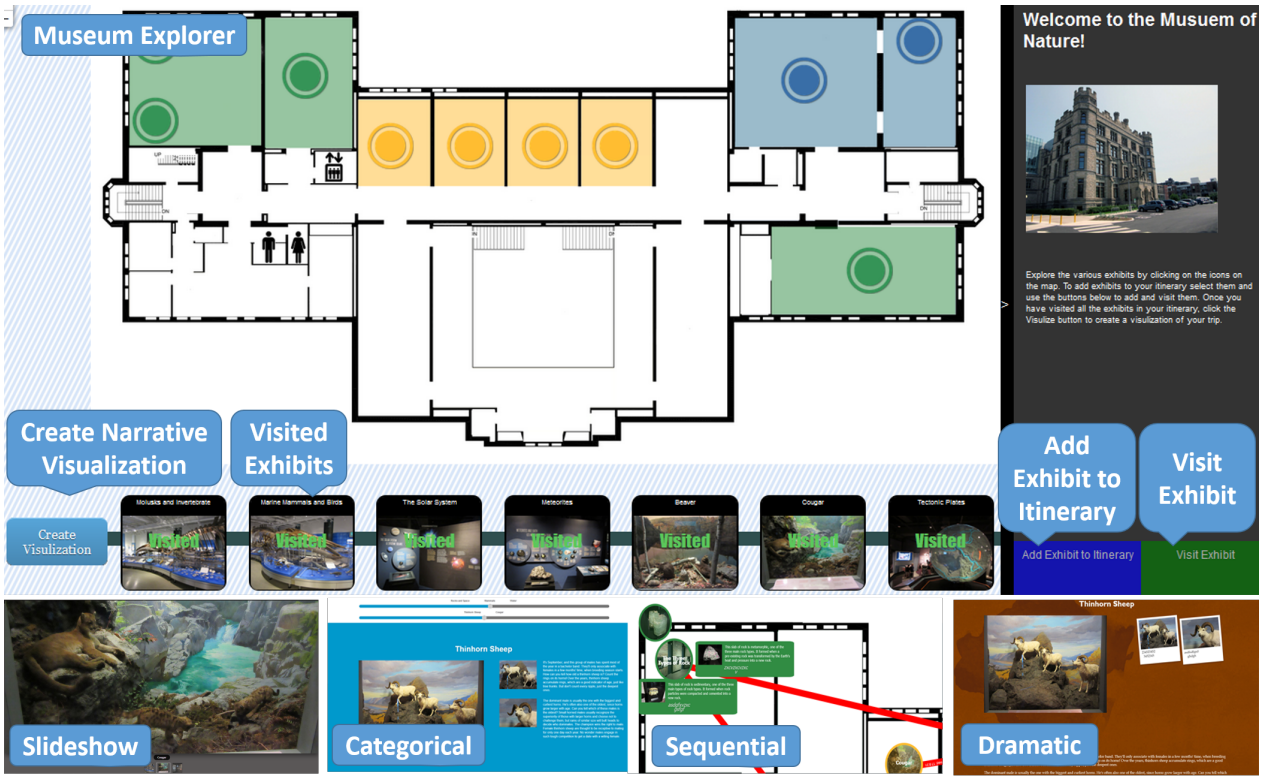


Figure 8.1: The Museum Explorer with the four visualization styles.

aspects.

We are using Gerroir's software as a case study for our MADE evaluation techniques. We wished to explore how our techniques might give insight about the design of the software and how it might be improved. In particular, we wish to explore the effect of the narrative modality in the educational context. Gerroir gave four different narrative visualization styles for the museum experience, but did not specifically evaluate the objective of making the users more engaged and reflective, thinking about what they learned, or the user's affective experience.

8.4 Museum Explorer and Affective Walkthrough

In the first step, we use this software as a case study for our *Affective Walkthrough*. This is based on Wharton et al.'s cognitive walkthrough [150], and Dormann and Biddle's affective

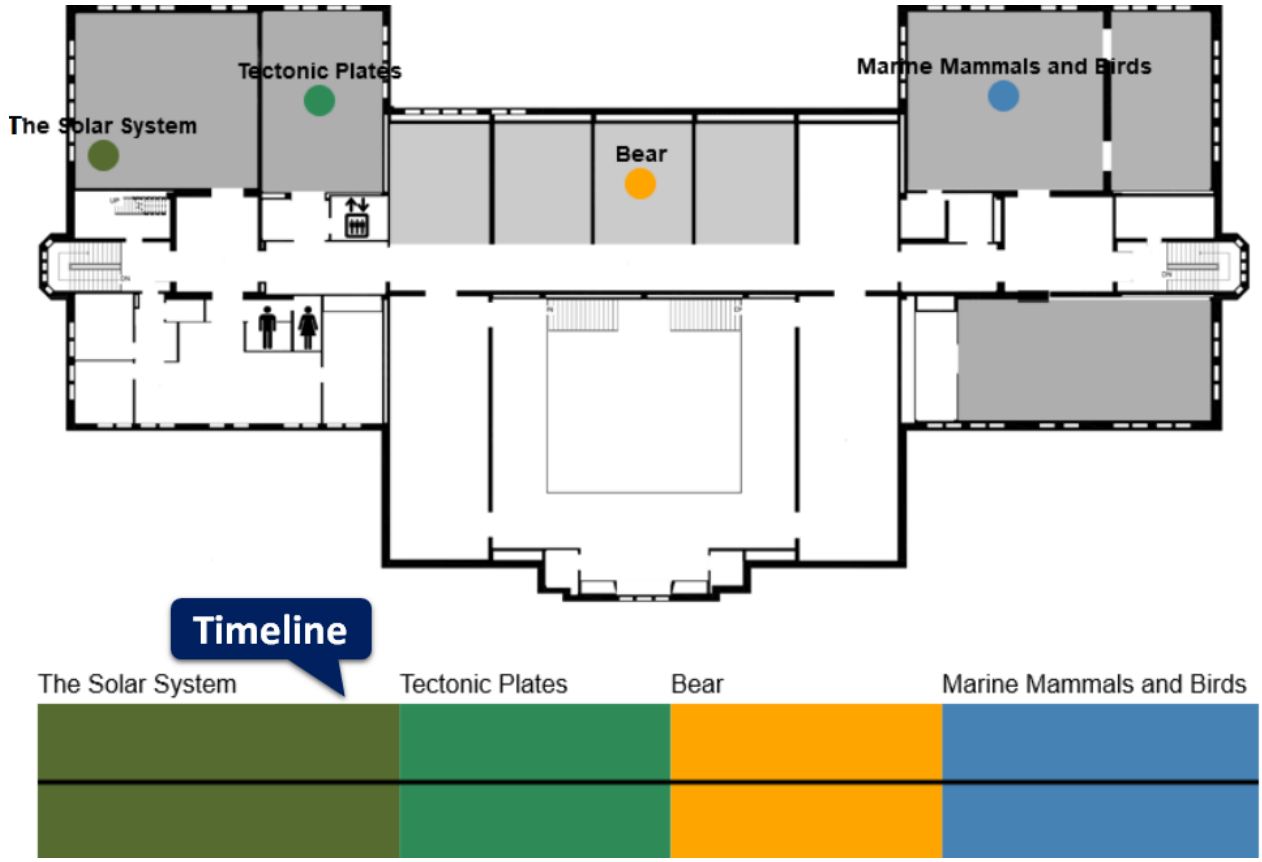


Figure 8.2: The map and the timeline.

walkthrough [40]. It follows Kort et al.'s affective model [79]. This identifies four phases of learning and the affective character of each. The first phase is encouraging exploration with positive affect. The second phase introduces challenges, and negative affect is expected. The third phase is to support overcoming challenges and reduce the negative affect, and the fourth phase is to affirm learning and restore positive affect. In the affective walkthrough, the role of the evaluator is to go through some tasks, while looking at four questions.

8.4.1 Method

To examine the effectiveness of the walkthrough, we recruited participants to use our technique to evaluate the software. We recruited only people with HCI evaluation experience, but not members of our own research group. The study protocol was approved by our research ethics board (see appendix). We applied a qualitative approach, audio recorded and took

notes for our detail analysis to identify software issues.

The participants applied the walkthrough to evaluate the software. We as facilitators first started the process and gave the tasks that were chosen for testing. One person operated the software on a large screen, while the others discussed what to do and what they found. They walked through the software, answering each of the questions and considering the modalities and the teaching objectives.

8.4.2 Findings

With the first participants we ran into issues where they were focused on the *usability* rather than the affective dimensions, and found that the affective walkthrough was too concerned with usability (see Table 8.1).

The importance of a need for better solutions with regards to evaluating software systems beyond just their usability was clear. Therefore, to bring more emotional aspects of the users into focus (and not usability), we emphasized Kort et al.'s affective model [79] more and modified the wording in our technique, and identified new walkthrough steps. Kort et al.'s model highlights the relationships between emotions and learning, and was built on Russell's circumplex model. This model focuses on what emotional state the learner is in, and from that, what help he/she needs, and attempts to link the emotion and the cognitive aspects of the learning process. This gave the modified affective walkthrough (Version 2) in Table 8.2. We used these modifications for later participants. For each walkthrough question we now present some of the participants' comments, highlighting those that we felt reflected on the effectiveness of the walkthrough.

First question: Does the system use positive emotions to encourage the user to explore the learning environment? Participants typically disagreed, and their comments showed the question did influence them to consider the appropriate issues.

> **P:** The user has a fair amount of freedom to explore. But counteracting that is the lack of clear goals that would provide the motivation to explore. Also visualization is boring and like a form, nice to have a drag and drop interface here.

Table 8.1: The Affective Walkthrough (Version 1).

The Affective Walkthrough (Version 1)	
1	Will the learner be able to use the multiple sensory modality correctly (learnability in the user interface) and taking into consideration the learning objective that the instructor has set up and the cognitive affective learning strategies of the instructor that might help the learner to achieve the right effect that the subtask needs (learnability in the education)? E.g. what cognitive affective strategies might help the learner achieve the instructor's cognitive affective objectives? And does the learner understand that a subtask like specific hand gesture with correct distance is needed to reach the learner's goal? What might the learner experience emotionally?
2	Will the multiple sensory modality helpful and does the learner notice that the correct action is available? E.g. is the multimodal system sound clear and visible?
3	Will the learner understand and associate the correct action by using the multiple sensory modality with the effect trying to be achieved? (Learnability of the user interface) E.g. the multimodal system sound is clear and visible but the learner does not understand the text and sound, and will therefore not respond to it.
4	When the correct action is performed, will the learner know that they have done the right thing, receive affective strategies, and see that progress is being made toward solution of the task? E.g. does the learner get haptic feedback, sound or written message when the correct action is performed?

Table 8.2: The Affective Walkthrough (Version 2).

Walk through the system answering each of the questions considering the new modalities and the teaching objectives	
1st	Exploring: Does the system use positive emotions to encourage the user to explore the learning environment?
2nd	Challenging: Does the system provide more difficult material to challenge the user?
3rd	Overcoming: Does the system allow the user to persevere and overcome challenges?
4th	Affirmation: Does the system gives positive affective feedback to affirm successful learning?

P: There is not much information presented about the exhibits. I am wondering why we adding to the itinerary, what it does? Why we are doing it? Not clear why the need of these functionalities.

P: Give a dynamic sort of feel, for example using animations, as circle moving to itinerary, expands up or somethings that feels it is living more.

P: Dramatic slideshow encourages students to seek what they have missed out on and what to look for in the future. (Other participants saw this as a weakness in the software instead of a challenging experience for the learner.)

Second question: Does the system provide more difficult material to challenge the user? The participants again did understand the intent and offered insightful comments.

P: The act of exploring the exhibits is not challenging in itself.

P: The activities are a little bit passive, because all they doing is just clicking on things and they probably are not even reading them.

Third question: Does the system allow the user to persevere and overcome challenges? Their comments were similar:

P: There are no challenges to overcome inherent in freely exploring a virtual museum.

P: It does not prod you to think about these challenges to overcome. It does present data to you, does not seem challenging at all. It does not help you string together a story. Categorical visualization provides a holistic view of the material and then its individual break down by exhibit, and it prompts learning "*learn more here*". Allows for a holistic narrative, by theme instead of exhibit.

Fourth question: Does the system give positive affective feedback to affirm successful learning? The participants explained:

P: I do not remember any instance of positive feedback during the process of exploring the museum or creating the visualization.

P: I expected music in dramatic, the name of dramatic more interactive entertaining.

> **P:** It is good for individual exhibits, but not so much as an overall theme of learning. Provides tools for students to seek out learning, no direct feedback but it encourages learning by having shareable content. Multimedia could also help with this (music, videos, YouTube links).

Overall, the participants understood how to apply the walkthrough. They were able to imagine themselves in the place of the learners and consider which of the four questions related to each situation. Despite not being familiar with the work of Kort et al., the participants quickly understood the steps, their rationale and the connection between affect and learning. Moreover, they were able to connect the questions with the narrative quasi-sensory modality. We discuss other issues in more detail later in Section 8.6.

8.5 Museum Explorer and Affective Heuristic Evaluation

For evaluation of affect in software our second proposed technique, the *Affective Heuristic Evaluation*, takes a standard method of evaluating usability (heuristic evaluation), and adapts it for this particular issue, which we explained in Chapter 6 (see Table 8.3).

Table 8.3: The Affective Heuristic Evaluation (Version 1) (summarized from Chapter 6).

Affective Design	**Heuristics**
Visceral	*H8:* **Aesthetic and minimalist design**
	H10: **Help and documentation**
Behavioral	*H1:* **Visibility of system status**
	H3: **Learner control and freedom**
	H5: **Error prevention in learning**
	H6: **Recognition rather than recall**
	H7: **Flexibility and efficiency of use**
	H9: **Help users recognize, diagnose, and recover from errors**
Reflective	*H2:* **Match between multimodal software and the educational environment**

Table 8.3: The Affective Heuristic Evaluation (Version 1) (summarized from Chapter 6).

Affective Design	**Heuristics**
	H4: **Consistency and standards: Consistent and intuitive mapping**

8.5.1 Method

The affective heuristic evaluation is another of what Nielsen and Molich [104] term an inspection method, meaning the system is inspected by evaluators, rather than involving actual users. The method is therefore not as definitive, but the process is easier, cheaper, and shown to usually produce similar results.

We follow the method of Nielsen and Molich, but with our set of heuristics. In pre-testing we found that the wording of our heuristics were confusing, so we improved the wording, giving Version 2 of the heuristics, shown in Table 8.4. The new wording uses verbs to show the goals of the heuristics, making them more understandable, and re-numbered them to make a clearer order. We also changed the wording of some heuristics to better emphasize our particular focus (e.g. Version 1 H1 vs. Version 2 H3).

The proposed affective heuristics themselves are adapted, with affect supported by the work of Norman [105] and Kort et al. [79] and multimodal design based on the work of Sankey [132]. The rational, and more detail, has been provided earlier in Chapter 6.

Table 8.4: The modified Affective Heuristic Evaluation (Version 2).

Affective Design	**Heuristics**
Visceral	*H1:* **Only include design elements and modalities that support the learning objective** *H2:* **Provide help and documentation only where it is likely to be needed**
Behavioral	*H3:* **Maintain visibility of progress achievements and challenges** *H4:* **Allow the user freedom to explore but also to return to the previous step** *H5:* **Avoid or prevent actions with no learning value** *H6:* **Visualize options clearly to support exploration** *H7:* **Tailor frequent actions considering different skill-levels of learners** *H8:* **Challenge learners and provide constructive feedback if they fail**
Reflective	*H9:* **Match the learners world view in language and multimodal support** *H10:* **Maintain interface consistency to support task continuity**

We took a qualitative approach, and both audio recorded and took notes for later detail analysis to identify issues. The sessions involved participants acting as evaluators and two facilitators representing the instructor and the software developer.

The evaluators applied the proposed inspection method to evaluate the Museum Explorer software. We as the facilitators first gave an introduction to the software. The evaluators individually went through the affective heuristic evaluation (Table 8.4) and found the interface elements violated, the design problems, gave a solution and rated the severity, following Nielsen's process.

We monitored the evaluators and collected data to see how the evaluation methods worked, trying to improve the inspection technique with every evaluation session. This was an iterative process; we were recoding, going back and discussing what they have learned from the observations; we modified the proposed heuristic evaluation with evaluators' feedback and reshaped the proposed heuristics during the study.

As we mentioned earlier, in the MADE inspection technique we are considering multimodal and affective aspects. For instance, an evaluator may think that when the software did nothing, it should have done something extra, and there should be some sort of multimodal support more than what is there. The multiple sensory modalities we are considering are such as music, sound or movies, and the quasi-sensory modalities are narrative or storytelling (does it help a story?), or persuasion (was it persuading about the right or wrong thing?). Below we discuss our findings using the affective heuristics evaluation.

8.5.2 Findings

In the study participants read the heuristics, explored, and then reflected on the software. They then wrote comments about the system filling in a form (see Figure 8.3) with respect to the different heuristics. For heuristic severity the participants used Nielsen's severity ratings for usability problems [102], changing their emphasis to reflect learning impact. For each

MADE Heuristic(s) Involved	Interface Element	Usability Problem	Solution	Severity (0 to 4)
.	.	.	.	.
.	.	.	.	.
.	.	.	.	.

Figure 8.3: The Affective Heuristic Evaluation Form. The results are shown in Tables 8.5 and 8.6

heuristic, we show below participant comments that reflect on the effectiveness of the heuristics (see Tables 8.5 and 8.6). Overall, participants were able to understand the heuristics and apply them sensibly. Each of the ten heuristics led to a useful comment at least once. The heuristics invited more reflection on modality as well as affect. However, some participants

commented more than expected about simple usability concerns; see H2 and H6 as noted below. Of course, poor usability will inevitably have an effect on learning.

8.6 Discussion

This chapter described an inspection method used to evaluate a Museum Explorer system. Instead of evaluating the usability we were trying to evaluate the effectiveness of the affective modality in this educational software. The system helps the students to make a visual story of their trip to the museum, so the modality that we were interested in was the narrative aspect.

Our primary purpose in this case study was not to evaluate the Museum Explorer system, but rather to test and refine our MADE inspection techniques. Therefore, we are interested in whether our method revealed issues of concern, and not those issues that are unimportant or irrelevant. Of course, as an inspection method, the limitations are reliance on general principles, rather than on detecting specifics issues using testing.

Affective Walkthrough: In the first part, the participants using our inspection method, found a number of issues in the software that might lead to significant improvements. For example: they found the software did encourage at the beginning and provide a sense of completion at the end, but had no engaging challenges in between. They also identified many places where multimodality (animation, music, video) would have been beneficial, but was not provided. The participants found the software seemed surprisingly passive and even boring (just adding pictures and comments, but no humor, no music, no video), and did not bring the engagement expected from narrative. They also found inconsistencies and lack of continuity in the interface, which also detracted from engagement. Moreover, the participants also thought the design would not be motivating for users. For example, when some participants understood the visual story idea, they responded by commenting: "*it should be exciting like making a movie.*"

However, as our sessions progressed, we discovered several important ways to improve our inspection methods. In a normal cognitive walkthrough there are four questions that apply

Table 8.5: Affective Heuristics result, participant comment, and severity (Part 1).

Affective Heuristic Evaluation		
Heuristics	Participant comment	Severity
H1: Only include design elements and modalities that support the learning objective	P: The interaction flow does not make creating the visualization the focus for the learner, e.g. the activity should be something like "let's make a slideshow"	3
H2: Provide help and documentation only where it is likely to be needed	P: The objective of adding exhibits and then visiting them is unclear, as a user I would simply click on the circles and try to explore directly. This is for the main page. Visualizations are clear in this regard [Note the usability emphasis.]	1
H3: Maintain visibility of progress achievements and challenges	P: Challenges and achievements aren't highlighted in all visualizations, some of them are such as the dramatic visualization, while others just display what has been done. It tells you what you found, and it tells you are missing something which you have to go back and explore again	2
H4: Allow the user freedom to explore but also to return to the previous step	P: Should allow the user to generate a more personal story of their experience at the museum including photos they have taken, comments, etc.	2
H5: Avoid or prevent actions with no learning value	P: The map metaphor doesn't help the students engage with the topics	3
H6: Visualize options clearly to support exploration (somethings this was misinterpreted as leading to usability and not educational aspects)	P: All of the buttons on the page are the same size. Similar shape and size visually groups these buttons together, but they have very different functions [Note the usability emphasis.] P: The information sidebar is not well-integrated into the rest of the interface. It updates when you click on the timeline or the map [Note the usability emphasis.]	3 1

Table 8.6: Affective Heuristics result, participant comment, and severity (Part 2).

Affective Heuristic Evaluation		
H7: Tailor frequent actions considering different skill-levels of learners	P: Does not specify tasks for different learning levels i.e. all task provided are meant for one level and advanced users can be limited by this	3
H8: Challenge learners and provide constructive feedback if they fail	P: There is no [challenge in the] task itinerary, it's just adding exhibits to a list	3
H9: Match the learners world view in language and multimodal support	P: No dynamic visual representation of dynamic objects like animals, big bang P: There is an emphasis on images, and a lack of use of multimedia such as audio, movies, videos, etc., which is less like the real world where multimedia is employed to help students learn. This applies to the dramatic, categorical and possibly sequential	4 2
H10: Maintain interface consistency to support task continuity.	P: The timeline isn't available in each exhibit which affects the visit exhibit element. Having the timeline would allow for navigation between previously visited exhibits and allows for more freedom to explore. The timeline has been employed in other places such as the visualizations so it is not consistent with the remainder of the user interface [Comment also related to H4.]	2

Table 8.7: The Affective Walkthrough (Version 3).

Walk through the system answering each of the questions considering the new modalities and the teaching objectives	
1st	What is the learning goal of this task?
2nd	Where in affective cycle of learning is this task? (i.e. exploring, challenging, overcoming, and affirmation)
3rd	Is the appropriate affective support provided?
4th	Does the affective support work as intended?

to each task. In our modified affective walkthrough (Version 2), the questions were more holistic. That is to say that in the beginning of each session the evaluators should have a general idea of the system and the tasks, as the walkthrough is following four steps of learning: exploring, challenging, overcoming, and affirmation. For each task only one of those steps might be appropriate. This perhaps caused the walkthrough to be more difficult.

Once our participants understood this, they could make helpful comments. For example, where a task clearly related to exploring, one participant said: "*Yes, I think the software encourages the user to learn by providing positive emotions, it says the task I want is available to me and I can do something with it...*". Moreover, where participants recognized that the task should be more challenging, to ensure real learning was taking place, they identified that the Museum Explorer did not really have any challenges: "*There are no achievements or challenges to complete... I thought it was just filling in what I visited (or what I plan to visit), which would be the closest thing to a challenge.*" However, because there were no challenges, there was nothing to overcome, and therefore, it was hard to evaluate persuasion and encouragement. This can require delicate design, because different learners will find different levels of challenge possible. We felt that it would be better if all the questions applied at every step, as happens with a cognitive walkthrough.

Therefore, we propose the final version (Version 3) shown in Table 8.7 that can confidently be used to evaluate an educational system.

Affective Heuristic Evaluation: For the second part, the participants using our inspection method, found a number of issues in the system that might lead to significant improvements. For example: *"the map metaphor doesn't help the students engage with the topics"*. They also identified many places where multimodality (animation, music, video) would have been beneficial, but was not provided. The participants also found the system surprising for having no dynamic visual representation of dynamic objects like animals; being passive and even boring (just adding picture and comments, but no humor, no music, no video), and did not bring the engagement expected from narrative. They also found inconsistencies (e.g. the timeline) and lack of continuity in the interface, which also detracted from engagement. Overall, the participants also thought the design would not be motivating for users. They commented that creating a visual narrative should be exciting like generating a more personal story of their experience at the museum.

Overall, we were pleased and feel the points identified above would lead to significant improvements in the Museum Explorer. However, as our sessions progressed, we discovered several important ways to improve our inspection methods. During the first observation sessions we noticed that participants interacting with the system during the evaluation focused a lot on usability issues. This resulted in a substantial loss of emphases on modality and learning. For instance, one participant commented: "The buttons for add exhibit to itinerary and visit exhibit do not look buttons, but create visualization does look like a button, which might confuse the user."

In typical usability inspection methods the goal is clear: usability. In our MADE method, however, the issues are complex. In particular, we assume that there is a specific educational objective and a modality with some known advantages is being used to help support that objective. We found that in order for participants to use our methods, we needed to state the objectives and the modality advantages more clearly and explicitly. For example, in the case of the Museum Explorer system, the objective was to make learners more engaged and reflective (think about what they learned) and the strategy was narrative, to help the learners tell a visual story. We found that the more explicitly we stated these objectives, the better our participants would understand and usefully apply the inspection methods. Therefore, we modified the inspection method, and came up with the heuristics described in Table 8.8. The

general structure is similar, but it has a stronger focus on the particular learning strategies and modalities involved. It also emphasizes Kort et al.'s affective model [79], depicted in Figure 8.4. This identifies four phases of learning and the affective character of each. The first phase is encouraging exploration with positive affect. The second phase introduces challenges, and negative affect is expected. The third phase is to support overcoming challenges and reduce the negative affect, and the fourth phase is to affirm learning and restore positive affect.

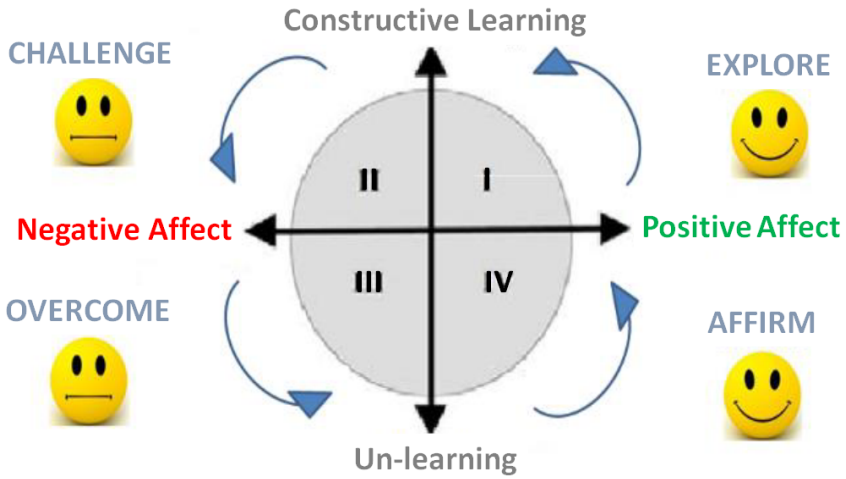


Figure 8.4: The affective model of education from Kort et al.

We improved the affective heuristics and adopted it in later evaluation sessions. Table 8.8 shows our final version of the affective heuristics.

Table 8.8: The final version of the Affective Heuristic Evaluation (Version 3).

Affective Design	**Heuristics**
Visceral	*H1:* **Design elements and modalities should support the affective learning strategy** *H2:* **Ensure help and documentation is provided where needed but does not distract from affective learning strategy**
Behavioral	*H3:* **Maintain visibility of progress, affirming challenges already overcome, and those remaining**

Table 8.8: The final version of the Affective Heuristic Evaluation (Version 3).

Affective Design	**Heuristics**
	H4: **Allow the user freedom to explore but also to return to the previous step**
	H5: **Avoid or prevent actions with neither feedback to help overcome, nor affirmation when success**
	H6: **Visualize options clearly to encourage exploration**
	H7: **Tailor actions to be encouraging at first and efficient later, while learners are attempting to overcome challenges**
	H8: **Challenge learners and provide constructive feedback if they fail, and affirming success when they succeed**
Reflective	*H9:* **Match the learners world view in affective strategy and multimodal support**
	H10: **Maintain interface cohesion to support affective strategy**

8.7 Summary

This chapter focused on two evaluation techniques for affect in educational technology development. We identified a need to de-emphasize ordinary usability, and state explicitly the teaching objectives, the educational strategies and modality advantages involved.

It focused on two inspection evaluation techniques for evaluating affect in educational software. We presented a case study where participants used the technique to evaluate a Museum Explorer system. The system featured narrative and visual stories to support greater engagement and reflection. Our inspection techniques were a cognitive walkthrough and a heuristic evaluation based on techniques from usability evaluation but adapted for multimodal affective systems.

In our case study participants were able to apply the inspection techniques and make useful

comments that would significantly improve the system.

However, by using qualitative analysis of our observations and transcripts of participant think-aloud comments, we were able to identify several ways to improve our inspection techniques. In particular, we identified a need to de-emphasize ordinary usability, state explicitly the educational strategies and modality advantages involved, and align with the model of Kort et al. [79] on the emotional trajectory of learning.

8.8 Addendum: Carleton Virtual

Our next case study of our evaluation techniques was intended to be of Carleton Virtual [4]. This was a 3D virtual immersive environment, similar to Second Life, customized for use as a teaching platform in our university. We obtained ethics approval for our study, and conducted interviews with two instructors who used this system in teaching classes. Just as we were about to begin the main part of our study, the system was closed down by the provider, Avaya. Although we were not able to complete the study, we did gain some useful insights.

The main way the system was used in teaching was to simulate authentic real-world situations for the students to role-play within. The multimodal strategy was therefore, the 3D immersive world. The quasi-sensory modalities involved the students role-playing narrative and persuasion. For example, in language learning, the environment was used to support scenarios for shopping, presenting, and discussion of objects. The cognitive strategy appeared to be situated cognition, as we discussed in Chapter 3, Section 3.3.2.

The affective strategy, however, was less clear. We heard anecdotal stories that suggested students might be less shy, and might find it easier to role-play other identities. Students apparently enjoyed the system and the simulated environments, such as a market place (with fruits and vegetables) and a real-estate agent's open-house. From an affective point of view, one of the limitations of the system was the avatar's inability to express emotion naturally. For instance, facial gestures were not possible, so there was no way that avatars could show happy or sad faces. Body language was also limited. As a result, misunderstanding and

miscommunication could follow. On the other hand, we heard that difficulties with the system led students to help each other in learning how to work within the constraints of the software.

We regret that we were not able to complete this study by applying our evaluation methods. We speculate that one issue that would have arisen would be the importance of the evaluators having the learning objectives and the learning strategies made explicit. This was the issue we identified in the Museum Explorer case study and addressed in Version 3 of our techniques. We are unsure of how well our techniques would identified the affective issues discussed above. The challenge is the generality of a system such as Carleton Virtual, compared to systems with a more narrow focus.

Instead of the Carleton Virtual case study, we conducted a study of Minecraft Hour of Code, as presented in the next chapter.

Figure 8.5: Carleton Virtual: avatars in a virtual learning space. `https://carleton.ca/learningspaces/virtual-learning-spaces/`

Chapter 9

Case Study 3 – Minecraft Hour of Code: Evaluating Software for Learning Computer Programming

9.1 Introduction

This chapter presents an evaluation of Code.org's Minecraft Hour of Code. The software was created to encourage and support people for initial learning of computer programming. In particular, this web-based software uses a spatial model world, where the learner's programs manipulates the world. We applied the *Affective Walkthrough* and the *Affective Heuristic Evaluation* proposed earlier in Chapter 6 and refined with the previous case study in Chapter 8. Our previous studies were conducted with participants to refine our methods, and inform us about the collaborative processes. In this new study, we simply applied this knowledge and conducted the evaluations ourselves. Our findings provided illumination about the Minecraft Hour of Code approach, highlighting some aspects that are successful, and others where improvement appears necessary. We also gained insight about the evaluation methods and their effectiveness.

Code.org is a non-profit organization that encourages learning of computer science, especially programming and computation thinking. It's "Hour of Code" is an initiative that emphasizes short introductory tutorials aimed at total beginners, and has been widely used, claiming almost 100 million learners, including US President Barack Obama [152].

The Hour of Code approach is strongly visual in nature, both for the program itself, and for what the program does. This is reasonable, as it is established that such visual environments might assist in novice engagement, and also support them to demonstrate skills and strategies in a familiar or easily understood context [18].

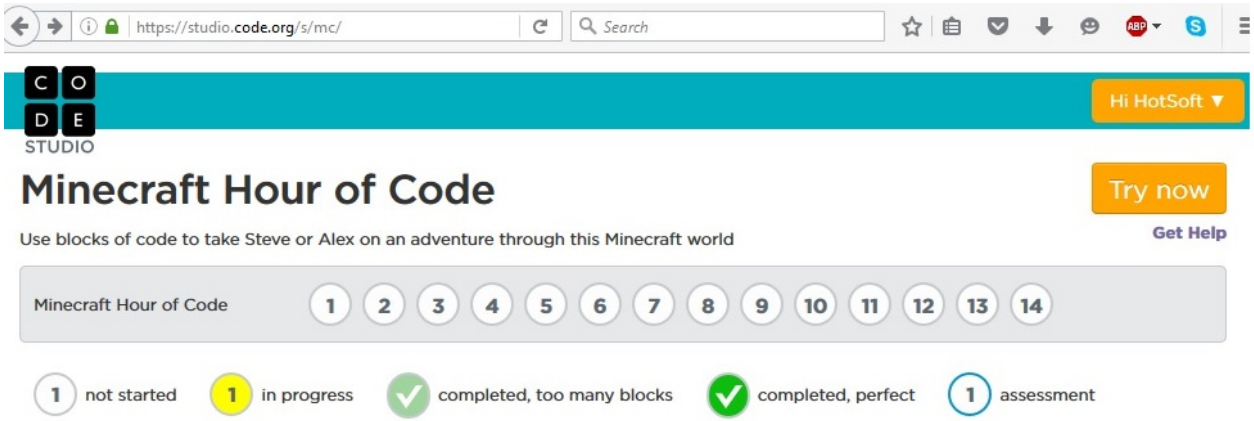


Figure 9.1: The initial interface of the Minecraft Hour of Code tutorials.

One aspect of Hour of Code's visual nature relates to the program itself, and it uses a variant of the "jigsaw" or "block" approach used in Scratch [121], Alice [34] and other learning systems. The other aspect is the spatial model world, where such worlds include graphical elements that serve as integral parts of a computational context, problem, and solution. The Hour of Code tutorials include many variants of this, for example including ones resembling many popular games, e.g. "Angry Birds", " Plants vs. Zombies". In this chapter we focus on the "Minecraft" variant — see Figures 9.1 and 9.2. This is of special interest because Minecraft itself involves programming as a core part of the appeal of the "game".

Our interest is in *affective* learning in multimodal educational software, focusing on the emotional elements in the software. This reflects calls for more attention to emotion in the domain of Human-Computer Interaction (HCI) and education [154]. In this chapter we present our application of two such methods, the affective walkthrough and the affective heuristic evaluation proposed earlier to the Minecraft Hour of Code software.

9.2 Overview of Minecraft Hour of Code

The Hour of Code Minecraft tutorial was developed to engage students in learning programming. It introduces players to basic programming concepts, allowing them to navigate, mine, craft, manipulate and explore in a 2D Minecraft world by plugging together blocks to complete all actions and generate computer code.

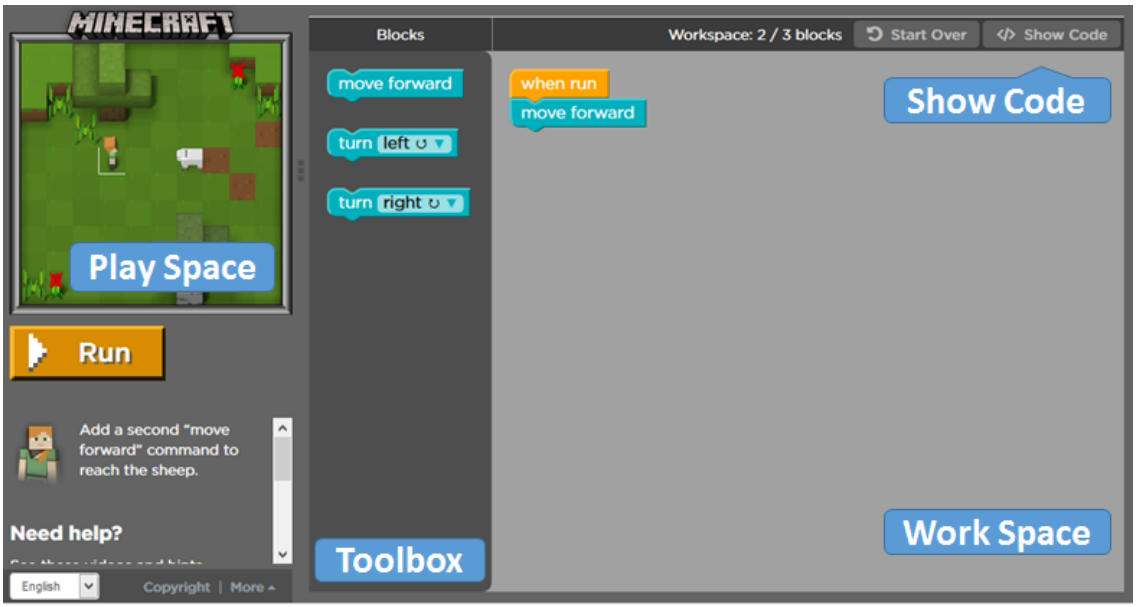


Figure 9.2: The Hour of Code Minecraft interface, with model world on the left, and the program on the right.

The system provides a sequence of tutorials, illustrated by a horizontal line with nodes representing each tutorial step, as shown in Figure 9.1.

For each tutorial, the learner uses an interface such as shown in Figure 9.2. On the left top is the model world ("play space"), resembling the rectilinear blocks as used in the full Minecraft game, representing objects such as grass, trees, sheep, etc. On the left bottom is a panel with instructions to the learner. When a learner begins the tutorials, they select an avatar, either "Alex" or "Steve". The chosen avatar then appears in the model world, and the program controls their actions within the world.

On the right is an area consisting of a "toolbox" of jigsaw-like pieces, and a "workspace". The pieces in the toolbox are for simple statements or commands, as well as loop ("repeat") and conditional ("if") statements, and the shape of the pieces shows how they connect with other pieces. These can be copied and dragged into the workspace, and assembled in sequence to form programs.

When assembled, the jigsaw pieces do resemble traditional (textual) code, and in most Hour of Code this is the only code shown. In the Minecraft Hour of Code tutorial, however,

whenever a tutorial step is completed satisfactorily, a pop-up window shows the equivalent code in JavaScript, which is what is actually used to program Minecraft.

In Minecraft Hour of Code, the tutorial steps reflect teaching and learning goals, in sequence. The sequence is fairly traditional, beginning with sequences of commands, moving on to loops, and then conditional structures. At first, the introduction is strongly guided, and then more freedom allowed to encourage a program solving approach.

In the interface and the tutorial steps, several modalities are involved. Firstly, the interface is graphical and spatial, with the model world and the jigsaw piece program. The character of the graphics is appealing, with a coarse lo-fi and fun game-like appearance that might help engagement. Secondly, the model world suggests a "world of action", like a board game, where the blocks might be moved and interact. Thirdly, and related, is that the tutorial instructions build narrative elements where the player's avatar should accomplish goals by moving and acting in the model world. Over the sequence of the steps, these build into an overall story. This approach would appear to help learner motivation, understanding, and possibly supporting reflection and creativity about what other possibilities might be programmed. Finally, there are audio elements with calm background music and occasional effects to emphasize events.

9.3 Minecraft Hour of Code and Affective Walkthrough

The affective walkthrough proposed earlier in Chapter 6 was based on Wharton et al.'s cognitive walkthrough [150], and Dormann and Biddle's affective walkthrough [40]. Following case studies, it was refined in Chapter 8 to de-emphasize normal usability issues, and to acknowledge the contextual role of teaching goals and potential modality benefits. It follows Kort et al.'s affective model [79]. This identifies four phases of learning and the affective character of each. The first phase is encouraging exploration with positive affect. The second phase introduces challenges, and negative affect is expected. The third phase is to support overcoming challenges and reduce the negative affect, and the fourth phase is to affirm learning and restore positive affect.

9.3.1 Method

As with the earlier cognitive walkthrough by Wharton et al., the main idea is to select user tasks and then step through the software considering key questions, and making notes and observations, as shown in Table 9.1. The walkthrough was conducted by us (GhasemAghaei with Co-supervisor Biddle), role-playing evaluators. We used a large screen display to collaborate in the same manner as done by participants described in Chapter 8.

Table 9.1: The Affective Walkthrough (Version 3).

The Affective Walkthrough (Version 3)	
1	What is the learning goal of this task?
2	Where in the affective cycle of learning is this task? (i.e. exploring, challenging, overcoming, and affirmation)
3	Is the appropriate affective support provided?
4	Does the affective support work as intended?

The objective of our study was to evaluate affective learning in Hour of Code Minecraft tutorials, with special attention to the supporting modalities involved. We presented, above, the basic design of the tutorials and the interfaces, and identified the modalities involved and their potential benefits. We then outlined the basis of the affective walkthrough, and the steps involved.

To select the tasks, we simply used the Hour of Code Minecraft tutorial steps, although we grouped the 14 steps into 4 larger tasks, bringing together steps that had strongly related educational purposes.

We then followed the walkthrough as shown in Table 9.1, and considered each question in turn. We made Walkthrough Evaluation Sheets to record our answers and related observations. In the section below, we review the nature of the tasks, and the results of our walkthroughs.

9.3.2 Findings

First Task: To begin, we select a character. We have two choices, e.g. we can pick either Alex or Steve. The first task is about "commands". The steps are shown in Table 9.2 and the results of the walkthrough are shown is Table 9.3.

Table 9.2: First task: walk around.

First Task: Command Sequences	
1	Add a second "move forward" command to reach the sheep.
2	Then walk to the tree and use the "destroy block" command to chop it down
3	Use the "shear" command to gather wool from both sheep.
4	Cut down all trees.

By completing the first part of Task 1, the software gave emotional feedback such as happy sounds, a jumping character, green highlight on the achievement bar and a pop-up window providing feedback. Figure 9.3 shows the pop-up window with the achievement bar and our character reaction at the bottom. If we were not able to complete the step with the minimum lines of code, the software will let us know with a challenging message, but it does not help us to complete the step with the minimum lines of code (see Figure 9.4). But if it was really going to encourage exploration, it should allow the user to see all of the options he/she could select and do, but it does not.

For every task, the software only provides us with the number of blocks we need. Therefore, they are not encouraging exploration all that much. They are not providing us with other alternatives. That is a bit of a criticism if we are encouraging people to explore. If we want to encourage exploration, we do not constrain people. We give them a bunch of different things, so that they explore to see what happens. We are surprised they did not give all the blocks that are relevant to let us explore a bit (e.g. *shear* sheep and *cut down* tree). We were a bit confused that we did not know all the possible things or at least some small subset of the possible things.

In the second part of the first task, the software took away *cut down*, and added *shear*.

It provides the user with more information about the reason of doing the task, e.g."wood is

Figure 9.3: Task 1 - Step one.

a very important resource. Many things are made from it." The third part of the first task uses "shear" command instead of "destroy block".

As we are not using repeat loops in this task, the last part of task one needed lots of "moveForward()" blocks. It would be easier if we could just tell the computer to perform the move forward command the number of times we needed. It would be much easier to transition to repeat loops, if we had told the computer to "move forward" a specific number of times. We will need thousands and thousands of blocks to create a new world in Minecraft.

Second Task: The second task is about building the rest of a house from the material available. The "repeat" command will come in handy. For the last part of the task we can select "Easy", "Medium" or "Hard". The steps are shown in Table 9.5 and the results of the walkthrough are shown is Table 9.4.

The second and third tasks are about how to come up with algorithms, giving a problem and solve it using repeat loops. For the first part of task two, the software provides the reason and time to complete this part by saying: "We need to build a house before the sun goes down. Houses require a lot of wood." For the second part of this task it suggests: "Every

Table 9.3: First Task Walkthrough Results: walk around.

Walkthrough— First task: Command Sequences	
1	What is the learning goal of this task? About learning programming and being engaged. In this task we used commands and add them in the work space to help the character move around the screen. we were able to see the code that we created with using blocks. It can be done, but there were no motivation to it, and considering the role of narrative, which could be better. *Summary: The learning goal of this task was understood to learn how to add new commands in the program, but no motivation.*
2	Where in the affective cycle of learning is this task? (i.e. exploring, challenging, overcoming, and affirmation) It cover exploring and gives feedback through sound and text if I did wrong and otherwise an affirmation sound with jumping character and the character being happy. *Summary: The affective cycle of learning was exploring and affirmation.*
3	Is the appropriate affective support provided? Yes, by adding blocks from toolbox in the work space, we can see how our character moves in the play space, and receiving positive feedback or negative that is still encouraging by providing how to try again with character's emotional reaction. *Summary: appropriate affective support is provided with different multimodal support.*
4	Does the affective support work as intended? No, affective support did not work as intended. The emotional support was well provided, but the task did not provide reasons of using the commands. *Summary: even with the affective support provided, the purpose of the task was not clear.*

Table 9.4: Second task walkthrough results: build a house.

Second Task: Repeat Loop and Commands — Simple Task	
1	What is the learning goal of this task? The learning goal of this task was to help you learning about repeat loops and commands in programming. *Summary: learning about repeat loops and commands.*
2	Where in the affective cycle of learning is this task? (i.e. exploring, challenging, overcoming, and affirmation) The affective cycle of learning in this task covers all the four cycles. Minecraft Hour of Code lets us explore, and provides with three different easy to medium, and hard questions covering the challenging cycle. Then it helps us to overcome with popup windows. And at the end it gives us affirmation. *Summary: it has exploring, challenging, overcoming, and affirmation.*
3	Is the appropriate affective support provided? Yes, it starts with a question and a hint, i.e. "We need to build a house before the sun goes down. House requires a lot of wood. Cut down all the 3 trees." *Summary: affective support was provided*
4	Does the affective support work as intended? No, it does not work as intended. It really does not help us to build a house, just create some walls. Other issue with the software it tells us the task should be done before the sun goes down, but it does not work as intended, and we do not have a time limit or a change in the background. *Summary: the purpose of this task is to build a house, which it did not help us to do so.*

Table 9.5: Second task: build a house.

Second Task: Repeat Loop and Commands — Simple Task	
1	Build the first part of your house by putting the "place" and "move forward" commands inside the repeat loop.
2	Let's build a house. Choose the floor plan (easy, medium or hard) for your house.

Figure 9.4: Task completion but not with minimum lines of code.

house starts with a wall."

This part of the software is not telling the user how much wood he/she gets, and why we need to chop down all the trees. We do not have variables. We do not have any idea how much wood we have. Also we can not really build a house; we can just build some walls, so the story is not correct.

Third Task: This task is again about using "Repeat loops" and "commands", but in a more challenging way using some creativity and freedom. It has four steps and each step takes you to different interactive spaces including: 1. Plant crops on both sides of water, 2. Move past the Creepers and reach safety at home, 3. Move underground placing torches and mine coal, 4. Avoid walking into molten and place cobblestone to create a bridge, and mine iron blocks. The steps are shown in Table 9.6 and the results of the walkthrough are shown is Table 9.7. This task has gamification steps to bring more engagement to the user.

Forth Task: Finally, the last task is about "commands", "repeat loops", and learning about "if" statements; a fundamental part of learning to program. It introduces us to a concept, which requires more problem solving. "if" statements help all computers make decisions. We

Table 9.6: Third task: plant some crops.

Third Task: Repeat Loop and Commands — Challenging Task Commands	
1	Plant crops on both sides of the water so you don't get hungry later on. (It's good to plan ahead.)
2	Running into a Creeper is a bad idea. Carefully move past the Creepers and reach the safety of your home.
3	You'll find the most valuable resources underground, but it can get dark. Place at least 2 torches and mine at least 2 coal.
4	Walking into molten lava is a bad idea. Place cobblestone to create a bridge, then mine at least two of the iron blocks.

Table 9.7: Third Task Walkthrough Results: plant some crops.

Third Task: Repeat Loop and Commands — Challenging Task Commands	
1	What is the learning goal of this task? The learning goal of this task was to practice using different commands with repeat loops bringing challenging steps. *Summary: providing challenging steps.*
2	Where in the affective cycle of learning is this task? (i.e. exploring, challenging, overcoming, and affirmation) This task only covers affirmation and somehow challenging cycles, the steps are not encouraging enough and there is no clear help to overcome the challenges. *Summary: it covers challenging and affirmation cycles.*
3	Is the appropriate affective support provided? Yes, there are similar affective support provided that are: giving a short story about the task and then emotionally trying to engage us in the space *Summary: it was provided.*
4	Does the affective support work as intended? There was a story at the beginning of each step, but narrative is not considered well to make it an exciting experience. *Summary: even with bringing challenge in this task it was not motivating and good story is missing.*

Table 9.8: Fourth Task: "if" statement.

Fourth Task: "if" statement	
1	Lava is hiding beneath some of these blocks, which you'll need to cover up before moving forward. An "if" command will come in handy here. Add a "move forward" command in the correct place to mine these blocks.
2	Now things are getting tricky. Mine 3 redstone, but don't fall in the lava. Use an "If" command to place cobblestone over any lava you uncover.

are able to use "if" statement in the code to make our character react to what they see in the world, e.g. if there is rock in front of her, she can turn left or right and watch her steps. It has two steps; first step is simple and second step is more challenging, which are including: 1. Lava is hiding beneath some of these blocks, which you'll need to cover up before moving forward, 2. Mine redstone but don't fall in the lava by placing cobblestone over any lava you uncover. The steps are shown in Table 9.8 and the results of the walkthrough are shown is Table 9.9. This task has gamification steps to bring more engagement to the user.

9.3.3 Discussion

Reviewing the results of our walkthroughs, we are able to make several general observations. Most simply, we were pleased with several strengths: good visual effects and artwork, and sensible audio. Beyond those, however, we found more insight, described in sections below.

World and Story: Perhaps our main positive finding was the interplay between the model world of the "playspace" and the narrative aspects of the programming tasks given in the instructions. We felt it was clear and motivating when the learner was asked to create programs to move the avatar, chop down the tree, build a wall, and so on. The world and the story seemed to go together well, almost the way that narratives often work in computer games. This worked very well here, and showed some limitations in tutorials in Scratch, for example, which typically start with a blank canvas, rather than a world ready for a story.

While acclaiming this aspect, we found two weakness. One is that there was often little

Table 9.9: Fourth Task Walkthrough Results: "if" statement.

Walkthrough— Fourth task: "if" statement	
1	What is the learning goal of this task? The learning goal of this task was to use command, repeat loop, and "if" statement. *Summary: learning about "if" statement.*
2	Where in the affective cycle of learning is this task? (i.e. exploring, challenging, overcoming, and affirmation) This task covers the challenging (providing two steps: simple and challenging) and affirmation, but does not encourage in exploration or help to overcome. *Summary: covers challenging and affirmation, but exploring and overcoming is not very well done.*
3	Is the appropriate affective support provided? Yes, we saw different affective support during the task. *Summary: there is appropriate affective support provided.*
4	Does the affective support work as intended? no, the affective support did not provide a clear narrative to engage us in doing this task. *Summary: no, the affective support did not provide a clear narrative*

motivation for the actions requested in the instructions. The learner is asked to chop down trees, for example, before any mention of building walls for a house. Second is that there was no overall story, no eventual goal to accomplish, despite this being so common in computer games. Even in Mario, the player knows they are not done until the princess is saved.

Challenge but not Overcome: We very much appreciate the way in which the tutorials helped the learner get started, offering a clear narrative goal, exactly the right tools to accomplish it, and providing affirming feedback when done. In several places, however, it seemed that this limited exploration: it encouraged simple constrained programming, but did not add the context of a large number of possibilities in the "toolbox". Learning to choose a strategy in the presence of many possibilities, with different strengths, is essential to learning about problem-solving.

We also liked the ideas of greater challenge by suggesting it might be done with fewer steps: we felt it really would make the learner try to do better. But whereas a real teacher might

monitor and provide hints when the learner gets frustrated, the tutorial offered nothing equivalent. Considering Kort et al.'s cycle, there was support to challenge, but not to overcome. As we show in Figure 9.5, we see the support for the cycle strong in some places (Challenge, Affirm) but weak in others (Explore, Overcome).

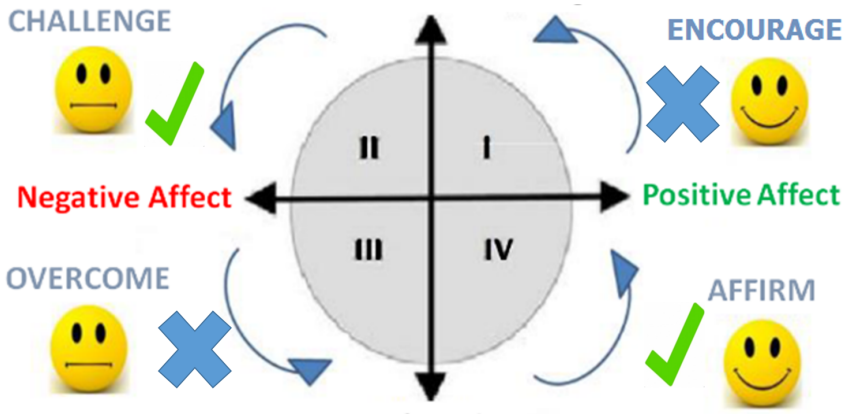


Figure 9.5: Minecraft Hour of Code and Kort et al.'s steps [79].

Simple Programming: As the name "Hour of Code" suggests, the tutorial only addresses the simple beginnings of programming — "First Hour of Code" might be more appropriate. We noticed no real introduction to variables, for example, despite the potential of having them associated with objects in the model world. There was also no lead-in to object orientation, the most common practical paradigm for programming, again despite the potential of objects in a model world — this is exploited well in Alice, for example. In reviewing comments on official videos about the Hour of Code tutorial, this was a common criticism, that it wasn't introducing "real" programming:

> I'm not convinced that this kind of a tutorial could help anybody in understanding of programming paradigms. Modern programming is mostly object-oriented, and this film could at most give a little of vision, how sequential programming should look alike. IMO it's not proper to teach only sequential programming, without any code and without even telling, that another styles of programming exists. Moreover, I think that beginning learning of programming with sequential programming instead of object-oriented programming could make OOP harder to

understand *Youtube Comment*

However it is still an open question as to whether OOP or sequential programming is best for absolute beginners. Similarly, other programming paradigms are also not addressed. These issues are beyond our limited scope in this study.

One last point is perhaps more relevant. By basing the tutorial on Minecraft, many beginners might hope to get started with actually programming in Minecraft, which is done in JavaScript. The tutorials do show pop-up windows showing JavaScript code equivalent to the jigsaw blocks, but go no further. We understand that the complexities and dependences in real JavaScript would present great difficulties to address within the Hour of Code framework, but we also anticipate that some beginners might find this disappointing.

Table 9.10: The Affective Heuristics (Version 3).

Affective Heuristics (Version 3)
H1: Design elements and modalities should support the affective learning strategy
H2: Ensure help and documentation is provided where needed but does not distract from affective learning strategy
H3: Maintain visibility of progress, affirming challenges already overcome, and those remaining
H4: Allow the user freedom to explore but also to return to the previous step
H5: Avoid or prevent actions with neither feedback to help overcome, nor affirmation when success
H6: Visualize options clearly to encourage exploration
H7: Tailor actions to be encouraging at first and efficient later, while learners are attempting to overcome challenges
H8: Challenge learners and provide constructive feedback if they fail, and affirming success when they succeed
H9: Match the learners world view in affective strategy and multimodal support
H10: Maintain interface cohesion to support affective strategy

9.4 Minecraft Hour of Code and Affective Heuristic Evaluation

This section addresses the use of the affective heuristic evaluation proposed in Chapter 8 to evaluate affect in the user interface, educational design, and content of Minecraft Hour of Code to see if the software supports the educational objectives, narrative and persuasion to make the learners more engaged in learning programming. The affective heuristic evaluation is based on Nielsen and Molich [104], and adapted with affect supported by Norman [105] and Kort et al.'s emotional cycle of learning model [79] as well as multimodal design based on Sankey [132]. We have explained the full rationale in more detail earlier, in Chapter 6.

9.4.1 Method

The study was performed in the same laboratory as earlier. In this study, we (GhasemAghaei with Biddle observing) applied the evaluation method as discussed earlier. We reviewed the heuristics, explored, and then reflected on the software. We then wrote comments about the environment filling in a table with respect to the different heuristics. For heuristic severity we adapted Nielsen's severity ratings for usability problems [103], changing their emphasis to reflect emotional educational impact. We explicitly emphasized on the learning objective and the modality that was employed.

For the learning objective we considered the following. This system was designed to be fun for a student working alone or in a classroom where the instructor has minimal preparation or computer science background. It has fourteen tutorials to learn the basics of JavaScript programming concepts that are: use of commands, repeat loops and if statements. It aims to create a quick and enjoyable experience for students and instructors who are new to computer science.

The modality involved a 2D block world with programming using jigsaw pieces. There is also an avatar to represent the user, animation of the block world, and engaging audio samples. Moreover, it used a narrative and persuasion quasi-modalities.

We went through the interface systematically. For example we added blocks to the workspace, trying to think of the intended user. We checked in each step if the system state to consider

the learning objective and the modality employed. At the end of the suggested tasks, we filled in a table with issues based on the ten heuristics. Table 9.10 shows the affective heuristic evaluation. This is the set of heuristics following revisions discussed in chapter 8.

9.4.2 Findings

In this study, the evaluators read the heuristics, explored, and then reflected on the environment. They then wrote comments including the interface element that was violated, problems that illustrated poor considering of affect and modality, suggestions and recommendations for solutions, and the severity based on Jakob Nielsen's five-step rating scale [103]. Tables 9.11 and 9.12 show the affective heuristic evaluation results. Each of the ten heuristics led to a useful comment at least once.

9.4.3 Discussion

The heuristics invited reflection on modality as well as affect. Our findings show that the affective heuristic evaluation led to identification of many problems and potential solutions as shown in the tables. By reviewing these results, we can make some general observations.

Narrative and continuity: There was no clear continuity between the steps. One of the tutorial steps was to chop down trees (Figure 9.6 top left), and the next one was to *shear* sheep (Figure 9.6 top right). The wool that resulted from shearing was never used or mentioned again. It did say that wood was useful for building and later steps did build walls for a house, but there was no indication of the wood came from chopping down the trees. It could have been made more continuous making it clear that the trees were used to build the house and also by using the wool to make carpets etc. They are building a house without using the wood or the wool, and therefore, there is no strong continuity and storytelling between the steps (Figure 9.6 bottom). This might also have been a useful opportunity to introduce quantities of wood or wool as variables.

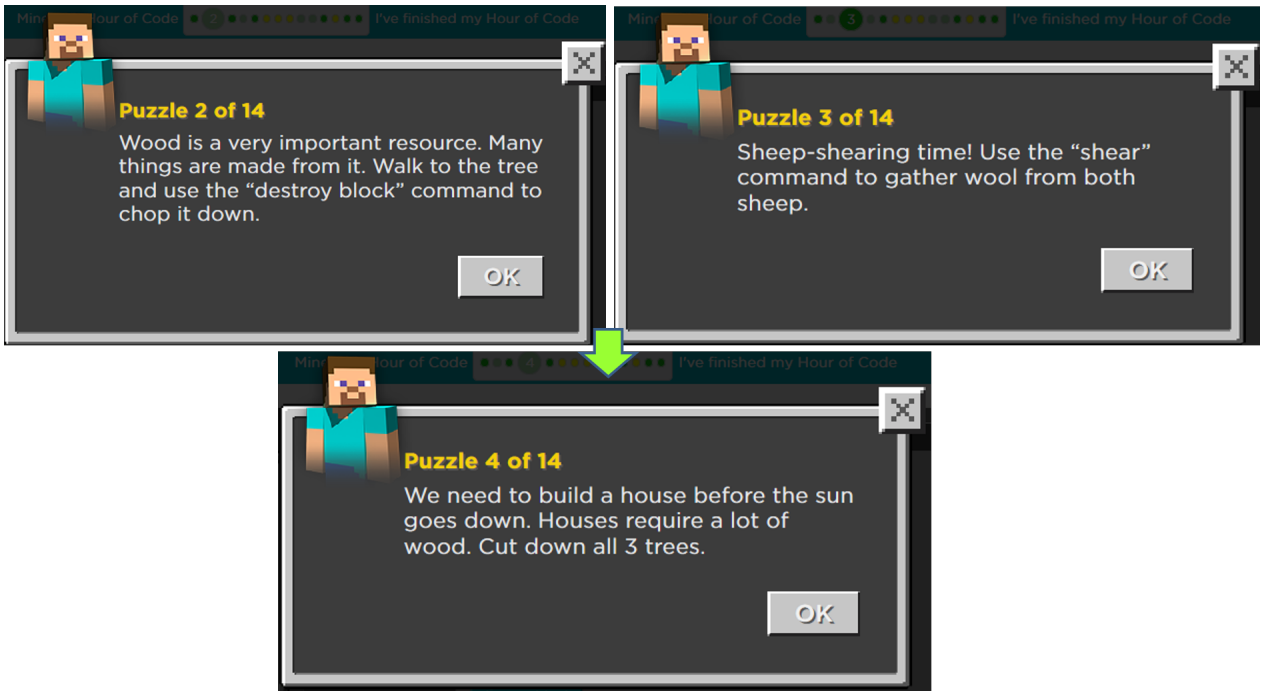


Figure 9.6: Narrative and continuity.

Encouraging Exploration: There was not enough encouraging to explore at each step, and no clear story for exploration. The commands provided were specific to each step, meaning there were not all commands provided in each tutorial e.g. *shear* was provided in the sheep-shearing task but not in other tasks. The commands were changed for every tutorial, which is a real limit to learning as it restricts exploration, and people would expect specific commands for each programming situation. Figure 9.7 shows the command blocks for tutorials two and three. They can use "destroy block" in the second tutorial step but it is not seen in the third tutorial step where they have to use "shear" instead.

JavaScript Code: As we mentioned, the real Minecraft software is programmed using JavaScript. Minecraft Hour of Code uses the jigsaw command blocks. But when the task is complete, is shows the real JavaScript code. But a learner would not know what they have to do with this code; there is a poor connection between the JavaScript code and the block commands that represents that code, and there is no affective feedback provided to the learners. Therefore, the connection between them is not made; we can not edit or change the JavaScript code. They are not really learning the JavaScript language. The JavaScript code

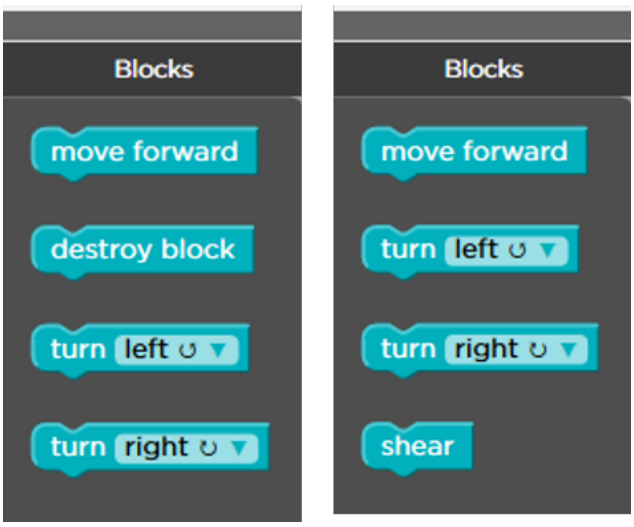


Figure 9.7: Encouraging exploration.

is not even nicely formatted, and there is no interaction with the learners, which can lead to disappointment for them (see Figure 9.8).

9.5 Summary

In this chapter we presented our evaluation of the multimodal affective learning support in Code.org's Hour of Code tutorial set based on the popular game *Minecraft*. We applied our two proposed evaluation methods called the affective walkthrough and the affective heuristic evaluation, modeled on the widely used cognitive walkthrough and heuristic evaluation. The new methods keep the procedural framework of the cognitive walkthrough and heuristic evaluation, but as the evaluator steps through, the questions and heuristics are about emotional support in education.

The Minecraft Hour of Code software applies various modalities to support learning, talking a visual approach with a engaging game-like graphics, a model world, and a programming language using jigsaw-like pieces (as do Scratch and Alice). The tutorial instructions add a narrative aspect, whereby the programming tasks involve acting out a story in the model world.

Our experience in conducting the evaluation was positive and enlightening. We recognized

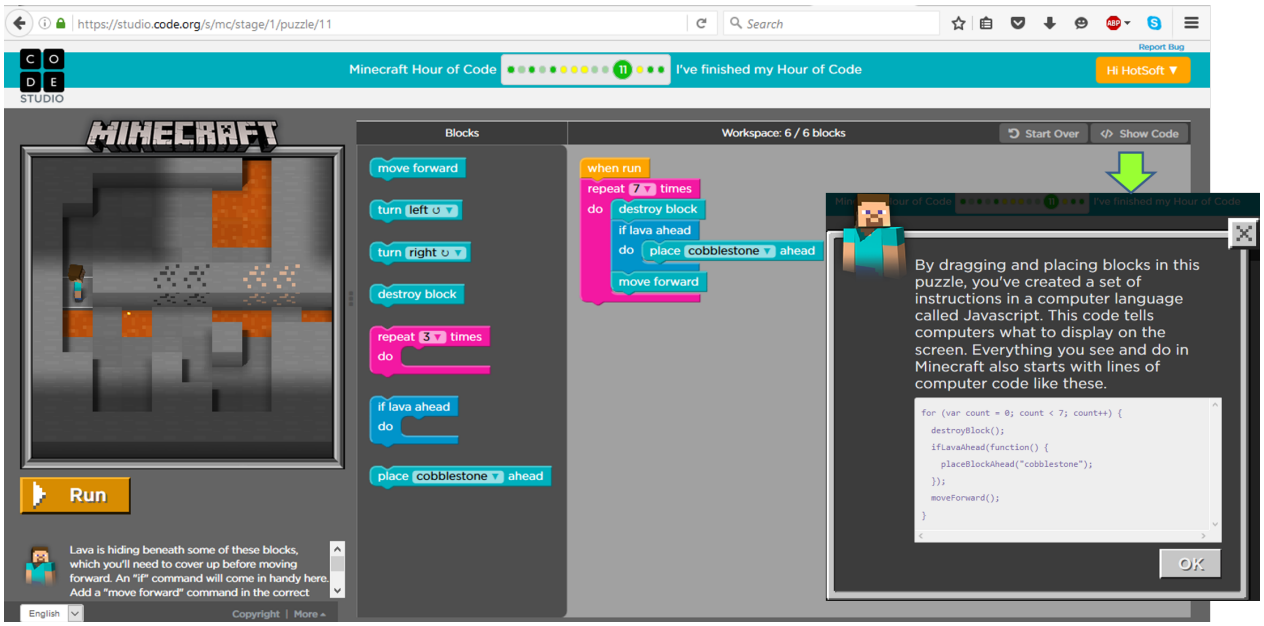


Figure 9.8: Blocks with the generated code.

elements in the Hour of Code design that we would not otherwise have noticed, and also identified ways in which the design might be improved. We see as especially valuable the interplay between the spatial model world, and the narrative stemming from the instructions. On the other hand, we felt that some extra freedom in command choice might in several places be helpful. Also, where challenges were provided by suggesting shorter solutions were possible, it would be helpful to add a capability for the learner get hints so they can overcome the challenges if they are stuck. We also recognized the limits of the tutorials, which really only learners with the very early steps in learning to program.

We appreciated the focus of the affective walkthrough and affective heuristic evaluation focus on modality benefits and teaching goals, and most importantly its use of Kort et al.'s model of the affective cycle in learning. We feel they offer a helpful and systematic approach to evaluating affective learning in multimodal software.

We conducted this case study with Version 3 of both the Affective Walkthrough and Affective Heuristic Evaluation. These were based on refinements from our earlier case study, but we found no reason to refine them further.

Table 9.11: The Affective Heuristic Evaluation H1 to H5.

Affective Heuristic Evaluation Involved	Interface Element	Problems not considering affect and modality	Suggestions for Solution/ Comments/ Recommendations	Sev (0 to 4)
H1: Design elements and modalities should support the affective learning strategy	Show code	The JavaScript code is provided but without affective support	Have it be more interactive Some kind of interaction to get them more engaged Use more physical/material design elements	4
H2: Ensure help and documentation is provided where needed but does not distract from affective learning strategy	Affirmation pop-up window	It does not provide all the possible things or at least some small subset of possible things	Give hints and use affective persuasion	3
	Tasks, command sequences	Reasons of using commands not provided	There can be a multimedia content e.g. video and providing the purpose and goals of using commands	3
H3: Maintain visibility of progress, affirming challenges already overcome, and those remaining	Entire environment	Pretty good job for this part	Progress timeline can also show the steps	0
H4: Allow the user freedom to explore but also to return to the previous step	Tool Box	Software only provides us with number of blocks we need, and it is not encouraging exploration all that much	It should provide us with other alternatives	3
	Minecraft Hour of Code tutorials list	No label next to each task in the timeline	Better to have label	4
H5: Avoid or prevent actions with neither feedback to help overcome, nor affirmation when success	Affirmation pop-up window	To overcome the challenge with minimum lines of code there is no feedback to help overcome	Provide some feedback to how to be able to get to minimum lines of code	3

Table 9.12: The Affective Heuristic Evaluation H6 to H10.

Affective Heuristic Evaluation Involved	Interface Element	Problems not considering affect and modality	Suggestions for Solution/ Comments/ Recommendations	Sev (0 to 4)
H6: Visualize options clearly to encourage exploration	Tasks, command sequences	No motivation and role of narrative could be better	Provide a good narrative about using command sequences in this task	2
	Tasks, repeat loops	Narrative is not considered well to make it an exciting experience	Provide a good narrative about using repeat loops in this task	2
	Tasks, “if” statement	No clear narrative.	Provide a good narrative about using "if" statement in this task	2
H7: Tailor actions to be encouraging at first and efficient later, while learners are attempting to overcome challenges	All tasks	No clear narrative and connection between the steps	Easier tasks are earlier but there is no clear connection between the steps	2
H8: Challenge learners and provide constructive feedback if they fail, and affirming success when they succeed	Puzzle 2 to 3	No challenge	add some challenge to puzzle 3	0
H9: Match the learners world view in affective strategy and multimodal support	Puzzle 4	Starts well, but then give examples, and then does not continue well, e.g. not building a house. Sun does not go down. Story is not correct	It is not complete but it is pretty good at creating a child world.	1
H10: Maintain interface cohesion to support affective strategy	Tool box	No learning to choose a strategy in presence of many possibilities with different strengths	Add more affective strategies	3
	Show Code	Also JavaScript not bringing engagement and motivation	Using affective strategies to have interaction with users	4
	All steps	Narrative is not well done	Have a better storytelling and continuity	2

Chapter 10

MADE Teacher's Dashboard - Emotion Dashboard

10.1 Introduction

While conducting our case studies of our evaluation techniques, we begin to recognize some inconsistencies between our approach to evaluation and to design. On reflection, we realize that the common HCI techniques we adapted were not wholly complementary. For example, while cognitive walkthroughs are a well-respected technique for evaluation, there is no direct equivalent technique for design. In our affective walkthrough technique, we found the importance of Kort et al.'s model [79], but we did not consider it in the design methods. Therefore, we needed to go back and support that in the design methods as well. Moreover, these issues led us to recognize the need for better support for the teacher, as well as the student.

10.2 Task Sequencing

Realization: The problem was that Kort et al.'s model and phases proved valuable in our evaluation techniques, but were not considered in our design techniques. Our design did not support task sequences as well as we could have done.

Action: We returned to our design case study, and added new design elements. To support Kort et al.'s model, software might provide a sequence of tasks with difficulty and affective support that vary. We therefore changed our software to take the learner through a sequence of tasks, each presenting a goal for the learner, along with affective feedback. We also added support in our database for these tasks and sequences. By doing this, the user is able to learn ratios and proportions by following a customized task sequence with receiving feedback

in each step.

With our ratio study, first the system might start helping the user to explore. Next, giving a simple task and afterwards a complex task. Finally, providing the affirmation feedback when all the tasks are done. The changes in MADE Ratio included the following items:

- Ratio – First, there will be a simple task, and after the task completion there will be a more challenging task.
- Feedback success – If the task has completed successfully, teacher can provide customized feedback to a user.
- MaxTime – Each user has a time period to complete a specific task.
- Feedback Failure – If a user is not able to complete a task in a time period, the system will show a failure feedback. The teacher is providing feedback to explain and encourage.

We wanted the teacher to come up with the task sequences, and customize the feedback messages with affective strategies. We illustrate this sequence in Figure 10.1. The system provides each user a customized list of tasks from simple to challenging, when they are interacting with the system.

Later, we also realized our design techniques had a focus on learners, but did not address the teacher's role in a supporting way. For example, one issue we recognized was that while our design technique emphasized affective personas, we did not consider how they might be created or maintained.

Realization: Up until this point, our design techniques called for different personas to be accommodated, but no way for them to be easily created or maintained by teachers. Our software simply allowed a database to store each learners settings of affective messages or media.

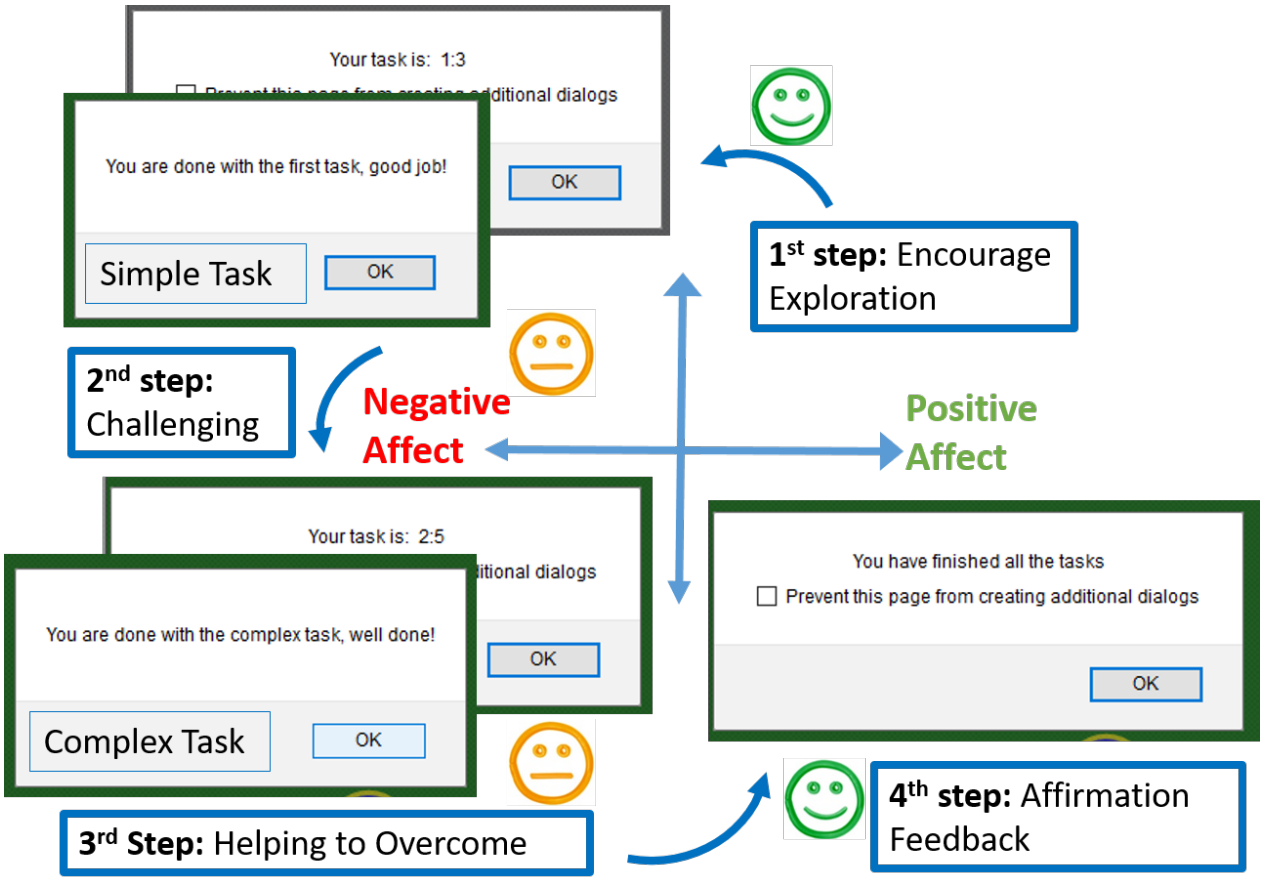


Figure 10.1: Simple-complex task model for a specific learner.

Action: We decided that learner and personas was properly the responsibly of teachers, and should be supported as such. We therefore created a teacher's dashboard for this purpose.

The dashboard allows teachers to access learners and personas, and manage the affective settings. In particular, the teacher may link a learner's setting to that of a persona. Thus, creating a persona may be done once, and it can be reused for the students that match that persona, as several students match the affective strategies from the affective persona model. For example, a teacher might create and reuse personas for a students lacking confidence, student more interested in music, or students with who have trouble reading complex text.

Figure 10.2 illustrates the dashboard support for personas. This figure shows the class list on the left side, in the center are the detail settings, and on the right side, the list of personas. The teacher can change the settings and have a specific affective strategy for a specific learner

or persona. While recognizing the need for this kind of support for teachers, our work on

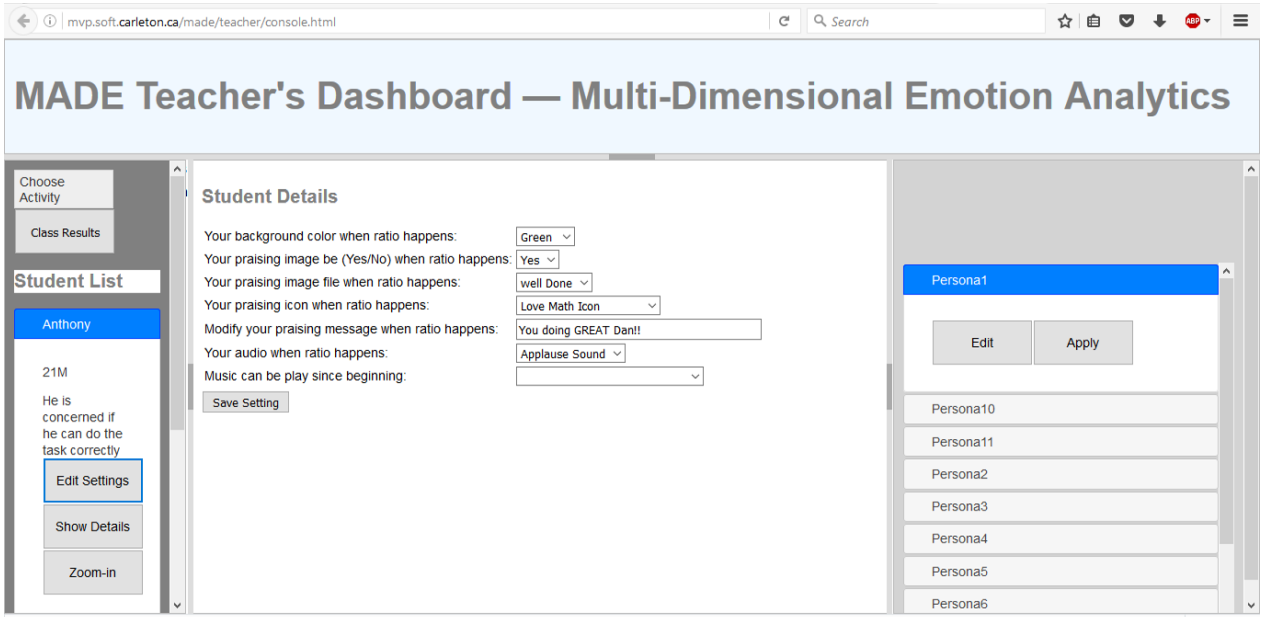


Figure 10.2: Dashboard persona-learner edit settings.

evaluation made us realize a more far-reaching kind of support was indicated. We address this in the next section.

10.3 Emotion Dashboard

In our case studies of evaluation presented in Chapters 8 and 9 we found that assessment of affective software required the teacher's perspective. It occurred to us that we could extend our teacher's dashboard to support this directly. We could allow a teacher to *see* the emotional state of a student as they use educational software.

A student's emotions, such as sadness, happiness, and surprise, often significantly impact her or his behavior and performance in a learning environment. When you are a teacher in a traditional classroom, you can judge the emotional status of the students; but when your classroom is online, how do you know what the students are feeling? When there is a co-located environment, the teaching relationship can be managed in a way that is impossible in an online environment. While learning online, students experience emotional reactions just

as they do in an ordinary classroom. For example, difficult e-learning tasks can lead to failure and frustration, resulting in fear. Recognition of student emotions in such situations would help teachers. Eventually, students might learn to help themselves to develop more self-awareness and self-management [154]. It may be possible to create e-learning systems with affective intelligence to support this goal. Our approach is more immediate, and rather than embedding affective adaptability in the system, we propose better supporting the teacher. Our suggestion is that the online educator should be able to monitor learning and emotion, and customize each student's learning tasks accordingly. While students work on the tasks, the system could show a graph of their emotions over time and display it with different visualizations showing proportions of emotions in specific time periods. In comparison to other research on learning dashboards, e.g. Muldner et al. [98], our approach does not use artificial intelligence, but rather visual analytics for the teacher. To support this approach, we extended the MADE Teacher's Dashboard to monitor emotion in e-learning. It is a visual analytics system developed to let the teacher see emotional states of the student. In our studies of these techniques, we found that even with careful design and evaluation for affective education, there was still a need for teachers to monitor and manage the learning process, motivating our work on the Teacher's Dashboard.

In the following sections, we first review related work, and identify our design rationale. We next present our design process, from sketch to functional prototype system, and especially our innovation information visualization. We then discuss the issues that emerge, and plans for future work.

10.4 Emotion Visualization Background

10.4.1 Models of emotion

Mehrabian and Russell [92] described what has become a common model of the practical dimensions of emotion. These consist of positiveness or valence (comparing states of pleasure e.g. happiness, with states of displeasure e.g., sadness), arousal (degree of excitement; states of low arousal e.g. quiet, with states of high arousal e.g. surprised), and dominance (degree

of aggressiveness; state of low dominance e.g. in control, with states of high dominance e.g. controlled). These are associated with a person's affective reaction to a wide variety of stimuli, and are pervasive in organizing human judgment for a wide range of perceptual and symbolic stimuli. Of course, we acknowledge that affect and emotion are topics rich in depth and detail, and we expect more understanding to emerge from ongoing research. In our work on affect, we focus on the emotions that arise in the learning process.

10.4.2 Emotion and learning

As discussed in Chapter 4, Bloom's taxonomy [10] breaks learning objectives for students into three domains, being cognitive, affective and psychomotor. The affective domain is our focus, and while Bloom's work includes learning in the affective domain itself, our particular interest is in the use of affective teaching strategies to support all kinds of learning. We are particularly motivated the model of affect in education from Kort et al. [79]. This model was very useful in helping us refine our techniques during our case studies. Kort et al.'s model identifies four phases of emotion commonly cyclically involved in learning. This is based on the circumplex model of emotion, first developed by Russell, who proposed an emotion model based on the two dimensions of arousal and valence (see Figure 10.3. Kort et al. make connections between what emotional state the student is in and, from that, raises consideration of what help he/she needs. They proposed an affective model of interplay between emotions and learning; a four quadrant learning spiral model in which emotions change when the learner moves through the quadrants and up the spiral. It is well known that students' results can be improved with the right encouragement and support. Figure 10.3 shows this model, which attempts to merge the emotion and the cognitive dynamics of the learning process.

To simplify: when students start the task, they are happy, then things go wrong and they are sad. Later, things start to pick up again and they happy again. We also see as related the emotional flow state [31], where the two main elements influencing flow are challenge and skill. Flow emerges when skills match the challenges of a learning situation. As regards to flow, an educational system requires conditions between activity and individual, and internal experience, and flow condition include: clear goals, immediate feedback, balanced between

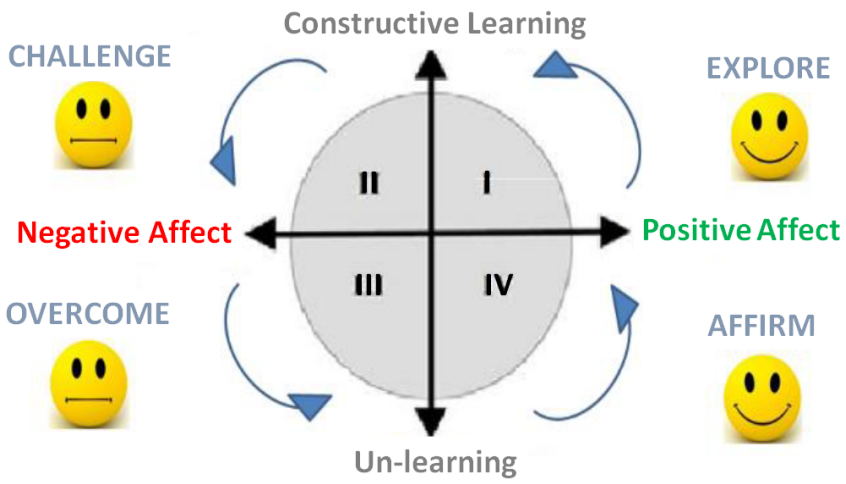


Figure 10.3: Kort et al.'s affective model of education.

challenges and skills, easily controlled, merging of action and awareness, concentration on task at hand.

10.4.3 Emotion and educational software

Modeling student emotion has increasingly important for computational teaching systems. One limitation in these systems is they do not systematically examine the relationship(s) between student affective state and learning outcomes [98, 154]. Theories and technologies are needed to understand and integrate the knowledge of student affect (e.g., surprised, happy or sad) into learning events for the educators. Woolf et al. [154] identified emotion dimensions (valence and arousal) that, together with knowledge on-task engagement, can represent desirable or undesirable states related to student learning. These would also related to physical behaviours connected to emotion, indicating and corresponding to student affect. Their approaches to affect recognition include human observation, hardware sensors and machine learning techniques. We believe that a simpler approach may now be possible that requires no special hardware and allows the teacher to monitor student emotions.

10.5 Basis for Design

10.5.1 Facial expression of emotions

Micro-expressions were first identified by Haggard and Isaacs [60]. They discovered the micro-momentary expressions while scanning motion picture films. Later, Ekman et al. [48] found a selection of emotional labels that fit facial expressions to be universal including happy, sad, surprise, fear, disgust and anger. There has been interest in human ability to modify these gestures, including simulated expressions, neutralized expressions, and masked expressions. While we understand the importance of these possibilities in a range of applications, including teaching, we do not explore them further in our current work. Instead, we focus simply on the canonical forms of Ekman's categorization of emotions based on analyses of facial expressions. Figure 10.4 shows examples of these six emotional facial expressions.

Figure 10.4: Basic facial expression phenotypes. 1: disgust; 2: fear; 3: joy; 4: surprise; 5: sadness; 6: anger. Posed images from Kanade et al. [72]. Copyright images used with permission.

10.5.2 Computer detection of facial expression of emotions

Based on the work above, understanding facial expressions can be an important behavioral way to study emotion, cognition, and social interaction. Bartlett et al. [5] showed that computer vision can identify facial gestures. More recently, there has been work on software and algorithms for many practical applications in this area. For example, there are commercial products such as Noldus FaceReader (www.noldus.com/facereader), which can be used in research applications where tracking emotional state of participants is important. Saragih et

al. [133] introduced algorithms based on Constrained Local Models (CLM) for facial gesture recognition. Audun Mathias Øygard [111] developed Clmtrackr, software based on the algorithms of Saragih et al., written in JavaScript and able to be used within web applications. Clmtrackr is open source and freely available, and supports tracking of facial features and detection of emotions in real time, identifying the six different emotions discussed above. This set of emotions emphasizes valence (pleasure/displeasure), and other dimensions (arousal and dominance) are less explicitly identified.

10.5.3 Unfolding the model of Kort et al.

As we discussed above, Kort et al.'s model [79] suggests several phases of emotion in the learning process. In brief, first it starts with positive affect, then there are the challenge and overcome phases, which we have some negative affect, and finally the affirmation that has positive affect. We are interested to track the learner's progress over time. Hence, we want to "unfold" the model with time along the x-axis. Figure 10.5 shows our suggestion. The valence of emotions progresses from positive (happy and surprised) through negative (sad and angry).

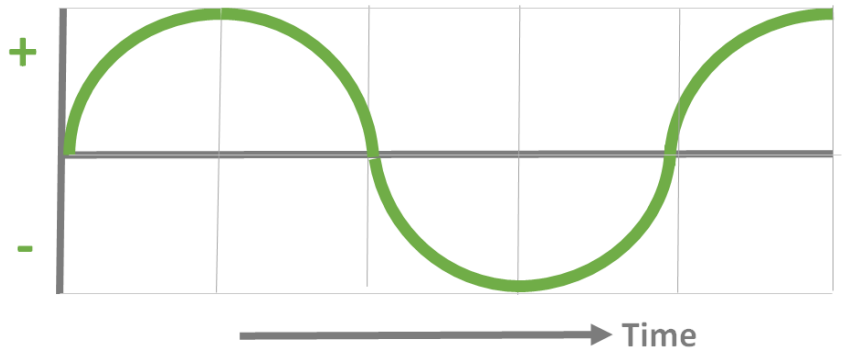


Figure 10.5: Model of the emotional state in learning based on unfolding the model of Kort et al. [79].

10.6 MADE Teacher's Dashboard – Multi-Dimensional Emotion Analytics

10.6.1 Architecture

Figure 10.6 presents a schematic of the context for our current teacher's dashboard prototype. Technically, there are four main components to the data architecture: the teacher's dashboard, the student's screen, the web server, and the database. In the figure, the educational software that the students use is shown on the right. We show several educational applications, corresponding to three case studies that we have conducted. This software is augmented with library code to access input from a camera, and identify emotional state as discussed earlier. The software runs from within a web browser, and communicates with a web server as shown in the center; this stores emotional information in the database. Our teacher's dashboard is shown on the left side. Its also communicates with the web server which accesses the database. The information from the student application can then be displayed on the teacher's dashboard. As an example and test for our architecture and the

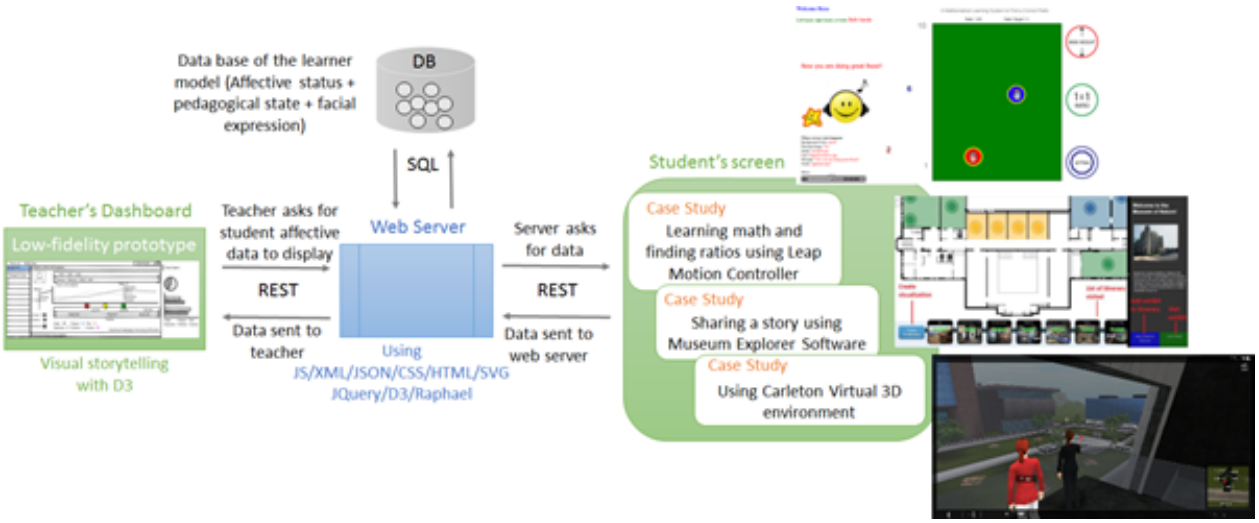


Figure 10.6: Architecture: overview of the Teacher's Dashboard in context.

Teacher's Dashboard, we used our MADE Ratio software explained in Chapter 7, which is web-based software for learning about mathematical ratios, designed using Affective Personas and Affective Essential Use Cases. To adapt the software to our architecture, we used the Clmtrackr software library; this accessed the webcam, and in real-time analyzed the image of

the user's face, rating the intensity (on a scale from 0.001 to 1.000) of 6 emotions (happiness, sadness, surprise, disgust, and anger). These ratings were calculated every second, and for efficiency these were uploaded to our database in bulk once every minute. This process is simply uploading one array of sixty numbers per minute per user to our server and database. This is a small load on any reasonable server configuration. For example, testing on our modest server (Intel i686 2.6 GHz with 8 GB RAM) suggested an ability to handle one hundred simultaneous users without difficulty. This is also much more efficient than uploading video images, and will present fewer privacy issues.

We also modified the MADE Ratio software to upload details for every significant event that occurred as the software was used. Every time the user began a new task, succeeded or failed, etc., an event was entered into the database. This enabled exploration by the teacher of correlations between cognitive learning events and emotional states.

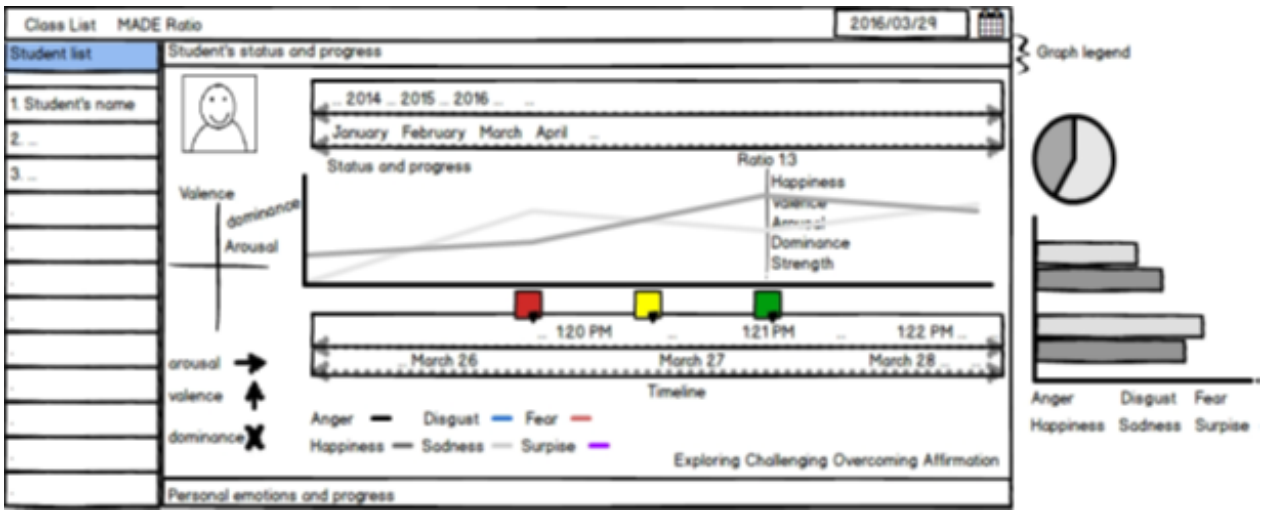


Figure 10.7: Teacher's Dashboard low-fidelity prototype.

10.6.2 Low-fidelity prototype

To design the teacher's dashboard, we followed standard practice in HCI. We first identified the user tasks, then developed "low-fidelity" sketches, and then walked through how they would be used to perform the tasks. We used an iterative approach, with our findings from the walkthroughs being used to change or refine the sketches. We used the Balsamic Mockups

software (https://balsamiq.com/) to develop the sketches. Figure 10.7 shows one of the later sketch prototypes. On the top we show the learning activity (e.g. MADE Ratio), on the left we show a list of students and their characteristics, and in the main window we imagined a graph with time on the horizontal, and the emotion dimensions graphed on top. On the right we imagined summary graphs of the emotion distributions.

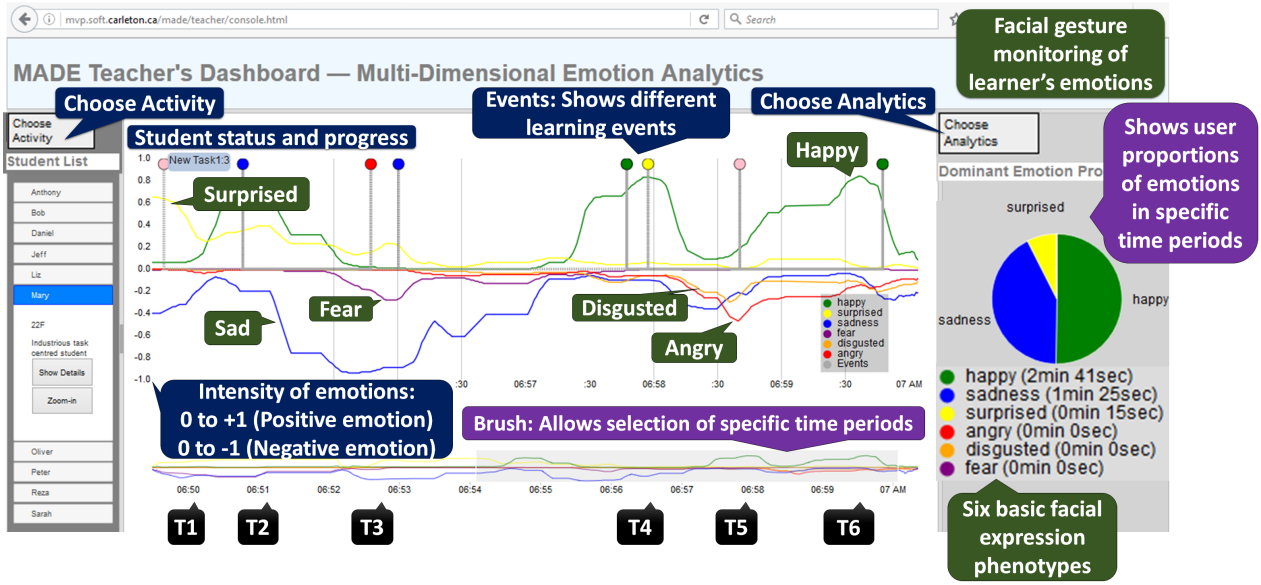


Figure 10.8: The Teacher's Dashboard functional prototype.

10.6.3 Functional prototype

Figure 10.8 shows the main screen of our functional prototype. This was based on our low-fidelity sketch, but is fully implemented and works within the architecture we outlined above. On the left side of the dashboard we have the student list; the instructor is able to see the details of a specific learner on the graph and can zoom-in. For the selected activity and learner, it shows six different emotions shown by colored lines as they vary over time. Emotions associated with positive valence (e.g. happy) are shown above the axis (0 to +1), and emotions associated with negative valence (e.g. sad) are shown below the axis (0 to -1). Below the main graph is a visual analytics "brush", which allows the user to select and zoom in on a specific time period. Vertical lines with "balloons" show specific learning events, allowing the teacher to related cognitive and emotional states. When the teacher hovers

their cursor on the balloon, a tooltip with descriptive text is shown. On the right side of the dashboard we have the analytics, showing user proportions of emotions in specific time periods (which can be selected by moving the brush).

Figure 10.9 shows the effect of brushing to select a ten minutes region including the start of a task (pink balloon), seeing both hands (blue balloon), losing hands (red balloon), completing task (green balloon), and receiving positive feedback (yellow balloon).

Figure 10.10 shows a closeup of the summary pie graph at the right of Figure 10.9, and the alternative bar graph view. These graphs show the proportions of dominant emotions in the time period selected. We identify a dominant emotion as one with a level of greater than 0.5 or less than -0.5. In the graphs, the category of *none* represents the amount of time where no emotion was dominant. The category of *mixed* represents the amount of time where more than one emotion is dominant. In our informal studies we found periods of mixed emotion were infrequent.

On the right hand side of the dashboard, we have the *emotion pie chart* or *emotion bar graph* showing user proportions of emotions in specific time periods.

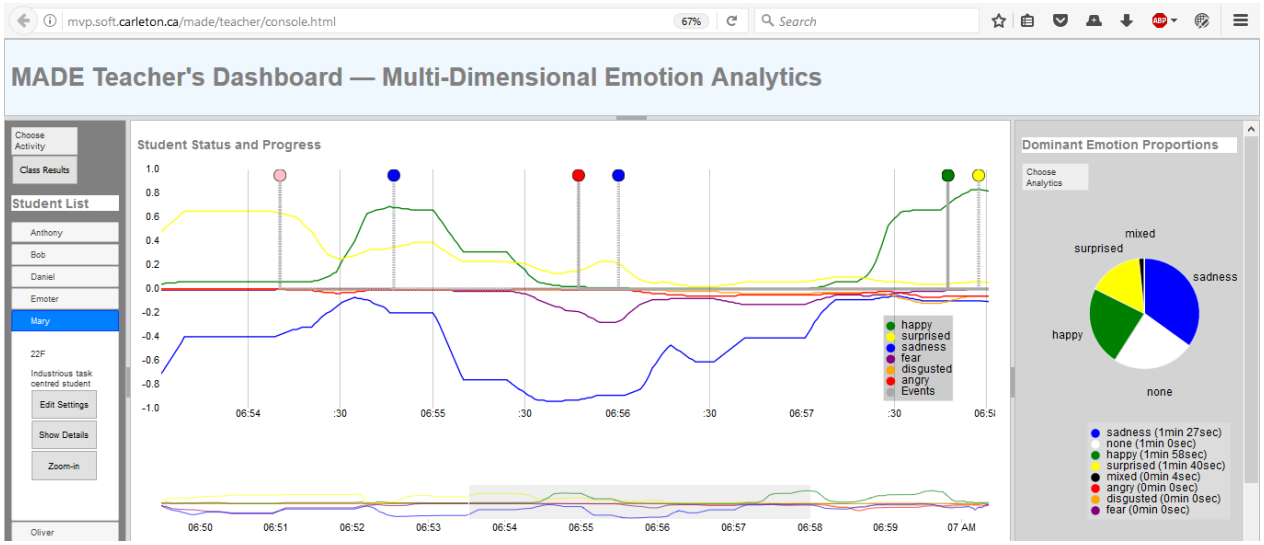


Figure 10.9: Brushed area showing range between beginning of a new task (pink balloon) and completion with affective feedback (yellow balloon).

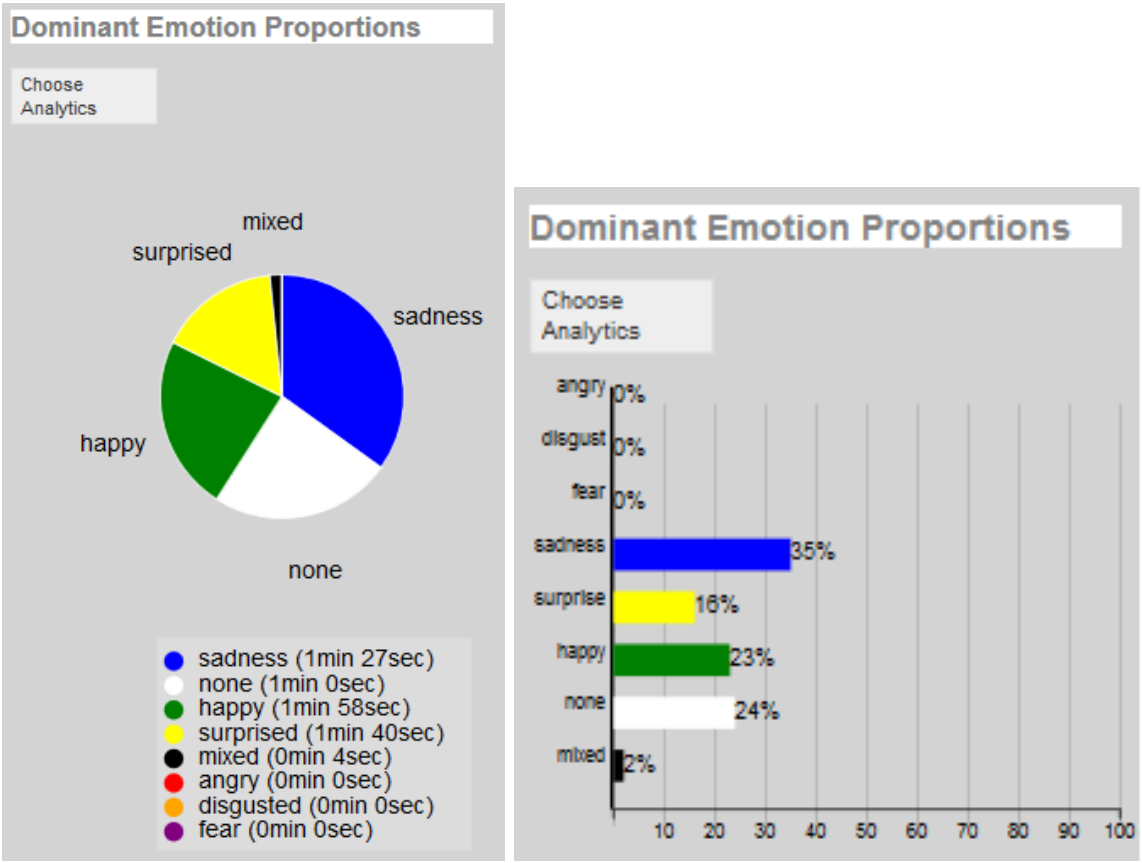


Figure 10.10: Closeup of pie chart and bar graph from Figure 10.9.

10.6.4 Implementation technology

The time series visualizations for emotional data were developed using Data-Driven Documents (D3) (https://d3js.org/), a JavaScript library for producing dynamic, interactive data visualization for the web (bears a resemblance to lower-level graphics libraries such as Processing and Raphael) with HTML5 for page content, JavaScript for interaction, SVG for vector graphics, Cascading Style Sheets (CSS) for aesthetics, and the jQueryUI JavaScript library for screen layout. The student software and the teacher software communicate with the web server using the ReST (Representation State Transfer) protocol. The web server communicates with the database using SQL. Vision based emotion detection is done using the Clmtrackr JavaScript library.

10.6.5 Usage walkthrough

To illustrate how the teacher's dashboard can help, we now present a step-by-step walkthrough of how it might be used. We start with the teaching strategy; what the instructors' expectations might have been, and how the dashboard can help them understand what is actually happening in the e-learning software. Consider a teacher using our MADE Ratio software, which helps learners explore mathematical ratio by tracking learner hand positions. In that software, the teacher might select tasks for the student, such as finding ratios of simple 1:3, or ones that might appear more difficult to understand, e.g. 3:9 or 2:5. The teacher can also select affective strategies for each student. For example, a teacher might know that a student, Mary, likes frequent feedback with audio elements, such as praising text messages, Justin Bieber music, and applause. The teacher can set these to accompany appropriate learning events that occur in use of the MADE Ratio software. Imagine that the teacher has chosen those tasks, and those affective feedback settings. The teacher can then monitor or review what happened using the teacher's dashboard, as illustrated in Figure 10.8. We can now follow the events shown on the timeline from left to right over a 10 minute period marked T1 through T6 on our figure.

- T1: First, Mary is given an easy task (pink balloon, finding ratio 1:3). Initially, she is surprised (yellow line).
- T2: Mary becomes happy (green line), and we can see her hands are being successfully tracked (blue balloon). After that, she becomes less happy, and increasingly sad (blue line), and we see her hands stop being tracked (red balloon).
- T3: Mary's hands start being tracked again (blue balloon), and she continues to work with the system, initially with neutral emotions, but eventually becoming happier (green line).
- T4: Mary succeeds at the task (green balloon), and is happy, then receives positive feedback (yellow balloon), and becomes happier still (green line). Her happiness then diminishes.
- T5: A new task is then assigned (pink balloon, finding ratio 3:9). Initially she is a

bit angry at the challenge (red line), but she engages, the anger goes away, and her happiness increases (green line).

- T6: Mary completes task (green balloon). She is happy (green line), although then the happiness dissipates.

If you were the teacher, you would know the tasks, and know about Mary. You might be aware of the model of Kort et al., and expect emotions to vary during the course of a learning activity: exploring, challenging, overcoming and affirmation. In our example, much that happens reflects an ideal state. Figure 10.11 shows how emotions rise, fall, or stay constant in each time period. However, between T2 and T3, or between T3 and T4, perhaps more could have been done. You do not want the learner to give up; you want her to be on-task and active, engaging and involving her. For example, perhaps there could be another affective intervention here, where it could be saying: "you should keep trying" or "you are almost there". Alternatively, you could take a meta-cognitive strategy, and suggest that Mary might take a different approach, or review earlier learning materials before returning to the tasks. More generally, you might adopt the concept of Flow from Csikszentmihalyi [31], and consider whether Mary is either too bored or too frustrated, or whether she is in the desirable Flow channel.

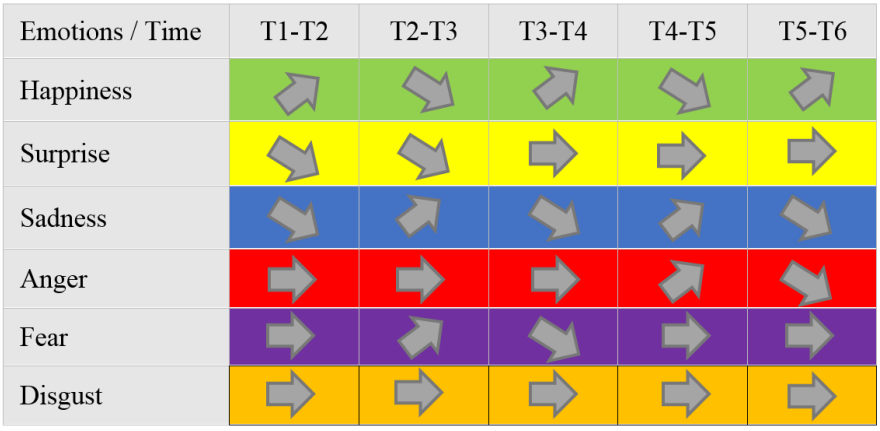


Figure 10.11: Mary's emotions during the task.

We also allowed Self-Assessment Manikin (SAM) emotion queries [14], the most popular visual self-report measure in the education software (see Figure 10.12). We displayed those

results on the teacher's dashboard as an alternative to the results from the camera-based data. SAM can be used to also report dimensions such as arousal, and dominance. This was implemented in our MADE Ratio software to allow learners to easily enter their affective state at the end of each task.

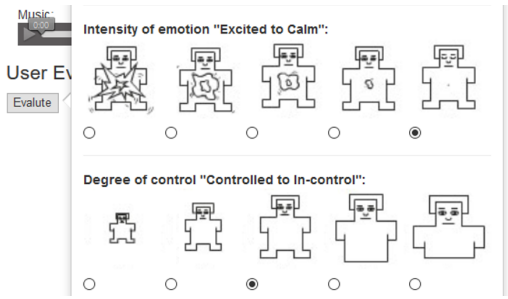


Figure 10.12: Evaluating arousal and dominance with SAM.

10.7 Summary

In an ordinary classroom environment, it is easy for teachers to monitor the emotional state of students as they work, and teachers can provide affective support accordingly. In an online environment, this is all much more difficult. In this chapter we presented our work to address this challenge. It involved augmentation of learner software with computer vision software to identify learner emotions, and presents this emotion data, together with cognitive learning software events, in a visual form with an affective e-learning teacher's dashboard. Our principal contributions are the use of web-based emotion tracking for education software, and the visualization techniques displayed on the teacher's dashboard.

Our prototype system is working as described, but there is clearly much more research needed. We conducted preliminary testing on the MADE Teacher's Dashboard, with ourselves as participants, to check the functionality and that the emotion recognition was working reasonably. Much further study is needed. One issue related to the privacy concern, that we first considered in Chapter 7. It might be that people would be less concerned, because this approach records data rather than photographs or video, but it may also be that any recording of emotion may be challenging. Another issue related to the interpretation and utility of the visualizations. Would teachers understand our proposed graphs? Would they want additional

features, such as aggregation over students or tasks? This question requires more explicit study. Finally, of course, the algorithms and software for emotion detection are areas of research themselves, but beyond our scope.

Chapter 11

Conclusions

To conclude this thesis, we look at our experience with the framework for multimodal software design and evaluation with a focus on affect in education, and application in several case studies. In particular, we described how our techniques were refined after each case study, and how the evaluation techniques led us to revisit our design. We gained new insight about the role of the teacher, leading us to develop a new affect visualization dashboard for teachers. We believe that these experience can help inform the design and evaluation of affective educational applications and may also be applicable to other types of learning systems.

11.1 Review of Thesis

First, we reviewed the work in multimodal systems and affective learning environments. Next, we developed our proposed framework, design methodologies and evaluation techniques for multimodal affective educational systems. Later, we explored some case studies and then refined our design and evaluation techniques, and finally showed the dashboard for the teacher. Throughout this thesis, we have considered how affective strategies lead to improvement, because the system encourages and motivates different learners. In the following subsection we briefly explore our framework, design and evaluation methodologies, the case studies, and the dashboard.

11.1.1 Framework

To address support for affect in multimodal education systems, we proposed a framework called MADE (Multimodal Affect Design and Evaluation). We began with Bloom's taxonomy. Our proposal suggests that a learning objective and learning strategies (cognitive, affective and multimodal) should precede system design. We then suggest system design should involve an HCI approach with both design and evaluation. The stakeholders involve are: students, teachers and system designers. The intent of our framework is to support the system designer to create a system that helps the teacher in facilitating learning activities for the students.

11.1.2 Design Methodologies

One of our goals was that our framework should help designers of multimodal software for affective learning. We decided to adapt existing HCI techniques. Specifically, we chose to focus on Goal-Directed Design (GDD) with an adaptation of personas, and Usage-Centered Design (UsageCD) with an adaptation of essential use cases.

11.1.3 Evaluation Techniques

As well as identifying techniques for design, we also explored techniques for evaluating multimodal software that aims to support affective education. In particular, we focused on adapting inspection techniques from usability research: these have proven effective in evaluation without requiring extensive testing with actual users. We took the two most popular inspection techniques, the Cognitive Walkthrough and Heuristic Evaluation, and adapted them using our framework.

11.1.4 Case Studies

We conducted several case studies. In each, we designed or evaluated systems for affective multimodal education, but our primary goal was to explore the effectiveness of the design

and evaluation techniques themselves. We found a need to refine our techniques.

Case Study on the Design Methods

This case study addressed the use of multiple sensory modalities and affective strategies in learning mathematics by starting from low and high fidelity prototyping with mockups, and later on developing the system considering both Affective Personas and Affective Essential Use Cases.

MADE Ratio: The idea was to drive both the understanding and solving cognitive conflicts between student's implicit assumption and her/his own observable enactment, providing experiences that were recast in terms of emerging mathematical concepts. In the design of this system, we used two design methodologies, Affective Personas and Affective Essential Use Cases. Overall, we were pleased that our design techniques drove the design in valuable ways we might not otherwise have considered. For example, we found the Affective Personas led to our inclusion of feathers to allow customization of multimedia feedback, which was used and praised by our evaluators. Affective Essential Use Cases helped us to consider inserting affect in the interactions in novel ways. In future, we would like to see what other relevant tasks can be specified and used with this system, and if the framework and the system could also be applicable to other domains.

Refinements: As part of the development process, we took an iterative approach, and produced several versions, and did user testing. However, we later realized that several design elements indicated a need for more support not for the student, but for the teacher. For example, there was a need for the teacher to specify personas, and to specify task sequences. We addressed these needs in Chapter 10, by providing support for these in a teacher's dashboard.

Case Studies on the Evaluation Methods

Following our development of the two approaches to evaluation: Affective Heuristic Evaluation and Affective Walkthrough, we conducted two case studies, which we describe in the

following. Instead of evaluating the usability we were trying to evaluate the effectiveness of the modality in this educational software.

Museum Explorer: Our first case study was of software recently developed by our colleagues in order to study how narrative might support school students plan and reflect on a trip to a local museum. To examine our evaluation techniques we recruited participants to evaluate the Museum Explorer system. The participants applied first our walkthrough, and then our heuristic evaluation. We took a qualitative approach, and both audio recorded and took notes for later detail analysis to identify issues.

Refinement 1: With the first participants we ran into issues where they were focused on the usability rather than the affective dimensions, and found that the Affective Walkthrough was too concerned with usability. Therefore, to bring more emotional aspects of the users into focus (and not usability), we emphasized Kort et al.'s affective model [79] more and modified the wording in our techniques. This identifies four phases of learning and the affective character of each: encouraging exploration, introducing challenges, supporting overcoming challenges, and affirming learning. With our heuristic evaluation, we had identified wording issues in pre-testing, and address them. This gave new versions of our techniques, which we used for later participants.

Refinement 2: While Version 2 of our walkthrough succeeded in moving the emphasis away from usability, we found another problem. The nature of the model of Kort et al. means that the affective focus is different at various steps, so we re-worded the walkthrough steps accordingly. We also found a need to refine our heuristics further. We found that in order for participants to use our methods, we needed to state the objectives and the modality advantages more clearly and explicitly. For example, in the case of the Museum Explorer system, the objective was to make learners more engaged and reflective (think about what they learned) and the strategy was narrative, to help the learners tell a visual story. We found that the more explicitly we stated these objectives, the better our participants would understand and usefully apply the inspection methods. These refinements led us to Version 3 for both the Affective Walkthrough and Affective Heuristic Evaluation. These are the versions we used in Chapter 9.

Minecraft Hour of Code: The second case study of our evaluation techniques involved a professional and heavily used multimodal education system, Code.org's Minecraft Hour of Code. This was created to encourage and support people initial learning of computer programming.

We applied Version 3 of our Affective Walkthrough, and found it straightforward and productive. As expected with such a well-respected system, our technique highlighted design success, but also areas for improvement. For the heuristic evaluation, we reviewed Version 3 of the heuristics, then followed the process, writing comments about the system, along with severity ratings and potential ways to address the issues. Two issues that emerged serve to illustrate our findings were narrative and continuity, and encouraging exploration. There was no clear continuity between the steps. There was not enough encouraging to explore at each step, and no clear story for exploration. However, we were satisfied with our two evaluation techniques, and found no further refinement necessary.

11.1.5 Teacher's Dashboard

While conducting our evaluation studies, we had realized that some context was necessary to properly evaluate systems in this domain. It was important to know the learning objectives, and also how the multimodal design was intended to help. We also realized that varying affective personas and task sequences needed support for teachers. Together, all these lessons suggested that we might support teachers in other ways as well. Where teachers are involved in determining affective strategies, they also need support in evaluating their effect. We therefore created a dashboard to support the teachers to evaluate how well affective strategies work.

Teachers are interested to track the learner's progress over time. Hence, we "unfolded" the Kort et al.'s model of emotional state in learning with time along the x-axis as shown in Chapter 10. By providing this emotion visualization graph on the dashboard, the teacher can monitor or review what happened. We are not aware of any other work that takes this approach. In our dashboard, the data on emotion came from analysis of facial expression of emotion. We were able to see the data results in real-time as the user was interacting

and working on a specific task. There are still a number of issues to explore in terms of limitations, including accuracy and efficiency, as well as learner privacy.

We next summarize our research contributions and show how these met the objectives set forth in this thesis. In closing, we discuss research directions based on this work, and offer concluding remarks.

11.2 Research Contributions

This thesis proposed a framework for helping system designers in creating a system for affective multimodal education. It should in turn help teachers in facilitating learning activities for the students.

Our specific research questions were "What interaction design processes would support multimodal software for affective education?, what principled framework would be appropriate?, how might these principles be leveraged for design?, how might these principles be leveraged for evaluation?". We made four main contributions, summarized below:

Contribution 1: Catalogued existing multimodal frameworks, focusing on affect and education, but finding primarily consideration of technical architecture (Chapter 2). We then examined work on affect and multimodality in education (Chapter 3), and created the MADE framework (Chapter 4).

Contribution 2: Proposed two novel design methodologies (Chapter 5), and built and tested the design methodologies on a case study (Chapter 7).

Contribution 3: Proposed two novel evaluation techniques (Chapter 6), and tested and refined the evaluation techniques on two case studies (Chapter 8 and 9).

Contribution 4: Identified the importance of the teacher's role in design and evaluation, and created the MADE Teacher's Dashboard, including novel information visualization of emotion (Chapter 10).

Overall, the practical benefits of this work are articulation and refinement of several techniques for system design and evaluation for affective multimodal education. These are of direct benefit to system designers and evaluators, and of indirect benefit to teachers and students.

There are a variety of limitations to the work in it's current state. Perhaps most importantly, while much of the software explored is aimed at children, we did not involve children in any of our studies. Children may respond differently than we expect, and this deserves specific attention. As we mentioned at the end of Chapter 10, further study is also needed on our MADE Teacher's Dashboard. The accuracy of emotion recognition needs research, and the usefulness of the visualization needs to be explored with teachers. The use of cameras to detect emotion also raises issues of privacy, despite no images being recorded, and these concerns need to be addressed.

11.3 Research Directions

This thesis has contributed to multimodal affective education literature, but it has also raised further questions. In this section, we describe other research questions to be explored.

Explore other design techniques. In this thesis, we adapted two well-known design techniques from Goal-Directed Design and Usage-Centered Design. In future work, we might explore other design techniques. Good candidates would include story maps, which emphasizes sequences of activities [113], and Information Architecture (IA), which emphasizes the structure of a user interface [120]. In particular, we suggest that story maps might be adapted to directly support the model of Kort et al. [79] as a design technique especially suited to affective education.

Conduct user studies of the MADE Teacher's Dashboard. We have created the Teacher's Dashboard to show student emotions while they use educational software. The next step should be to conduct studies with real teachers to check the effectiveness of visualizations. Do teachers understand the graphs? Are teachers able to relate the graphs to student experience while working on the tasks?

Explore other issues with the MADE Teacher's Dashboard. We have to conduct work to check the reliability of facial emotion algorithms to see how well they detect the right emotions. Furthermore, will the learners accept this approach or visualization, and allow their emotions being tracked in this way? If we have the same kind of data for many learners, we could then explore analysis about how to improve the educational experience.

Other application areas. In this thesis our subject was affect in multimodal educational software. However, our affective design and evaluation strategies may also be useful in other areas. Norman has discussed how affect can be used in the design of many kinds of artifacts [105]. Our approach maybe useful in domains related to education, were affect may also be helpful. For example, such domains might include robotics, health-care systems and smart homes.

Bibliography

[1] Dor Abrahamson and Dragan Trninic. Toward an embodied-interaction design framework for mathematical concepts. In *Proceedings of the 10th International Conference on Interaction Design and Children*, pages 1–10. ACM, 2011.

[2] Mattias Arvola, Johan Blomkvist, Stefan Holmlid, and Giovanni Pezone. A service walkthrough in Astrid Lindgren's footsteps. In *ServDes. 2012: Service Design and Innovation Conference, Espoo, Finland, February 8-10, 2012*, pages 21–29. Linköping University Electronic Press, 2012.

[3] Ali Arya, Jesse Gerroir, Efetobore Mike-Ifeta, Andres Adolfo Navarro-Newball, and Edmund Prakash. Meseum: Personalized experience with narrative visualization for museum visitors. In *International Conference on Human-Computer Interaction*, pages 179–190. Springer, 2016.

[4] Ali Arya, Peggy Hartwick, Shawn Graham, and Nuket Nowlan. Collaborating through space and time in educational virtual environments: 3 case studies. *Journal of Interactive Technology and Pedagogy 2*, 2012.

[5] Marian Stewart Bartlett, Joseph C Hager, Paul Ekman, and Terrence J Sejnowski. Measuring facial expressions by computer image analysis. *Psychophysiology*, 36(02):253–263, 1999.

[6] Brian R Belland, Krista D Glazewski, and Jennifer C Richardson. A scaffolding framework to support the construction of evidence-based arguments among middle school students. *Educational Technology Research and Development*, 56(4):401–422, 2008.

[7] David Benyon, Kristina Höök, and Laurence Nigay. Spaces of interaction. In *Proceedings of the 2010 ACM-BCS Visions of Computer Science Conference*, page 2. British Computer Society, 2010.

[8] Randolph G Bias. The pluralistic usability walkthrough: Coordinated empathies. In *Usability inspection methods*, pages 63–76. John Wiley & Sons, Inc., 1994.

[9] Robert Biddle, James Noble, and Ewan Tempero. Essential use cases and responsibility in object-oriented development. In *Australian Computer Science Communications*, volume 24, pages 7–16. Australian Computer Society, Inc., 2002.

[10] Benjamin S Bloom. *Taxonomy of educational objectives: The classification of education goals. Cognitive Domain. Handbook 1.* Longman, 1956.

[11] Richard A. Bolt. "put-that-there": Voice and gesture at the graphics interface. *SIGGRAPH Comput. Graph.*, 14(3):262–270, July 1980.

[12] Elizabeth Bonsignore. Sharing stories in the wild: a mobile storytelling case study. In *CHI'11 Extended Abstracts on Human Factors in Computing Systems*, pages 917–922. ACM, 2011.

[13] Elizabeth Bonsignore, Alexander J Quinn, Allison Druin, and Benjamin B Bederson. Sharing stories "in the wild": A mobile storytelling case study using storykit. *ACM Transactions on Computer-Human Interaction (TOCHI)*, 20(3):18, 2013.

[14] Margaret M Bradley and Peter J Lang. Measuring emotion: the self-assessment manikin and the semantic differential. *Journal of behavior therapy and experimental psychiatry*, 25(1):49–59, 1994.

[15] John Seely Brown, Allan Collins, and Paul Duguid. Situated cognition and the culture of learning. *Educational researcher*, 18(1):32–42, 1989.

[16] Renate Nummela Caine and Geoffrey Caine. Making connections: Teaching and the human brain. 1991.

[17] Jonathan Chang, Karon MacLean, and Steve Yohanan. Gesture recognition in the haptic creature. In *Haptics: Generating and Perceiving Tangible Sensations*, pages 385–391. Springer, 2010.

[18] Po-Yao Chao. Exploring students' computational practice, design and performance of problem-solving through a visual programming environment. *Computers & Education*, 95:202–215, 2016.

[19] JC Coelho and FJ Verbeek. Pointing task evaluation of leap motion controller in 3d virtual environment in creating the difference. In *Proceedings of the ChiSparks 2014 Conference*, pages 78–85, 2014.

[20] Philip R Cohen, Michael Johnston, David McGee, Sharon Oviatt, Jay Pittman, Ira Smith, Liang Chen, and Josh Clow. Quickset: Multimodal interaction for distributed applications. In *Proceedings of the fifth ACM international conference on Multimedia*, pages 31–40. ACM, 1997.

[21] J Gordon Coleman Jr. All seriousness aside: The laughing-learning connection. *International Journal of Instructional Media*, 19(3):269–76, 1992.

[22] OpenURC Technical Committee. Openurc alliance, 2009. `http://www.openurc.org`.

[23] Douglas Connery, Brian Farrell, Dennis Pratt, and Sheri Johnson. Social approaches to learning, 2009. `http://www.slideshare.net/farrellb/social-approaches-to-learning-2410584`.

[24] Sunny Consolvo, Katherine Everitt, Ian Smith, and James A Landay. Design requirements for technologies that encourage physical activity. In *Proceedings of the SIGCHI conference on Human Factors in computing systems*, pages 457–466. ACM, 2006.

[25] Larry L Constantine and Lucy AD Lockwood. *Software for use: a practical guide to the models and methods of usage-centered design*. Pearson Education, 1999.

[26] Alan Cooper, Robert Reimann, and David Cronin. *About face 3: the essentials of interaction design*. John Wiley & Sons, 2007.

[27] Alan Cooper, Robert Reimann, David Cronin, and Christopher Noessel. *About Face: The essentials of interaction design*. John Wiley & Sons, 2014.

[28] Matthieu Courgeon, Jean-Claude Martin, and Christian Jacquemin. Marc: a multimodal affective and reactive character. In *Proceedings of the 1st Workshop on AFFective Interaction in Natural Environments*, page 20, 2008.

[29] Joëlle Coutaz, Laurence Nigay, Daniel Salber, Ann Blandford, Jon May, and Richard M. Young. Four easy pieces for assessing the usability of multimodal interaction: The care properties. In Knut Nordby, Per Helmersen, David J. Gilmore, and Svein A. Arnesen, editors, *Human—Computer Interaction: Interact '95*, pages 115–120, Boston, MA, 1995. Springer US.

[30] Monica Cowart. Embodied cognition. *The Internet Encyclopedia of Philosophy*, 2004.

[31] Mihalyi Csikszentmihalyi. Flow: The psychology of optimal experience. *Harper&Row, New York*, 1990.

[32] Fredy Cuenca. The cogenive concept revisited: A toolkit for prototyping multimodal systems. In *Proceedings of the 5th ACM SIGCHI symposium on Engineering interactive computing systems*, pages 159–162. ACM, 2013.

[33] Fredy Cuenca, Davy Vanacken, Karin Coninx, and Kris Luyten. Assessing the support provided by a toolkit for rapid prototyping of multimodal systems. In *Proceedings of the 5th ACM SIGCHI symposium on Engineering interactive computing systems*, pages 307–312. ACM, 2013.

[34] Wanda P Dann, Stephen Cooper, and Randy Pausch. *Learning to Program with Alice (w/CD ROM)*. Prentice Hall Press, 2011.

[35] Edirlei Soares de Lima, Bruno Feijó, Simone Barbosa, Fernando Ginez da Silva, Antonio L Furtado, Angelo EM Ciarlini, and Cesar T Pozzer. Multimodal, multi-user and adaptive interaction for interactive storytelling applications. In *Games and Digital Entertainment (SBGAMES), 2011 Brazilian Symposium on*, pages 206–214. IEEE, 2011.

[36] Rodrigo De Oliveira, Mauro Cherubini, and Nuria Oliver. Movipill: improving medication compliance for elders using a mobile persuasive social game. In *Proceedings of the 12th ACM international conference on Ubiquitous computing*, pages 251–260. ACM, 2010.

[37] Marco de Sá, Carlos Duarte, Luís Carriço, and Tiago Reis. Designing mobile multimodal applications. *Multimodality in Mobile Computing and Mobile Devices: Methods for Adaptable Usability*, pages 106–136, 2009.

[38] John Dewey. *Democracy and education*. Courier Corporation, 2004.

[39] C. Dion. *Laughter and humour in second-language learning and teaching*. PhD thesis, Université de Montréal, Quebec, Canada, 2007.

[40] Claire Dormann and Robert Biddle. Understanding game design for affective learning. In *Proceedings of the 2008 Conference on Future Play: Research, Play, Share*, pages 41–48. ACM, 2008.

[41] Claire Dormann and Robert Biddle. Humor and computer games: Play, laugh, and more—a review. *Simulation & gaming*, 2009.

[42] Paul Dourish. Embodied interaction: exploring the foundations of a new approach to HCI. *Transactions on Computer-Human Interaction on "HCI in the New Millennium*, 1999.

[43] Marcy Perkins Driscoll and Marcy P Driscoll. *Psychology of learning for instruction*. Pearson Boston, 2005.

[44] Bruno Dumas, Denis Lalanne, and Sharon Oviatt. Multimodal interfaces: A survey of principles, models and frameworks. In *Human machine interaction*, pages 3–26. Springer, 2009.

[45] Bruno Dumas, Beat Signer, and Denis Lalanne. Fusion in multimodal interactive systems: an hmm-based algorithm for user-induced adaptation. In *Proceedings of the 4th ACM SIGCHI symposium on Engineering interactive computing systems*, pages 15–24. ACM, 2012.

[46] Joost Dumeez, Maxime Bernaert, and Geert Poels. Development of software tool support for enterprise architecture in small and medium-sized enterprises. In *Advanced Information Systems Engineering Workshops*, pages 87–98. Springer, 2013.

[47] Sidney D'Mello, Tanner Jackson, Scotty Craig, Brent Morgan, Patrick Chipman, Holly White, Natalie Person, Barry Kort, Rana el Kaliouby, Rosalind Picard, et al. Autotutor detects and responds to learners affective and cognitive states. In *Workshop on Emotional and Cognitive Issues at the International Conference on Intelligent Tutoring Systems*, pages 306–308, 2008.

[48] Paul Ekman, Wallace V Friesen, and Phoebe Ellsworth. *Emotion in the human face: Guide-lines for research and an integration of findings: guidelines for research and an integration of findings*. Pergamon, 1972.

[49] B.J. Fogg. *Persuasive technology: using computers to change what we think and do*. Morgan Kaufmann, 2002.

[50] Brian J Fogg. Persuasive computers: perspectives and research directions. In *Proceedings of the SIGCHI conference on Human factors in computing systems*, pages 225–232. ACM Press/Addison-Wesley Publishing Co., 1998.

[51] Mary Ellen Foster. State of the art review: Multimodal fission. *COMIC project Deliverable*, 6(09), 2002.

[52] Claude Frasson and Pierre Chalfoun. Managing learner's affective states in intelligent tutoring systems. In *Advances in Intelligent Tutoring Systems*, pages 339–358. Springer, 2010.

[53] Ann Fry, Sonia Chiasson, and Anil Somayaji. Not sealed but delivered: The (un) usability of s/mime today. In *Annual Symposium on Information Assurance and Secure Knowledge Management (ASIA'12), Albany, NY*, 2012.

[54] Helene Gelderblom and Paula Kotzé. *Designing Technology for Young Children: What we can Learn from Theories of Cognitive Development*. ACM SAICSIT, 2008.

[55] Nahum Gershon and Ward Page. What storytelling can do for information visualization. *Communications of the ACM*, 44(8):31–37, 2001.

[56] Shlomo Goltz. A closer look at personas: What they are and how they work. *Smashing magazine*, 2014.

[57] Jan Gulliksen, Bengt Göransson, Inger Boivie, Jenny Persson, Stefan Blomkvist, and Asa Cajander. Key principles for user-centred systems design. In *Human-Centered Software Engineering—Integrating Usability in the Software Development Lifecycle*, pages 17–36. Springer, 2005.

[58] Jan Gulliksen and Bengt Göransson. User-centred systems design- a process with focus on users and usability, 2003. `http://www.it.uu.se/research/hci/acsd/englishsummary.html`.

[59] Saurabh Gupta, Robert P Bostrom, and Mark Huber. End-user training methods: what we know, need to know. *ACM SIGMIS Database*, 41(4):9–39, 2010.

[60] Ernest A Haggard and Kenneth S Isaacs. Micromomentary facial expressions as indicators of ego mechanisms in psychotherapy. In *Methods of research in psychotherapy*, pages 154–165. Springer, 1966.

[61] Anita J Harrow. *A taxonomy of the psychomotor domain*. D. McKay Company, 1972.

[62] Toby Hede and Andy Hede. Multimedia effects on learning: Design implications of an integrated model. In *Untangling the web: Establishing learning links. Proceedings ASET Conference*, pages 7–10, 2002.

[63] Kristina Höök. Knowing, communication and experiencing through body and emotion. *Learning Technologies, IEEE Transactions on*, 1(4):248–259, 2008.

[64] Mark Howison, Dragan Trninic, Daniel Reinholz, and Dor Abrahamson. The mathematical imagery trainer: from embodied interaction to conceptual learning. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*, pages 1989–1998. ACM, 2011.

[65] OpenRemote Inc. Openremote, 2009. `http://www.openremote.com`.

[66] Ivar Jacobson. *Object oriented software engineering: a use case driven approach*. Addison-Wesley Professional, 1992.

[67] Alejandro Jaimes and Nicu Sebe. Multimodal human–computer interaction: A survey. *Computer vision and image understanding*, 108(1):116–134, 2007.

[68] Carey Jewitt, Gunther Kress, and Diane E Mavers. *The Routledge handbook of multimodal analysis*. Routledge London, 2009.

[69] Timo Jokela, Netta Iivari, Juha Matero, and Minna Karukka. The standard of user-centered design and the standard definition of usability: Analyzing iso 13407 against iso 9241-11. In *Proceedings of the Latin American Conference on Human-computer Interaction*, CLIHC '03, pages 53–60, New York, NY, USA, 2003. ACM.

[70] Charles Juwah, Debra Macfarlane-Dick, Bob Matthew, David Nicol, David Ross, and Brenda Smith. *Enhancing student learning through effective formative feedback*. Higher Education Academy (Generic Centre), 2004.

[71] Martin Kaltenbrunner, Till Bovermann, Ross Bencina, and Enrico Costanza. Tuio: A protocol for table-top tangible user interfaces. In *Proc. of the The 6th Int'l Workshop on Gesture in Human-Computer Interaction and Simulation*, pages 1–5, 2005.

[72] Takeo Kanade, Jeffrey F Cohn, and Yingli Tian. Comprehensive database for facial expression analysis. In *Automatic Face and Gesture Recognition, 2000. Proceedings. Fourth IEEE International Conference on*, pages 46–53. IEEE, 2000.

[73] Ashish Kapoor and Rosalind W Picard. Multimodal affect recognition in learning environments. In *Proceedings of the 13th annual ACM international conference on Multimedia*, pages 677–682. ACM, 2005.

[74] Harald Karner and Georg Droschl. Usage-centered interface design for knowledge management software. *J. UCS*, 8(6):634–643, 2002.

[75] Jack Katz. *How emotions work*. University of Chicago Press, 2001.

[76] Sandra Kerka. Somatic/embodied learning and adult education. *Trends and Issues Alert*, 32, 2002.

[77] David Kirsh. Embodied cognition and the magical future of interaction design. *ACM Transactions on Computer-Human Interaction (TOCHI)*, 20(1):3, 2013.

[78] Kenneth R Koedinger and Albert Corbett. Cognitive tutors: Technology bringing learning sciences to the classroom. In R Keith Sawye, editor, *The Cambridge handbook of the learning sciences*, New York, 2006. Cambridge University Press.

[79] Barry Kort, Rob Reilly, and Rosalind W Picard. An affective model of interplay between emotions and learning: Reengineering educational pedagogy-building a learning companion. In *icalt*, page 0043. IEEE, 2001.

[80] David R Krathwohl, Benjamin S Bloom, and Bertram B Masia. *Taxonomy of educational objectives, Book 2 Affective Domain*. Longman, 1964.

[81] Gunther Kress. *Multimodality: A social semiotic approach to contemporary communication*. Routledge, 2009.

[82] Christine Kühnel. *Quantifying Quality Aspects of Multimodal Interactive Systems*. Springer Science & Business Media, 2012.

[83] James A Larson, TV Raman, Dave Raggett, Michael Bodell, Michael Johnston, Sunil Kumar, Stephen Potter, and Keith Waters. W3c multimodal interaction framework. *W3C NOTE*, 6, 2003.

[84] Jean Lave. Situating learning in communities of practice. *Perspectives on socially shared cognition*, 2:63–82, 1991.

[85] Jin-Hee Lee, Eikyu Choi, Minseok Song, and Byeong-Seok Shin. Dreamware: edutainment system for children with developmental disability. *Multimedia Tools and Applications*, 68(2):305–319, 2014.

[86] Yingzi Lin. Affective driving. In *Emotional Engineering*, pages 263–274. Springer, 2011.

[87] Yannick Loriot. Kinect: How to install and use openni on windows – part 1, 2011. `http://yannickloriot.com/2011/03/kinect-how-to-install-and-use-openni-on-windows-part-1/`.

[88] Jason Lui and Zhi Su. Situated cognition/learning theory, 2009. http://etec.ctlt.ubc.ca/510wiki/Situated_Cognition/Learning_Theory.

[89] Robert L Mack and Jakob Nielsen. *Usability inspection methods*. Wiley & Sons New York, NY, 1994.

[90] Thomas Mahatody, Mouldi Sagar, and Christophe Kolski. State of the art on the cognitive walkthrough method, its variants and evolutions. *Intl. Journal of Human–Computer Interaction*, 26(8):741–785, 2010.

[91] Jean-Claude Martin, Luc Julia, and Adam Cheyer. A theoretical framework for multimodal user studies. In *Proceedings of 2nd International Conference on Cooperative Multimodal Communication, Theory and Applications, Tilburg, The Netherlands*. Citeseer, 1998.

[92] Albert Mehrabian and James A Russell. *An approach to environmental psychology*. the MIT Press, 1974.

[93] Elaine Van Melle and Sheila Pinchin. Writing effective learning objectives, 2008. http://http://www.entcanada.org/Word_Files/CreatingLearningObjectives.pdf.

[94] Rachel Metz. Free software ties the internet of things together, 2013. https://www.technologyreview.com/s/519666/free-software-ties-the-internet-of-things-together/%7D/.

[95] Julien Mineraud, Oleksiy Mazhelis, Xiang Su, and Sasu Tarkoma. A gap analysis of internet-of-things platforms. *Computer Communications*, 89:5–16, 2016.

[96] Punya Mishra and Kathryn A Hershey. Etiquette and the design of educational technology. *Communications of the ACM*, 47(4):45–49, 2004.

[97] Selene Mota and Rosalind W Picard. Automated posture analysis for detecting learner's interest level. In *Computer Vision and Pattern Recognition Workshop, 2003. CVPRW'03. Conference on*, volume 5, pages 49–49. IEEE, 2003.

[98] Kasia Muldner, Michael Wixon, Dovan Rai, Winslow Burleson, Beverly Woolf, and Ivon Arroyo. Exploring the impact of a learning dashboard on student affect. In *Artificial Intelligence in Education*, pages 307–317. Springer, 2015.

[99] Lars Müller. Pervasive monitoring to support reflective learning. In *Proceedings of the 2013 ACM conference on Pervasive and ubiquitous computing adjunct publication*, pages 349–354. ACM, 2013.

[100] Americo Talarico Neto and Renata Pontin de Mattos Fortes. Improving multimodal interaction design with the mmwa authoring environment. In *Proceedings of the 28th*

ACM International Conference on Design of Communication, pages 151–158. ACM, 2010.

[101] Christine M Neuwirth, David S Kaufer, Ravinder Chandhok, and James H Morris. Issues in the design of computer support for co-authoring and commenting. In *Proceedings of the 1990 ACM conference on Computer-supported cooperative work*, pages 183–195. ACM, 1990.

[102] Jakob Nielsen. How to conduct a heuristic evaluation, 1995. `https://www.nngroup.com/articles/how-to-conduct-a-heuristic-evaluation/`.

[103] Jakob Nielsen. Severity ratings for usability problems. *Papers and Essays*, 54:1–2, 1995.

[104] Jakob Nielsen and Rolf Molich. Heuristic evaluation of user interfaces. In *Proceedings of the SIGCHI conference on Human factors in computing systems*, pages 249–256. ACM, 1990.

[105] Donald A Norman. *Emotion design: Why we love (or hate) everyday things*. Basic Books, 2004.

[106] Donald A Norman. *The design of everyday things: Revised and expanded edition*. Basic books, 2013.

[107] Hartmut Obendorf. *Minimalism: designing simplicity*. Springer Science & Business Media, 2009.

[108] Zeljko Obrenovic and Dusan Starcevic. Modeling multimodal human-computer interaction. *Computer*, 37(9):65–72, 2004.

[109] Sharon Oviatt. *Multimodal Interfaces*. Third Edition, CRC Press, Human computer interaction handbook: fundamentals, evolving technologies, and emerging applications, Jacko, Julie A, 2012.

[110] Sharon Oviatt, Phil Cohen, Lizhong Wu, Lisbeth Duncan, Bernhard Suhm, Josh Bers, Thomas Holzman, Terry Winograd, James Landay, Jim Larson, et al. Designing the user interface for multimodal speech and pen-based gesture applications: state-of-the-art systems and future research directions. *Human-computer interaction*, 15(4):263–322, 2000.

[111] Audun Mathias Øygard. Fitting faces: An explanation of clmtrackr, 2012. http://auduno.com/post/61888277175/fitting-faces, accessed on 2016-04-22.

[112] Malgorzata Pankowska. Individual and corporate users in it product development process. *International Journal of Digital Society (IJDS)*, 4(3), 2013.

[113] Jeff Patton. The new user story backlog is a map. *http://www. agileproductdesign. com/blog/the_new_backlog. html> Acesso em*, 3(10):2012, 2008.

[114] Rosalind W Picard. Affective computing for hci. In *HCI (1)*, pages 829–833, 1999.

[115] Rosalind W Picard, Seymour Papert, Walter Bender, Bruce Blumberg, Cynthia Breazeal, David Cavallo, Tod Machover, Mitchel Resnick, Deb Roy, and Carol Strohecker. Affective learning—a manifesto. *BT Technology Journal*, 22(4):253–269, 2004.

[116] Rosalind W Picard and Roalind Picard. *Affective computing*, volume 252. MIT press Cambridge, 1997.

[117] Jenny Preece, Helen Sharp, and Yvonne Rogers. *Interaction design: beyond human-computer interaction*. John Wiley & Sons, 2015.

[118] Sara Price and Carey Jewitt. A multimodal approach to examining'embodiment'in tangible learning environments. In *Proceedings of the 7th International Conference on Tangible, Embedded and Embodied Interaction*, pages 43–50. ACM, 2013.

[119] Leah M Reeves, Jennifer Lai, James A Larson, Sharon Oviatt, TS Balaji, Stéphanie Buisine, Penny Collings, Phil Cohen, Ben Kraal, Jean-Claude Martin, et al. Guidelines for multimodal user interface design. *Communications of the ACM*, 47(1):57–59, 2004.

[120] Andrea Resmini and Luca Rosati. A brief history of information architecture. *Journal of Information Architecture*, 3(2), 2012.

[121] Mitchel Resnick, John Maloney, Andrés Monroy-Hernández, Natalie Rusk, Evelyn Eastmond, Karen Brennan, Amon Millner, Eric Rosenbaum, Jay Silver, Brian Silverman, et al. Scratch: programming for all. *Communications of the ACM*, 52(11):60–67, 2009.

[122] Alicja Rieger. Energize your classroom with humor, 2014. `http://www.facultyfocus.com/articles/effective-teaching-strategies/energize-classroom-humor/`.

[123] Mary Beth Rosson and John M Carroll. Scenario-based design. *The Human Computer Interaction Handbook: Fundamentals, Evolving Technologies and Emerging Applications*, pages 1032–1050, 2002.

[124] Mary Beth Rosson and John M Carroll. Scenario based design. *Human-computer interaction. Boca Raton, FL*, pages 145–162, 2009.

[125] Cyril Rousseau, Yacine Bellik, Frédéric Vernier, and Didier Bazalgette. A framework for the intelligent multimodal presentation of information. *Signal Processing*, 86(12):3696–3713, 2006.

[126] Jonathan P Rowe, Lucy R Shores, Bradford W Mott, and James C Lester. A framework for narrative adaptation in interactive story-based learning environments. In *Proceedings of the Intelligent Narrative Technologies III Workshop*, page 14. ACM, 2010.

[127] Jonathan P Rowe, Lucy R Shores, Bradford W Mott, and James C Lester. Integrating learning, problem solving, and engagement in narrative-centered learning environments. *International Journal of Artificial Intelligence in Education*, 21(1-2):115–133, 2011.

[128] Natalie Ruiz, Fang Chen, and Sharon Oviatt. Multimodal input. *Multimodal Signal Processing: Theory and applications for human-computer interaction*, pages 231–255, 2009.

[129] James A Russell. A circumplex model of affect. *Journal of personality and social psychology*, 39(6):1161, 1980.

[130] James A Russell. Core affect and the psychological construction of emotion. *Psychological review*, 110(1):145, 2003.

[131] Shima Salehi, Jain Kim, Colin Meltzer, and Paulo Blikstein. Process pad: a low-cost multi-touch platform to facilitate multimodal documentation of complex learning. In *Proceedings of the Sixth International Conference on Tangible, Embedded and Embodied Interaction*, pages 257–262. ACM, 2012.

[132] Michael D Sankey. How to develop 15 multimodal design heuristics in 3 easy (not) lessons. *International journal of Pedagogies and Learning*, 3(2):60–73, 2007.

[133] Jason M Saragih, Simon Lucey, and Jeffrey F Cohn. Face alignment through subspace constrained mean-shifts. In *Computer Vision, 2009 IEEE 12th International Conference on*, pages 1034–1041. IEEE, 2009.

[134] Akio Sashima and Koichi Kurumatani. Mobile monitoring of physiological signals for remote healthcare services. *Demonstrator presented at Pervasive*, 9:249–252, 2009.

[135] Rejeev Sharma, Vladimir Pavlovic, Thomas S Huang, et al. Toward multimodal human-computer interface. *Proceedings of the IEEE*, 86(5):853–869, 1998.

[136] Maxine Sheets-Johnstone. Animation: The fundamental, essential, and properly descriptive concept. *Continental Philosophy Review*, 42(3):375–400, 2009.

[137] Valerie Shute and Umit Tokac. Formative feedback, teacher training workshop, 2011. `www.cala.fsu.edu/files/formative_feedback.pptx`.

[138] Valerie J Shute. Focus on formative feedback. *ETS Research Report Series*, 2007(1):i–47, 2007.

[139] Eric Spaulding and Haakon Faste. Design-driven narrative: using stories to prototype and build immersive design worlds. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*, pages 2843–2852. ACM, 2013.

[140] Walter Dan Stiehl, Cynthia Breazeal, Kuk-Hyun Han, Jeff Lieberman, Levi Lalla, Allan Maymin, Jonathan Salinas, Daniel Fuentes, Robert Toscano, Cheng Hau Tong, et al. The huggable: a therapeutic robotic companion for relational, affective touch. In *ACM SIGGRAPH 2006 emerging technologies*, page 15. ACM, 2006.

[141] Wai Wa Tang, Kenneth WK Lo, Alvin TS Chan, Stephen Chan, Hong Va Leong, and Grace Ngai. i* chameleon: a scalable and extensible framework for multimodal interaction. In *CHI'11 Extended Abstracts on Human Factors in Computing Systems*, pages 305–310. ACM, 2011.

[142] Lawrence A Tomei. *Taxonomy for the technology domain*. IGI Global, 2005.

[143] Matthew Turk. Multimodal human-computer interaction. In *Real-time vision for human-computer interaction*, pages 269–283. Springer, 2005.

[144] Matthew Turk. Multimodal interaction: A review. *Pattern Recognition Letters*, 36:189–195, 2014.

[145] Kai Tuuri, Antti Pirhonen, and Pasi Välkkynen. Bodily engagement in multimodal interaction: A basis for. *Multimodality in Mobile Computing and Mobile Devices: Methods for Adaptable Usability: Methods for Adaptable Usability*, page 137, 2009.

[146] Florian van de Camp and Rainer Stiefelhagen. Gluetk: A framework for multi-modal, multi-display human-machine-interaction. In *Proceedings of the 2013 international conference on Intelligent user interfaces*, pages 329–338. ACM, 2013.

[147] Jelle van Dijk, Remko van der Lugt, and Caroline Hummels. Beyond distributed representation: embodied cognition design supporting socio-sensorimotor couplings. In *Proceedings of the 8th International Conference on Tangible, Embedded and Embodied Interaction*, pages 181–188. ACM, 2014.

[148] Virginie van Wassenhove, Ken W Grant, and David Poeppel. Visual speech speeds up the neural processing of auditory speech. *Proceedings of the National Academy of Sciences of the United States of America*, 102(4):1181–1186, 2005.

[149] Jeroen Vandaele. Humor mechanisms in film comedy: Incongruity and superiority. *Poetics Today*, 23(2):221–249, 2002.

[150] Cathleen Wharton, John Rieman, Clayton Lewis, and Peter Polson. The cognitive walkthrough method: A practitioner's guide. In *Usability inspection methods*, pages 105–140. John Wiley & Sons, Inc., 1994.

[151] Wikipedia. Open natural interaction, 2010. https://en.wikipedia.org/wiki/OpenNI.

[152] Cameron Wilson. Hour of code—a record year for computer science. *ACM Inroads*, 6(1):22–22, 2015.

[153] Susan Wilson. Components of cognitive apprenticeship: Scaffolding, 2008. http://etec.ctlt.ubc.ca/510wiki/Components_of_Cognitive_Apprenticeship:_Scaffolding.

[154] Beverly Woolf, Winslow Burleson, Ivon Arroyo, Toby Dragon, David Cooper, and Rosalind Picard. Affect-aware tutors: recognizing and responding to student affect. *International Journal of Learning Technology*, 4(3-4):129–164, 2009.

[155] Steve Yohanan and Karon E MacLean. The role of affective touch in human-robot interaction: Human intent and expectations in touching the haptic creature. *International Journal of Social Robotics*, 4(2):163–180, 2012.

[156] Wenhai Zhang and Jiamei Lu. The practice of affective teaching: a view from brain science. *International Journal of Psychological Studies*, 1(1):35, 2009.

Appendices

Appendix A

Background on Multimodal Systems Summary Table

The tables below summarize the frameworks discussed in Chapter 2.

Sections	Framework	Missing points
A Brief History	Bolt, 1980	User can communicate with a large screen display via speech and pointing gestures, however, the system does not provide an affective feedback.
	Cohen et al., 1997	A pen/voice system running on an early tablet, but does not consider affective states in their design.
Motivation for Multimodal Systems	Oviatt et al., 2000	Given a survey on multimodal interface design, but have not consider emotional aspects in the multimodal architectures.
	Ruiz et al., 2009	About multimodal input systems and user's cognitive load, but no consideration of affect.
	Turk, 2014	A review on designing multimodal interaction and cognitive load, but have not considered affect in the design.

Table A.1: Missing points summary: history and motivation (Part 1).

Sections	Framework	Missing points
Principles and Conceptual Models of Multimodal Systems	Dumas et al., 2009	Very briefly talked about affective aspects of communication such as emotion and attention, and future work that can be done in education;no consideration of design and evaluation methodologies in affective education.
	Rousseau et al., 2006	Designed a conceptual framework, which is based on four elements: information to present, interaction components, interaction context and behavior, but not considered affective strategies such as emotion, and the educational domain;no consideration of design and evaluation methodologies.
	Martin et al., 1998	Survey on ten multimodal user studies for multimodal interfaces, but have not considered, user's affective issues, or the educational domain. In addition, they have not explained design and evaluation methodologies.
	Tuuri et al., 2010	Discussed cognitive strategies of the user in bodily engagement multimodal interaction, but briefly explained some affective strategies, such as emotion and attention; no consideration of design and evaluation methodologies.

Table A.2: Missing points summary: principles and conceptual models (Part 2).

Sections	Framework	Missing points
Technical Architectures and Frameworks	Cuenca et al., 2013	Authors explained a toolkit for rapid prototyping of multimodal systems consisting of a framework and a graphical editor. But they considered technical architecture only.
	Tang et al., 2011	Developed i*Chameleon making possible programmers to prototype and test new interactive devices or interaction modes. But they considered technical architecture only.
	Van de Camp and Stiefelhagen, 2013	Presented a framework supporting touch, pointing like modalities, speech, person tracking, head pose estimation and face identification for interaction, allowing the design of interfaces for a wide range of display sizes. Present a control interactive room application as a example. But they considered technical architecture only.
	Dumas et al., 2012	Developed a framework and presented a novel algorithm for the fusion of input modalities in multimodal interactive systems and managing input data. But they considered technical architecture only.
	Open Natural Interaction, 2010	OpenNI currently supports sensor modules that are 3D sensor, RGB camera, IR camera, and a microphone or an array of microphones, but not yet applied to emotional sensors such as facial gesture recognition.
	Kaltenbrunner et al., 2005	TUIO facilitates interaction between people at different locations (e.g. possibility of distributed musical performance), but does not consider emotion of users.
	OpenRemote Inc., 2009	It is an open source framework for IoT, which creates customized user interfaces to control different devices with monitoring their status, but it does not provide cognitive or affective feedback to the users.
	OpenURC Technical Committee, 2009	OpenURC promotes the Universal Remote Console (URC) and associated standards and its application in products, this way facilitating user interfaces that are simple and intuitive to use, but without affective feedback.

Table A.3: Missing points summary: technical architecture (Part 3).

Sections	Framework	Missing points
UI Design for Multimodal Systems	Reeves et al., 2004	Considered cognitive strategies, but not affective strategies in their multimodal UI design; no consideration of design and evaluation methodologies.
Multimodal Systems and Education	Price and Jewitt, 2013	Considered cognitive strategies in learning but not affective strategies; no consideration of design and evaluation methodologies.
	Hede and Hede, 2002	Considered cognitive aspects of learning but not affective ones; no consideration of design and evaluation methodologies.
	Lee et al., 2012	Considered some cognitive and affective aspects of learning for motivating children with disability, but not applying to all users; no consideration of design and evaluation methodologies.
Multimodal Systems and Affect	Picard, 1999	Describe the benefits of multimodal affective computing in HCI and user emotion and experience, but have not considered design and evaluation methodologies.
	Woolf et al., 2009	Build an affect-aware tutors that use artificial intelligence for recognizing and responding to student affect; no explicit consideration of design and evaluation methodologies.
	James and Sebe, 2007	Explained affective HCI and considered affective strategies in their designed multimodal interaction system; no consideration of design and evaluation methodologies.
	Chang, 2010	Expressed an improved sensor system design. They explained affective touch and emotion, and described their first-generation gesture recognition engine, embodied in a fully functional and physically animated robot creature; not considered cognitive strategies, design and evaluation methodologies.
	Stiehl et al., 2006	Considered affective aspects by designing a robotic companion being showing positive benefits of companion animal therapy; not considered user's cognitive issues, or design and evaluation methodologies.
	D'Mello et al., 2008	Considers learners' affective and cognitive states, and it embodied pedagogical agent synthesizes affective responses through animated facial expressions and speech; but not talking about design and evaluation methodologies.
	Courgeon et al., 2008	Considered affective interaction between user and virtual agent. Not considered user's cognitive issues, or design and evaluation methodologies.
	Picard et al., 2004	The role typically seen for multimodality, however, simply relates to sensors that can detect emotion. They have not considered design and evaluation strategies for emotional multimodal learning systems.

Table A.4: Missing points summary: UI design, education and affect (Part 4).

Appendix B

Ethics Documents

This appendix contains the following documents:

- Ethics clearance certificates (pages 220 and 221).
- Case study 1 instructions (page 223).
- Case study 2 instructions (page 225).

Ethics Clearance for Case Studies

Carleton University
Research Ethics Office
Research Ethics Board
511 Tory, 1125 Colonel By Drive
Ottawa, ON K1S 5B6 Canada
Tel: 613-520-2517, ethics@carleton.ca

Ethics Clearance Form – New Clearance

This is to certify that the Carleton University Research Ethics Board has examined the application for ethical clearance. The REB found the research project to meet appropriate ethical standards as outlined in the *Tri-Council Policy Statement: Ethical Conduct for Research Involving Human, 2nd edition*, and the *Carleton University Policies and Procedures for the Ethical Conduct of Research*.

Date of Clearance: July 21, 2015
Researcher: Reza GhasemAghaei (Student Research: Ph.D. Student)
Department: Computer Science
University: Carleton University
Research Supervisor (if applicable): Prof. Ali Arya and Prof. Robert Biddle
Project Number: 103021
Alternate File Number (if applicable):
Project Title: Evaluation of Affective Educational Software
Funder (if applicable): NSERC Discovery Grant

Clearance Expires: May 31, 2016

All researchers are governed by the following conditions:

Annual Status Report: You are required to submit an Annual Status Report to either renew clearance or close the file. Failure to submit the Annual Status Report will result in the immediate suspension of the project. Funded projects will have accounts suspended until the report is submitted and approved.

Changes to the project: Any changes to the project must be submitted to the Carleton University Research Ethics Board for approval. All changes must be approved prior to the continuance of the research.

Adverse events: Should a participant suffer adversely from their participation in the project you are required to report the matter to the Carleton University Research Ethics Board. You must submit a written record of the event and indicate what steps you have taken to resolve the situation.

Suspension or termination of clearance: Failure to conduct the research in accordance with the principles of the Tri-Council Policy Statement: Ethical Conduct for Research Involving Humans, 2nd edition and the Carleton University Policies and Procedures for the Ethical Conduct of Research may result in the suspension or termination of the research project.

Louise Heslop
Chair, Carleton University Research Ethics Board

Andy Adler
Vice-Chair, Carleton University Research Ethics Board

Ethics Clearance for Case Studies

Carleton University Research Ethics Board (CUREB)

Certificate of Ethics Clearance

Principal Investigator	Department	StudyNumber
Reza GhasemAghaei	**Computer Science**	**15-184**

Co-Investigators and other researchers:

Researcher	Study Role	Position
Robert Biddle	Faculty Supervisor	Faculty

Study Title: **User study on affective learning software**

Approval Date: **10/19/2015** Expiry Date: Approval Type: Final

Submitted Date	Study Component	Approval Date

Validity Term:

Comments:

Certification

The protocol describing the above-named project has been reviewed by Carleton University Research Ethics Board and the research procedures were found to be acceptable on ethical grounds for research involving human participants.

Shelley Brown

Chair, Carleton University Research Ethics Board (CUREB)

This Certificate of Clearance is valid for the above term provided there is no change in the research procedures.

Close Print

Instructions for Case Study 1

Sample System Instructions

Description:

Thank you once again for your willingness to participate in the usability study for my PhD thesis research. This usability study is about improving software for learning different ratios in mathematics and understanding of proportional equivalence e.g. 1/3 = 2/6. You can find ratios e.g. 1/2, 2/6, 3/9 etc.

You are going through this software finding different ratios. You are able to select ratio and move your both hands above a motion sensor device to indicate this correct ratio. Moving one hand higher will result in a bigger distance and ratio. During the study and when the correct ratio happens, there will be affective strategies including: happy music, playing praising audio, showing encouraging messages, and using different colors and icons during the exercise. To change the affective feedback you are able to customize the settings by clicking on the "Setting" button. To access to the settings you need to enter your password.

The purpose of the system:

Using a multimodal affective learning system with having learner customization might increase the encouragement in learning, and will help students develop grounded understanding of different ratios and proportional equivalence.

Test:

In this system you have to do two parts:

- **Part 1:** Finding default ratio and having default affective feedback.
- **Part 2:** Finding your selected ratio, selected height and selected affective feedback.

1

Part 1: Finding default ratio and having default affective feedback

When the system starts, first the leaner has to enter his/her name. By clicking the "OK" button, the page refreshes, and shows the page with affective strategies needed for that specific learner (reset the system to its starting condition).

The system welcomes you and you can see which hand you are interacting above the motion sensor device on the top left corner of the interface.

In the bottom left side of the interface there are the affective strategies specific for you including: background color, praising image, image, icon, message and audio.

On the right side of the interface there are "Max Height", "Select Ratio" and "Setting".

- **Max Height:** main screen height can be changed between 0 and 99.
- **Select Ratio:** ratio can be changed between ratio 1:1 up to ratio 1:9 or 9:1.
- **Setting:** for accessing to settings you can use your own password or the course password that the instructor provides. By entering the correct password you are able to edit the affective settings.

Part 1 has two tasks:

- **Task 1:** Entering your name and checking to see the left/right hand/s above the motion sensor device.
- **Task 2:** Find the ratio 1:3.

2

Task 1: Entering your name and checking the left/right hand above the motion sensor device

In this section you have to enter a name from the database e.g. "Sarah" (your name as given to you by the researcher), and put your hand above the Leap Motion device to interact with the system with Sarah's affective settings that the instructor has set up (Sarah is the name we are using for an imaginary user of the system).

Your name is: . . .

Your password is: . . .

Task 2: Find the ratio 1:3

In this section you have to find the ratio 1:3 by moving your hands above the Leap Motion device.

When the correct ratio happens, you will receive affective feedbacks including:

Your background color when ratio happens: Green
Your praising image be (Yes/No) when ratio happens: Yes
Your praising image file when ratio happens: well Done
Your praising icon when ratio happens: Love Math Icon
Modify your praising message when ratio happens: You doing GR
Your audio when ratio happens: Applause Sound
Music can be play since beginning: Pharrell Williams Song - Happy
Update

The system saves the ratio, and the number of times the user finds/loses the correct ratio.

3

Part 2 –Finding your selected ratio, selected height and selected affective feedback

This part has three tasks:

- Task 1: Changing Height
- Task 2: Changing Ratio
- Task 3: Using your password and changing Affective Settings

Task 1: Changing Height

In this section select the height of the main screen that interacting.

As a default it is "10". It can be between 0 to 99.

Task 2: Changing Ratio

In this section you can select and find different proportional equivalences 1:2, 1:3, 1:4 and 1:5. As a default the system has set to 1:3 ratio. You move your both hands above the motion sensor device to indicate the correct ratio (i.e. 1:3). Moving one hand higher will result in a bigger distance and ratio.

Click on the "Select Ratio" on the right side of the interface, and select the ratio 1:2, and find this ratio.

Change the "Volume" in the bottom left side of the interface to 50%.

Task 3: Using your password and changing Affective Settings

You select the "Setting" button in the right side of the screen, and enter your own password.

You are able to change the settings with your preference. You can change the following affective feedbacks when the correct ratio happens. You can keep some empty (if preferred).

Your background color when ratio happens: Red
Your praising image be (Yes/No) when ratio happens: No
Your praising image file when ratio happens: No Image
Your praising icon when ratio happens: No Icon
Modify your praising message when ratio happens:
Your audio when ratio happens: No Sound
Music can be play since beginning: No Music

4

Instructions for Case Study 1

Researcher's Record

Participant number:

Part 1: Finding default ratio and having default affective feedback	
Tasks	**Comments**
Task 1: Entering your name and checking the left/right hand above the motion sensor device	
Task 2: Find the ratio 1:3	
Part 2 –Finding your selected ratio, selected height and selected affective feedback	
Tasks	**Comments**
Task 1: Changing Height	
Task 2: Changing Ratio	
Task 3: Using your password and changing Affective Settings	

5

Instructions for Case Study 2

Sample System Instructions

Thank you once again for your willingness to participate in the usability study for my PhD thesis research. This study is about improving a software for education; a narrative visualizations of a museum experience. We have two proposed usability inspection methods called *MADE Heuristics* and *MADE Walkthrough* (MADE: Multimodal Affect for Design and Evaluation). You as experts are going to do either a *MADE Heuristics* or a *MADE Walkthrough* that is provided in Part 1 (page 3) and Part 2 (page 5). I will be the facilitator.

This usability study is divided into two main sections. A *construction section* where you will be constructing a visual narrative of a simulated museum visit and a *viewing section* where you will view previously constructed visual narratives.

Construction section: In this section the user visits several exhibits in a virtual museum and collect data from each virtual exhibit in the way of pictures and text. Then view various ways they can be visualized and a narrative of their experiences presented.

Viewing section: In this section the user views visualizations based upon a premade trip and compare how each visualization presents the information.

The purpose of the system: To evaluate the extent to which the application of creating a visual narrative of users experience in a museum using interactive technology can enhance an individual's museum visit, and to communicate and share its contents.

Sub-Test 1: Construction section

To begin please navigate to the following website in Google chrome:

http://mvp.soft.carleton.ca/MuseumExplorer/

Please press the "*Reset System*" link before you begin. This will reset the system to its starting conditions. You should see a popup saying that the system has been successfully reset and ready for testing. Then click the "Museum 1" link to proceed to the system proper.

Task 1: Museum Explorer

1. Add the Beaver exhibit to your itinerary.
2. Remove the Beaver exhibit from your itinerary.
3. Add the following exhibits to your itinerary:
    - Three Types of Rock
    - The Solar System
    - Meteorites
    - Cougar
    - Beaver

    - Bear
    - Thinhorn Sheep
    - Marine Mammals and Birds
    - Mollusks and Invertebrate
    - Tectonic Plates
4. Click on an exhibit in your itinerary.
5. Visit the exhibit by clicking the '*Visit Exhibit*' button.

Task 2: Exhibit Explorer

1. Click on an artefact hotspot.
2. Take a picture of the artefact.
3. Add a note to the picture.
4. Return to the museum by clicking the "*Back to Museum*" button.

Task 3: Gathering Data

1. Visit all the other unvisited exhibits on your itinerary.
2. Take at least 1 more pictures and add at least 1 more notes.
3. When you are done exploring all the exhibits click the "*Create Visualization*" button.

Task 4: Visualization

1. View your visit.
2. Select which exhibits you'd like to include in your visualization.
3. Select one of the four visualization styles (it will open in a new tab).
4. Explore the visualization to get a sense of it.
5. Go back and explore the other four visualization styles until you have explored them all.

Sub-Test 2: Viewing section

To begin please navigate to the following website in Google chrome:
http://mvp.soft.carleton.ca/MuseumExplorer/

Have the Observer setup the system for you by completing the following:

1. Click the "Museum 1" link to proceed to the system proper.
2. The itinerary should already be filled with visited exhibits. Click the 'Create Visualization' button.
3. Select all of the exhibits as ones that you'd like to include in your visualization.

Each part basically involves viewing a visualization and trying to identify usability issues.

Part 1: Categorical Visualization

Part 2: Sequential Visualization

Part 3: Dramatic Visualization

Part 4: Slideshow Visualization

Part 5: Traditional Method

Part 1 – MADE Heuristics

In this part we are using heuristic evaluation to do the evaluation.

To begin please navigate to the following website in firefox:

http://mvp.soft.carleton.ca/MuseumExplorer/

Please do a heuristic evaluation on Sub-Test 1 and Sub-Test 2.

The proposed MADE heuristics is in the following table:

Affective Design	Heuristics
Visceral	**Aesthetic and minimalist design:** Learning should not contain information which is irrelevant or rarely needed. Every extra unit of information in a learning competes with the relevant units of information and diminishes their learning ability. Including additional music or sounds if that is an essential component of the learning interaction. The use of video may be preferred for a lecture style presentation.
	Help and documentation: Even though it is better if the multimodal software can be used without documentation, it may be necessary to provide help and documentation. Any such information should be easy to search, focused on the student's task, list concrete steps to be carried out, and not be too large.
Behavioral	**Visibility of system status:** The multimodal software should always keep learners informed about what is going on, through appropriate feedback within reasonable time, so the learner knows how he is doing, and what to do to get better.
	Learner control and freedom: Provide the student with some control over the learning environment considering the emotional aspects, ensuring that the instructional strategy is made clear. Learners often choose system functions by mistake and will need a clearly marked "emergency exit" to leave the unwanted state without having to go through an extended dialogue. Support undo and redo. The learner has to have the freedom to try new things.

	Error prevention in learning: Even better than good error messages is a careful design and learning objective providing cognitive strategies that prevent a problem from occurring in the first place. Either eliminate error-prone conditions or check for them and present learners with a confirmation option that the multiple sensory modalities provide before they commit to the action.
	Recognition rather than recall: Minimize the learner's memory load by using the cognitive and affective strategies making objects, actions, and options visible. The learner should not have to remember information from one part of the dialogue to another. Instructions for use of the multimodal software should be visible or easily retrievable whenever appropriate.
	Flexibility and efficiency of use: Acceleration strategy—unseen by the novice learner—may often speed up the interaction for the expert learner such that the multimodal software can cater to both inexperienced and experienced learners. Allow learners to tailor frequent actions.
	Help users recognize, diagnose, and recover from errors: Teacher should monitor the students with the multiple sensory modalities and if students face errors, motivate and persuade students with the affective strategies. Error messages should be expressed with text or sound, precisely indicating the problem, and constructively suggest a solution, e.g. providing a video clip.
Reflective	**Match between multimodal software and the educational environment:** The multimodal software and the teacher should speak the leaner's language using emotional and social factors, with words, texts, images and concepts familiar to the learner, rather than system-oriented terms. Follow real-world conventions, making information appear in a natural and logical order.
	Consistency and standards: Consistent and intuitive mapping Users should not have to wonder whether different words, situations, or actions mean the same thing. Follow platform conventions.

Instructions for Case Study 2

Part 2 – MADE Walkthrough

In this part we are using affective walkthrough to do the evaluation.

To begin please navigate to the following website in firefox:

http://mvp.soft.carleton.ca/MuseumExplorer/

Please do an affective walkthrough on Sub-Test 1 and Sub-Test 2.

We proposed the MADE walkthrough (MADEW) by considering the learnability and the learning objective (affective and cognitive strategies that the teacher has come up with), based on Wharton's work [Wharton, 1994]:

1. Will the learner be able to use the multiple sensory modality correctly (learnability in the user interface) and taking into consideration the learning objective that the teacher has set up and the cognitive affective learning strategies of the teacher that might help the learner to achieve the right effect that the subtask needs (learnability in the education)?

E.g. what cognitive affective strategies might help the learner achieve the teacher's cognitive affective objectives? And does the learner understand that a subtask like specific hand gesture with correct distance is needed to reach the learner's goal? What might the learner experience emotionally?

2. Will the multiple sensory modality helpful and does the learner notice that the correct action is available?

E.g. is the multimodal system sound clear and visible?

3. Will the learner understand and associate the correct action by using the multiple sensory modality with the effect trying to be achieved? (Learnability of the user interface)

E.g. the multimodal system sound is clear and visible but the learner does not understand the text and sound, and will therefore not respond to it.

4. When the correct action is performed, will the learner know that they have done the right thing, receive affective strategies, and see that progress is being made toward solution of the task?

E.g. does the learner get haptic feedback, sound or written message when the correct action is performed?

5