Combining pointing and gestures:
Novel interaction concepts for large high-resolution displays

by

Anton Stasche

THESIS

submitted in partial fulfillment of the requirements for the degree of

MASTER OF SCIENCE

in

PSYCHOLOGY

University of Konstanz,

Faculty of Sciences,

Department of Psychology

1st assessor: Prof. Dr. S. Sonnentag (Department of Psychology)

2nd assessor: Prof. Dr. H. Reiterer (Department of Computer & Information Science)

Konstanz, July 2008

Abstract

Interaction with large high-resolution displays offers the embracing of alternative input modalities due to the users' increased freedom of movement. As opposed to the standard mouse-keyboard combination, pointing devices such as laserpointers that possess motion-sensing capabilities can provide the advantage of integrating direct pointing with the eliciting of discrete commands to the system via gesturing. In the Human-Computer Interaction (HCI) domain, gestural interfaces have been largely developed and investigated using computer-vision and glove-based approaches. However, most of these findings do not regard the type of gestures applicable for the handheld motion-sensing device employed in the current study. This Master thesis explores the theoretical background of gestures for the sake of establishing a coherent classification of device gestures in HCI. These are then placed firmly into a framework for gestural interaction by augmenting the work of Karam (2006). Results and suggestions from a series of semi-structured interviews inform the creation of two collections of gestures for the context of electronic mind mapping performed on the Powerwall, a large high-resolution display at the University of Konstanz. Conclusions from a formal evaluation of the overall usability and learnability of two final gesture sets suggest the promise of using gestures as an additional input channel for large display interaction.

Table of Contents

1 Introduction 4
1.1 Motivation 5
1.1.1 Large, high-resolution displays 5
1.1.2 Laserpointer interaction 6
1.1.3 Computer supported cooperative work (CSCW) 7
1.1.4 Collaborative work with large display groupware 8
1.2 Overview 9
2 Gestures as an input modality 10
2.1 What are gestures? 10
2.2 History of gesture research 12
2.3 Gestures in Human-Computer Interaction 13
2.3.1 Karam's framework for designing gestural interactions 16
2.3.2 Relevance of the current Master's thesis 20
3 Classification of gestures 21
3.1 Levels of Analysis 23
3.2 Gesture typologies 27
3.3 A dimensional approach for classification of semiotic properties 32
3.3.1 Extending Karam's framework 34
4 Obtaining gestures 36
4.1 Assumptions 36
4.2 Pre-study method 37
4.2.1 Semi-structured interviews 37
4.3 Results of the pre-study 41
4.4 Discussion and conclusions of the pre-study 43
5 Establishing gesture sets 45
5.1 The arbitrariness distinction 46
5.2 The iconicity dimension 47
5.3 Creating the final gesture set 50
5.4 Individual mental models 51
6 Evaluating gesture interaction 52
6.1 Hypotheses 53
6.2 Method 55
6.2.1 Usability Test 55
6.3 Results of user test 67
6.4 Discussion of the user test 75
6.5 Conclusion of the user test 80
7 General discussion 80
8 General conclusion and outlook 85
9 References 87
10 Appendices 95

1 Introduction

Research and examinations of the use of gestures as a communication channel by humans has a long history (Kendon, 2004). Elaborate accounts of the variety of gesticulations can be traced as far back as ancient Rome with Quintilian's treatise on rhetoric (2004, p.18/19). In Psychology, Wilhelm Wundt (1904) included his considerations on the role of gestures for the development of languages in the first volume of *Völkerpsycholgie;* furthermore, anthropological investigations have long dealt with the defining features of gestures in different cultural and social contexts (Efron, 1945), while the controversy on the evolution of language and the role of gestures (Hewes, 1973) has been a major topic in Linguistics. Together with the analysis of the interplay between gestures and speech, these themes represent the main building blocks of a large body of research findings on human gestures. Approaches on how to define and classify gestures originate in all of these fields. It is hence logical to assume that this extensive body of knowledge provides the basis or at least informs the more recent research endeavours in the computing sciences. The use of gestures as a means to interact with computers, and more generally machines, grew largely due to progress in the development of the enabling technologies. The literature clearly shows that this has led researchers and developers to mostly concentrate on technical aspects of the recognition systems (Karam & schraefel, 2005b), although some authors have reported on their attempts to align the involved gestures with existing theories (Kettebekov & Sharma, 2000; Quek et al. 2002; Wexelblat, 1998). Reports from the Human-Computer Interaction domain present a similar picture as few accounts on the nature of the gestures used in interactions can be found (Latoschik & Wachsmuth, 1998; Martell, 2005). A promising first step in reconciling system-focused research and investigations into human gestures *per se* is put forth by Maria Karam (2006) in presenting a framework for gestural interaction research that combines guidelines for the classification, design, and evaluation of the defining parts of gesture interaction.

1.1 Motivation

With the current Master thesis it is my ambition to contribute to this reconciliation process by incorporating into my line of research the results of a thorough review of the various theories from Psychology, Anthropology, Linguistics and Psychophysiology about the nature of human gestures and their defining features. By augmenting and extending the framework presented by M. Karam, the groundwork for a methodological sound investigation of gestures as an input modality is laid.

Interacting with computers and machines by using gestures is marked by the specific types of gesture that are possible depending on which features of the users' gesturing are recognized by the system. Besides the gesture type, the context of use can influence the degree to which a specific means of eliciting gestures by the user is useful and practicable. The major motivating aspects for the current research spring from attributes (ISO 9241-11, 1998) of an intended context of use of gesture input that are presented as follows.

1.1.1 Large, high-resolution displays

The increased presence of large displays in academic, business, and public contexts calls for the readdressing of classic interaction concepts such as WIMP (windows, icons, mouse, pointing) in terms of the users' means for control, the affordances of the display and the context.

Large, high-resolution displays (LHRD) such as the 5x2m Powerwall at the University of Konstanz with a resolution of over 7 megapixels are capable of visualizing large and complex amounts of data. For typical exploration and analysis settings, a user or a group of users can observe detailed information from close-range or get an overview of the presented data from a distal position. Due to the display capabilities exceeding either the limited human visual acuity or the user's field of view depending on the distance to the screen, it is not possible to perceive both detail and overview simultaneously (König, Bieg, Schmidt & Reiterer, 2007). Hence, it presents both an opportunity and a challenge to address the users' freedom of movement when designing interaction techniques and input devices.

1.1.2 Laserpointer interaction

As opposed to the well-known and widespread mouse-keyboard combination, a more flexible alternative input device such as a laserpointer could harness this freedom of movement, enabling the user to control of what is visualized from any position in front of the display. The system developed in project inteHRDis of the HCI group (http://hci.uni-konstanz.de/intehrdis) at the University of Konstanz presents such an alternative by providing fast and accurate tracking of the infrared beam emitted by a custom-built laserpointer, which allows the direct mapping of pointing movements onto a cursor on the large screen. As a natural extension to the hand, this input device enables the user to perceive the immediate effect of his/her movements on the screen, which draws on the familiarity of laserpointers as used for presentation purposes (König, Bieg & Reiterer, 2007). Hence, its potential for immediate use without elaborate training is supporting the device's application in contexts of collaboration such as meetings and seminars that do not offer longer periods of practice through daily routine.

Depending on the on-screen interface used in the interaction, direct pointing can however be laborious if longer sequences of selecting and clicking are required. The possibility of reducing the interaction with graphical elements (widgets) would thus mitigate the drawbacks of the laserpointer's direct mode of control. Independently of the target application, a method to achieve this purpose could be to elicit discrete commands similarly to employing keyboard shortcuts instead of multiple selections of e.g. menu items with the mouse. Realizing the means to issue such commands could thus provide a more fluent control process, since the direct manipulation of the screen content must not be interrupted by distracting searches for menu commands or on-screen widgets. Although such discrete commands could potentially be realized with any input modality open to the user, employing an unobstructive method that can incorporate familiar representations of these commands, is desirable. One such method could be the use of gestures since they provide an easy-to-perform way to elicit any number of discrete symbols that can represent desired commands. Executing the movements for such gestures by using a laserpointer equipped with motion-sensing capabilities would thus combine the advantages of both direct control and discrete commands into a single input device. Given that an adequate method for text entry can be provided when required, the switching between devices such as in the case of the mouse/keyboard combination would be rendered obsolete.

1.1.3 Computer supported cooperative work (CSCW)

The potential of using computer systems to support collaboration prompted advocates of computer-supported cooperative work (CSCW) to establish guidelines for evaluating the crucial aspects of the interaction between users and the system (Neale, 2004). In the context of the Powerwall display, the interaction between users can be considered largely *co-located* (two or more persons in the same room) as well as *synchronous* (work is conducted at the same time) which leads to the question of how a specific form of input can support electronic tools realized on a shared display (Hilliges et al., 2007). Major issues in this concern are the input nature, and what difference simultaneous or alternating (turn-taking) control makes, as well as the impact on the communication, creative flow, and social dynamics (Streitz et al., 1999). For instance, Birnholtz, Grossman, Mak, Balakrishnan (2007) report that groups of users adapted their approach on performing a layouting task on a shared display depending on the number of input devices (single-mouse vs. one mouse for each user). In particular, user coordinated the tasks differently: more parallel work was performed when multiple mice were used, while a team-based method was adopted in the other case by leaving the mouse to a single user and giving verbal commands, as this was easier than continuously passing the device. However, as the authors point out, at times this was frustrating to those not controlling the mouse and enabled the "mouse-wielding" user to dominate the task. Providing users with a more flexible and mobile input device than the mouse could overcome the limitations in turn-taking scenarios by facilitating the passing of the device between users, and hence reducing the risk of a single user dominating the task. At the same time, issuing each user with an input device like a laserpointer could contribute to the advantage of parallel work on group efficiency. In fact, Vogt et al. (2004) found that user groups excelled in single-display collaboration that involved discussion and information-sharing, when using laserpointers.

1.1.4 Collaborative work with large display groupware

Apart from the challenges regarding the design of on-screen interfaces and applications that take advantage of the potential for collaboration and communication of large displays in academic and industry contexts, some of the hallmarks of large display groupware (LDG) imply the need for a suitable input method that reflects the particular aspects of the respective usage scenarios. Offering characteristics such as simultaneous visibility for multiple users as well as shared workspace, large displays represent a suitable vehicle for CSCW tools (Huang, 2006, PhD), that would benefit from an input modality like hand gestures that reduces the distractions of technology to leave the group's focus on the task. One of the potential scenarios of use involves the collaborative activity of brainstorming for e.g. problem-solving purposes, making use of the mind mapping technique (Buzan & Buzan, 1996). Enabling users to contribute equally to the process of creating and adjusting mind maps realized through an electronic mind mapping tool on a large display, and at the same time not impeding the coordination and communication efforts by all users hence requires a means of input that is marked by its mobility and inconspicuousness. In addition, because of potentially heterogeneous and ever-changing group configurations, tapping the benefits of shared mind mapping activities for meetings requires an input device that is straightforward to use and does not entail lengthy periods of habituation.

1.2 Overview

The current Master thesis provides first indications of how large high-resolution display interaction can benefit from gestures as an input modality in addition to direct pointing, by combining both in a single input device. After giving an account of the origins of gesture research across different disciplines, the historical development of using gestures as an input modality for human-machine interaction is summarized. Next, an overview of gesture classification attempts reported in the literature is presented, and the procedure and results of a line of interviews aimed at obtaining suitable gestures for computer input is discussed. It is assessed how prior classification schemes can be applied on the proposed gestures or which potential new classification structures can be drawn from them. In the next step, the creation of gesture sets using the proposed gestures is summarized, and classification features that could serve to differentiate potential gesture sets are established. The setup and procedure of a series of user tests stretching across two sessions are then reported, focusing on the overall usability and learnability of two distinct gesture sets. Following this, the results from the classification review, the interviews and the user tests are analyzed to assess how they can inform the assumption that one type of gesture is more suited for computer input than other(s). Finally, implications for the use of gestures as an input modality in human-computer interaction and more specifically for large display interaction are illustrated, and issues for further research are proposed.

2 Gestures as an input modality

Before being able to determine the nature of human gestures that qualify as a means of input for human-computer interaction, it is essential to arrive at a framework of definitions that not only delineates the different kinds of gesturing (Which human behaviours are called gesturing?) and the respective types of gestures (Along which features can types of gestures be distinguished?), but that also prescribes a fixed terminology and describes how prior schemes fit into a universal hierarchy.

2.1 What are gestures?

Given the diversity of focus and the vast spread of discussions of human gestural behaviour across scientific disciplines, it is not a simple task to straightforwardly define "gestures" without taking into consideration the context in which the term is used. Wilhelm Wundt (1904) names three classes of movements of expression: automatic, drive and voluntary ("Ausdrucksbewegungen […], in die drei Klassen der automatischen, der Trieb- und der Willkürbewegungen unterschieden werden"). In his discussion of voluntary movements he further considers only expressive movements made with the hands, that can form signs functioning like words independently of speech (Kendon, 2004, p.91).

Adam Kendon in his seminal work "Gesture: Visible Action as Utterance" (2004), initially gives a more general definition by describing human gesturing as "movement, whose communicative intent is paramount, manifest and openly acknowledged". It is clear from this description that any bodily movement can basically be interpreted as a gestural act, provided that as "deliberate expressiveness" (2004, p.15), it is apparent and its intentional character is directly perceivable. As the author puts it, "An action that is gestural has an immediate appearance of gesturalness".

For the sake of delineating the gestural movements performed in HCI contexts, the following prior assumptions are made:

- the movements are inherently deliberate, intentional and expressive (their purpose is to signal to the machine)
- the performing body part(s) are required as initial differentiators
- the importance of the relation to speech needs to be specified

A definition of gesture should hence be made in relation to these points whenever a discussion or investigation is initiated. Although the above characteristics are close to a first classification, this is not their purpose and the aim is solely to encourage an early denomination of the gestural movements under scrutiny. On the basis of Kendon's definition, gestural acts could for example be head and eye gestures, facial gestures, body motion, empty-hand gestures and object-aided hand gestures. Akin to the ground signaling performed for guiding planes by using rods, the gestural movements executed in the current context by using a special laserpointer can be called hand gestures, or more specifically, device gestures, as it is this object as an extension of the hand that is employed for the gestural act. As for the relation to speech it suffices to emphasize the independence of the gestural movements to whatever verbal communication is taking place simultaneously. This may however be different for cases where speech is of importance for the interaction, as in conversational interfaces with gestures as a complementary channel of communication (e.g. Wexelblat, 1995).

2.2 History of gesture research

In the introduction to "Gesture and Thought" (2005), the author David McNeill as one of the most active contemporary gesture researchers, considers the topics of recent investigations as being the result of moving away from a classical focus on gestures as a tool in oration dating back to first-century Rome. Kendon (2004) also mentions Aristotle in this context as having been averse to these aspects of oratorical skill (2004, p.17) and likewise cites the work of the Roman rhetorician Quintilianus as providing one of the early most complete accounts of *gestus*, which besides hand uses includes aspects such as body posture, head and face actions, and glance (2004, p.18). According to McNeill (2005), a first shift from this tradition is manifest in the anthropological investigations of David Efron (1941) who studied gestures in daily life and as part of conversation and other forms of communication against various cultural backgrounds. The second change in research focus is being linked by the author to the work of Adam Kendon. Although Wilhelm Wundt (1904, 1912) had discussed gesture in the early twentieth century, McNeill (2005) considers research on gestures in psychology and linguistics not as intensified until the early 1970s, when gestures were considered as "integral parts of the processes of language and its use" (2005, p.13).

In his extensive summary of the progress of gesture study beginning with this period, Kendon (2004) outlines three major themes that led to this revival of interest. The first was the revisiting of the idea that language originated as a form of gesture which, partly through the surprising success of teaching a form of American Sign Language (ASL) to a chimpanzee (Gardner & Gardner, 1969), led to a second theme of renewed interest, the analysis of sign language systems. The third focus had its origin in psychologists' and linguists' interest in the mental processes involved in language, with the development of the field of cognitive science encouraging the increase in investigations into the relation of language and thought (Kendon, 2004, p.73).

One of the recurring problems during these periods of heightened research efforts into human gesturing is the establishing of coherent and consistent classification schemes of the diverse gesture types. Various approaches have been proposed and although Kendon concludes that no one classification can serve as a general instrument for distinguishing gestures in any given context (2004, p.84), some of the more prominent and influencing schemes will be presented in the next chapter. In the following paragraph, an overview of the major themes of gesture research in the Human-Computer Interaction realm is presented.

2.3 Gestures in Human-Computer Interaction

The main interaction with computers has until recently almost always been realized through a mouse used for navigation and pointing in combination with keyboard input. However, research e.g. in the field of ubiquitous computing has challenged the classic notions of how humans can interact with computers, digital devices, and interfaces of various kinds. Acknowledging how computers become more and more present in everyday life, a promising way of rendering the interaction more accessible may be to let people use gestures to interact with the devices. Potential scenarios include the controlling of home entertainment systems, personal digital assistants, mobile phones, and public displays. Some of the latest web browsers also allow gesture interaction (Mozilla, 2007; Opera, 2007). In a recent development by videogame console manufacturer Nintendo, the wireless controller (WiiMote) of the Wii console (Nintendo, 2007) can be used as a pointing device in addition to performing gestures in three dimensions that are recorded by motion and tilt sensors.

In general, research endeavours in the computing sciences regarding the use of gestures for interaction purposes, have been guided to a large degree by progress in the underlying technologies. Examples can be found in the prototype developments of Bolt's

"Put-that-there" combination (1980), the GestureWrist and GesturePad (Rekimoto, 2001) and the improvements of recognition algorithms for pen-based (Rubine, 1991) and accelerometer-based gestures (Mäntylä, Mäntyjärvi, Seppänen & Tuulari, 2001). Nonetheless, some researchers have based the different gesture styles employed in the various studies into a framework of categories derived from the findings from other disciplines. These attempts will be included in the summary of classification endeavours in the next chapter. For the sake of providing an overview of the developments of gesture interaction as follows, a rough differentiation that follows the "high level breakdown" of Karam & schraefel (2005b) will be applied according to the technologies that provide the means to incorporate human gestures as an input modality.

Karam & schraefel (2005b) refer to *perceptual input* when users can perform the gestures to communicate with the system without requiring any physical contact to intermediary input devices or objects (2005b, p.12). Kettebekov and Sharma (2000) present one such system, enabling combinations of speech and gesture information to be extracted from TV weather narrations to provide indications for the suitability of computer vision for recognizing continuous gesturing in conjunction with speech. In a different vein, Smith and colleagues (1998) argue for electric field sensing to be used as noncontact gesture technology, which transmits a field through the human body to any number of receiving sensors. Placing receivers on screens for instance has been found to enable the tracking of finger movements to be used as an input alternative (Allport, Zimmerman, Paradiso, Smith & Gershenfeld, 1995).

In contrast to these approaches, *non-perceptual input* requires devices or objects to convey the gesture information (spatial and if required, temporal characteristics) to the machine. This can be realized as basic as using a mouse or pen to perform strokes which form discrete gestures (Cao & Zhai, 2007; Wobbrock, Wilson & Li, 2007; Rubine, 1991). In a similar fashion, that however entails a much higher flexibility, gestures that are realized through touch based input have been shown to offer great potential for interaction with mobile interfaces (Pirhonen, Brewster & Holguin, 2002) and table top computing (Wu, Shen, Ryall, Forlines & Balakrishnan, 2006). Developments in wearable computing proved already more than 25 years ago (Bolt, 1980) how simple gestures recognized by tracking the position of the users' arm can serve as input to a system. Wearable devices have been used to track head gestures (Brewster , Lumsden, Bell, Hall & Tasker, 2003) and finger bend and motion (Tsukada & Yasumura, 2002). Although the latter example bears some characteristics of glove-based approaches, the authors underline that their "Ubi-Finger"-device enjoys the benefits of increased comfort with the possible gestures being sufficient for the proposed context (2002, p.2). Glove-based gesture input has been discussed since the 1980s (Lingrand et al., 2006; Zimmermann, Lanier, Blanchard, Bryson & Harvill, 1986) and allows including information from single fingers, the whole hand and the wrist. A related but differently realized system using the tracking of finger markers attached to a glove has been reported to permit the recognition of finger gestures in addition to direct pointing (Foehrenbach, König, Gerken & Reiterer, 2008).

In the area of tangible interfaces, the manipulation of objects which is recognized by the system has included investigation into gestures (Fitzmaurice, Ishii & Buxton, 1995) but although many recent developments have furnished objects and devices with sensors that are capable to transmit gesture-related information to a computer, the manipulation of these devices is not afforded in the strict sense. Rather, these objects are intended as intermediate equipment that is utilized to transmit information about the gestures. The main differentiating characteristics are hence that they are not tangible and not wearable, but required to be held constantly in one or two hands. Work on gesture interaction employing such handheld devices have been reported for laser- or infrared-emitting appliances (Chen & Davis, 2002; Wilson & Shafer, 2003), passively tracked devices (Cao & Balakrishnan, 2003), portable devices such as smartphones (Ballagas, Borchers, Rohs & Sheridan, 2006) and motion-sensing devices such as some recent mobile phones (Apple, 2007; Nokia, 2006; Patel, Pierce & Abowd, 2004), game controllers (Nintendo, 2007) and custom-built prototypes. The latter developments are most akin to the laserpointer device designed at the HCI group of the University of Konstanz (König, Böttger, Völzow & Reiterer, 2008) and include the XWand (Wilson & Shafer, 2003), the BlueWand (Belgardt, Schwan & Reichardt, 2005) and the mCube (Kwon, Würmlin & Gross, 2007). Furthermore, there exist a reasonable number of publications reporting work that makes use of the SoapBox system (Tuulari & Ylisaukko-oja, 2002; Kela et al., 2006; Mäntiyjärvi, Kela, Korpipää & Kallio, 2004) and Nintendo's WiiMote controller (Kratz, Smith & Lee, 2007; Schlömer, Poppinga, Henze & Boll, 2008; Sreedharan, Zurita & Plimmer, 2007) to realize gesture input. The common theme among them is the particular nature of the gestures that can be performed by using any of these motion-sensing gadgets. In general that is, the gestures are not static as the device needs to be actively deployed. As in the laserpointer system used in the current study, information about the gesture movement is sensed by an accelerometer of either 3-axis (Belgardt, Schwan & Reichardt, 2005; Kwon, Würmlin & Gross, 2007) or 2-axis with additional sensors (Tuulari & Ylisaukko-oja, 2002; Wilson & Shafer, 2003). As sensors of this kind become ever smaller in size and price, gesture input by performing dynamic movements with motion-detecting handheld devices such as laserpointers, remote controls, presentation aids, and smartphones has the potential to be established as a viable alternative or supplement for the control of computers and digital artifacts.

2.3.1 Karam's framework for designing gestural interactions

In reviewing past work on gesture input, it can be observed that due to the diverse approaches adopted by the researches from the computing sciences and HCI, the different aspects of each system are hard to compare. Surely, a preliminary distinction on the basis of the underlying technology can serve to deliver an overview of the developments. However, if one intends to focus on the critical aspects of the interaction that is made possible and above all, the general suitability of the gesture modality for computer input, the characteristics of more than just the sensors, algorithms and devices need to be defined. In her doctoral dissertation, Maria Karam (2006) describes a first step in the process of unifying concepts, definitions and nomenclatures presented in the vast range of work on gesture input and interaction. The author introduces her aim of establishing a "theoretical framework for understanding and designing gesture-based interactions and the methods used in its development", for the purpose of guiding research and design of gestures and providing "a structure for understanding gesture systems and their interrelated concepts" (2006, p.80). Such a framework does however not imply the exclusion of existing ones. It is hence important to note that in general, extensions can be made and single constituencies be augmented. In this way, a framework for "research and design of gesture-based human-computer interactions" has the potential to evolve and remain up-to-date as conditions in gesture interaction change due to technological progress and theoretical insight. The main building blocks of Karam's framework will be introduced as follows. The author divides these into categories, subcategories and parameters (2006, pp.84-86). Subsequently, an extension to one of the subcategories ("Gestures") is proposed. The first category in Karam's framework (see Figure 1) concerns the Application domain (2006, pp.87-89). It contains parameters that aid in specifying the physical and cognitive requirements of the interaction context (represented in the respective subcategory), the criticality and complexity of the tasks at hand, and the existence of conflicts between the users' goals and any other parameters.

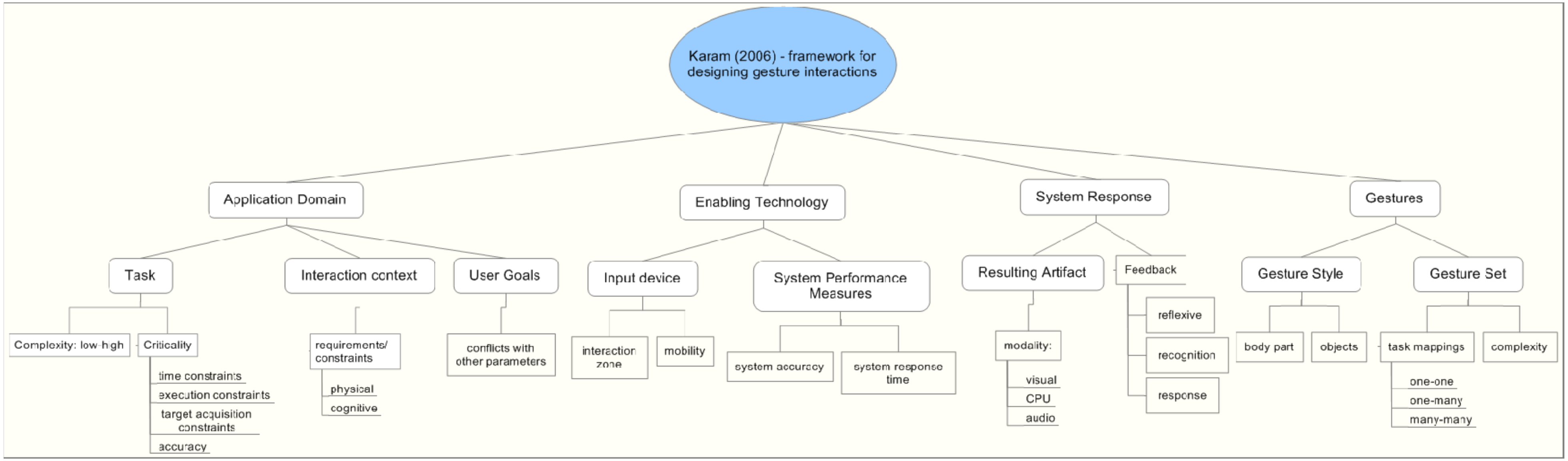


Figure 1. Structure of the original framework for "research and design of gesture-based human-computer interactions (adapted from Karam, 2006).

The second category deals with the technology employed for gesture input (2006, pp.90-92). The subcategory of input devices gives advice on how to specify the characteristics of the input device, that is, its interaction zone (whether contact with the device is required), and its mobility. Furthermore, aspects of system performance are introduced in a second subcategory, describing how the accuracy and response time of the gesture recognition impact on the design. Next, aspects of the gestural acts as such, independent of the technology, are subject of the category of Gestures in the framework (2006, pp.93/94).

Although the author refers to her adopted gesture styles elsewhere in her work, the Gestrure styles subcategory in the framework does not contain a parameter for the elements of this classification, but incorporates solely advice on how to specify the body parts and potential intermediary objects involved in the execution of the gestures. The alteration of this part of the framework to account for the major findings on gesture classification from multiple disciplines, will be discussed at the end of the next chapter. The second subcategory ("Gesture set") gives parameters for aspects of the gesture set such as the task mapping of the included gestures (does each gesture correspond to a single command), their overall number and physical complexity. Finally, the System Response category (2006, pp.95/96) outlines how the "outcome of a gesture interaction" can be specified in terms of the output system modality as well as through different stages of feedback.

2.3.2 Relevance of the current Master's thesis

The current study concerns the application of gesture input for large screen interactions. It involves the use of a custom-built laserpointer equipped with motion-detecting sensors in addition to an infrared laser beam that is tracked by a set of cameras behind the display. As such, reports on a combination of the interaction techniques of direct manipulation and gestures have been rare (Wilson & Shafer, 2003), and the particular domain of the WiiMote game controller (Nintendo, 2007) renders the rate of scientific publications involving this latter device rather insignificant. Furthermore, investigations about the nature of gestures that can be performed with these types of handheld motion-sensing devices are virtually non-existent, as most reports contain solely accounts on the underlying technology or the domain of application (Belgardt, Schwan & Reichardt, 2005; Kela et al., 2006; Kwon, Würmlin & Gross, 2007; Kratz, Smith & Lee, 2007; Schlömer et al., 2008; Sreedharan, Zurita & Plimmer, 2007; Wilson & Shafer, 2003). It is hence the aim of this Master thesis to provide guidelines of how to classify these types (or "styles" as Karam names them) of gestures, to enable researchers to better analyze the non-physical and non-technological features of the gestures employed for the interaction. These guidelines are then incorporated into Karam's (2006) framework for gestural-based interactions. In line with this framework, an empirical investigation has been conducted in the domain of large display interaction within the context of use of an electronic mind mapping session. The results of this investigation is presented in the latter part of the current thesis, and provide a first indication of the suitability of gesture interaction using a handheld motion-sensing device and the respective gestures.

3 Classification of gestures

Many frameworks and schemes for the classification of gestures have been presented in the literature, mainly outside the context of human-computer interaction. Due to their origins, the distinguishing properties of the proposed gesture types are derived from a diverse range of discipline-specific observations, as in the cultural differences recorded by Efron (1941) and the closeness to language analyzed by McNeill (2005). In general, which properties are under focus and employed for differentiation purposes is related to the aim of the researchers' investigation, or, as Kendon puts it, "the particular objectives of the inquiry” (Kendon, 2004, p.84). It is hence of major importance to first establish a set of goals whenever such an inquiry into gestures for HCI purposes is undertaken. Adopting the categories of the framework for gesture-based interactions described above, is one possibility to set an emphasis in investigations. For most of the existing studies, an alignment of their topics of interest according to the frameworks’ main building blocks is very well possible. Aspects such as system performance, interaction context, and employed input devices have been described in great detail in the respective publications. For some of the studies however, Karam’s (2006) framework lacks the subcategories (which further distinguish the various aspects of domain, technologies, gestures, and system response) and the related parameters (which define aspects of the subcategories). For instance, it would be difficult to apply the framework for Kettebekov and Sharma’s (2000) semantic classification of gestures used for manipulation of objects in a 2D environment, Kwo, Wuermlin and Gross’ (2007) glass metaphor and hand-held gestures for appliances, Payne et al.’s (2006) spatial symbolic gestures for videogame commands, Tsukada and Yasumura’s (2002) finger gestures for home appliances control and Kela et al.’s (2006) gathering of feasible gestures for an universal “vocabulary” for specific applications. After reading the results from the user interviews and experiments conducted for the current thesis, it will become clear that the category of “Gestures” and, more specifically, the subcategory of “Gesture Styles” is in need of elaboration. The result would allow accounting for all details of investigations as mentioned above, concerning the functional and physical nature of the involved gestures. Although Karam briefly mentions other classification attempts for gestures in her literature review (2006, p.12) and adopts a revised classification framework based on the work by Quek et al. (2002) to summarize the different gesture types (“styles”) in the existing literature dealing with gesture interaction (Karam, 2006, pp.13-17), she decides not to include elements of her “high-level classification” in the final framework to inform design and research. Furthermore, no clear statements are made about how to specify the approach to any classification endeavour including the actual properties of the gestures that are emphasized (Kendon, 2004, p.84). Acknowledging Kendon's claim that gestures "cannot be pinned down in any fixed way" (ibid.) but that any "particular classification systems developed are useful working instruments for a given investigation" (2004, p.85), does not however render the striving for classification guidelines futile. It rather makes every structured attempt at creating such guidelines for a given interaction context worthwhile, because such working instruments are still rare among the published studies in the domain of HCI. The following paragraphs contain both general advice for classification approaches, and the description of a potential working instrument for the classification of gestures that can be performed using handheld motion-sensing devices for large-screen intetion.

3.1 Levels of Analysis

An initial step that is crucial before attempting any classification of human gestures is the specification on which level of analysis this process is performed. This even precedes the aforementioned need to specify the dimensions, that is, the ties of the gestures that are under focus in any investigation. Among the existing typologies from the various disciplines, few mention explicitly the respective level on which the analysis of gestures has taken place. A hierarchy of levels will be proposed following the introduction of classification attempts in the literature.

Perspectives according to Efron
Spatio-temporal (movement characterstics)
Inter-locutional (interactional functions)

Box 1. Efron's perspectives of description (adapted from Kendon, 2004)

McNeill (2005), introduces four dimensions to be employed for distinguishing gesture types (see Box 2). David Efron (as cited by Kendon, 2004) deals with gestures of the hands and arms in conversation from three perspectives (summarized in Box 1). Ekman and Friesen (1969) consider three main principles to be crucial for understanding "nonverbal behavior" (1969, p.49). Box 4 outlines these aspects.

Dimensions according to McNeill (2005)
Relationship to speech
Relationship to linguistics properties
Relationship to conventions

Box 2. McNeill's dimensions (adapted from McNeill, 2005).

Karam, as has been mentioned above, accepts a high-level classification of the various gesture types (Karam, 2006, p.12) that builds on the distinctions put forth by Quek et al. (2002) based on the instrumental functions of gestures;
Likewise, Cosnier (1977) considers three different functions of gestures associated with the hand (1977, pp.35/36). Box 3 details their implications.

Cadoz (1994) introduces three methods to conduct a descriptive analysis of a set of gestures (a "corpus"). These are summarized in Box 5. In the realm of HCI, researchers have approached the classification problem mainly from two perspectives. The first concerns the physical aspects of the gestures and is manifested in coding schemes as summarized by Martell (2005) and in quantitative models e.g. for pen strokes (Cao & Zhai, 2007). The second perspective focuses on the functional role of the gestures in the interaction. As Quek et al. (2002) mention, most work concerned

Complementary functions of hand gestures according to Cosnier (1977)
Ergotique (modification and transformation of the environment; physical action)
Épistémique (acquisition of knowledge about the environment)

Box 3. Functions according Cosnier (1977). Translations by the author.

gestures entailing *manipulative* or *semaphoric* functions (with *deictic* as a special intermediate case). Manipulation is defined by the authors as involving "a tight relationship between the actual movements of the gesturing hand/arm with the entity being manipulated" (2002, p.172). An example can be found in the glove-based control of 3D models presented by Lindgrand and colleagues (2006).

Fundamental considerations according to Ekman and Friesen (1969)

Origin (learned or inborn)

Usage (external conditions, relation to verbalizations, performed consciously or not, communicative intent, observer feedback)

Coding (relationship between gesture act and its meaning)

Box 4. Ekman and Friesen's (1969) main principles

Semaphoric gestures are considered communicative by Quek et al. because they involve a set of "static or dynamic hand or arm gestures" (2002, p.172) to be communicated to the machine. Examples include the stroke episodes distinguished for recognition by Rubine (1991) and distinct movements that represent commands presented by Kela et al. (2006).

It is this last class of gesture to which the gestures presented in the current study are most akin to. However, keeping the classification endeavour on the level of distinction presented by Quek et al. (2002) would lead to not very fruitful insights concerning potential further distinctive features of gestures in HCI. For example, the question about the exact difference between various instances of semaphoric gestures would remain largely unanswered let alone shine any light on the nature or the connection between the actual gesture and its referent, that is, its meaning.

Methods according to Cadoz (1977)

Micro analytique (based on the structural linguistic model of "double articulation")

Macroanalytique (based on ethological methods to analyse habitual usage)

Box 5. Groups of methods for the analysis of a gesture corpus (Cadoz, 1994). Translations by the author.

This is what McNeill calls "character of semiosis" (2005, p.10). Semiosis, as defined in Merriam-Webster (2008), refers to any process involving signs. Due to the character of semaphoric gestures (from Greek *–phoros:* carrying) in HCI and in the current context in particular, the following discussion of classification schemes will contain mainly accounts of semiotic analysis (from Greek *sema:* marks, signs) of gestures and how the meaning of these "signs" is determined. Prior to presenting an overview of the relevant classification schemes, one last important argument about the levels of analysis is presented that should allow an initial definition of semaphoric gestures.

Building on the work of Louis Hjelmslev, Jean-Luc Nespoulous, and Andre Roche Lecours (1986) introduce four levels from which approaches to the semiotic analysis of gestures can be derived:

- **substance of expression**
  on this level, a gesture is described as the motion segment(s) that constitute it, independently of its signification (meaning) or function.
- **form of expression**
  those elements of the gesture that serve to distinguish and contrast against other movements are the focus on this level, requiring that they are shared by the people using these gestures in spite of variations between different persons; this is the "signifier" (Saussure).
- **form of content**
  on this level, the "signified" is described, as it constitutes the other side of the coin when analyzing a gesture form. The focus is now on the semiotic value, that is, the degree of being a "sign" for a community of people.
- **substance of content**
  this level concentrates on the signification that is conveyed by using gestures in order to learn about, as Nespoulous and Lecours (1986) call it, "the subjectivity of the individual".

The complementary character of these levels of analysis is underlined by Nespoulous and Lecours (1986) since an individual gesture can be scrutinized on multiple levels. They caution however to pay attention to the specific level when making an observation. Likewise, any change of level of analysis must be indicated. Moreover, the functional value of a gesture in a given communicative act may be determined inside a further framework, focussing more on to the context of the gestural act than on its nature. It is with additional regard to these functional attributes that some of the more prominent semiotic classifications of gestures have been proposed. Three of these existing typologies most relevant for the current research are presented hereafter. Presented first is the approach that is related to the abovementioned levels of analysis which is then followed by the proposed typologies of Ekman and Friesen (1969), Jaques Cosnier (1977) and Claude Cadoz (1994).

3.2 Gesture typologies

In what follows their discussion of the different levels of semiotic analysis for identifying gestural acts, Nespoulous and Lecours (1986) propose the distinction of gestures according to the degree of arbitrariness in the respective system of signs of the community of gesture users. It describes the connection between the “form of expression”, that is, the movement that constitutes a distinct gesture, and the “form of content”, the sign that is shared by the respective community.

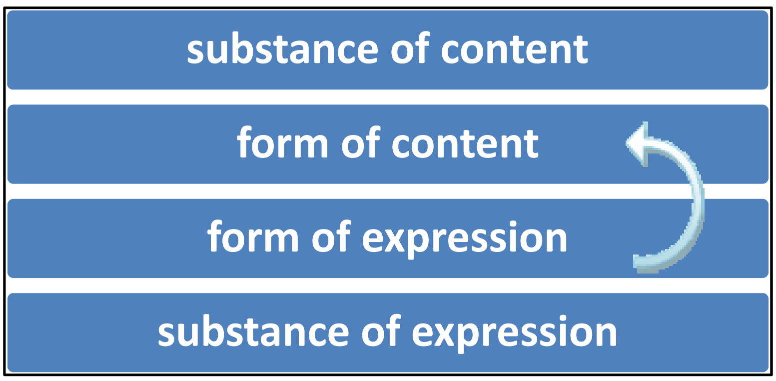


Figure 2. Arbitrariness of the link between form of expression and content

The authors distinguish three types of gestures according to this distinction: arbitrary, mimetic and deictic gestures (1986, p.56). An arbitrary connection between the gesture movement and its content exists e.g. in the case of the gesture for "money" when rubbing the thumb against index finger. Despite being conventionalized in certain cultural communities, this connection needs to be learned in order to be understood.

In contrast, a mimetic (Greek μίμησις: "imitation") gesture, such as the outlining of a metropolis' skyline by drawing it in the air or the mimicking of an action such as breaking something using the hands, is characterized by its transparent relationship (1986, p.56) with the referent. Nespoulous and Lecours (1986) further differentiate this mimetic class into "connotative gestures" as when sketching some partial feature of the referent in the air, and "strictly mimetic gestures" as in the case of mimicking some phase of an action (1986, p.57).

Nespoulous and Lecours (1986) propose a third type, deictic gestures, emphasizing the connection between the pointing gesture and the object in the environment. In particular the degree of arbitrariness of this connection is of importance. An example would be the difference between the pointing towards an object or location while referring to the object/location itself, and the pointing to an object or location which represents something abstract. The authors point out that without the information from the situation (context), no reference to the content can be established (1986, p.58).

Based on the work of Efron (1941) who studied the use of gestures in everyday discourse, Ekman and Friesen (1969) put forth a related distinction, and defined the "code which describes how meaning is contained in a non-verbal act" as "the rule which characterized the relationship between the act itself and that which it signifies" (1969, p.60). According to the authors, this coding may be arbitrary or iconic. Arbitrarily coded acts (the authors never use the term "gesture"), as has been mentioned above, do not relate to their signification through their appearance. In contrast, iconic coding visually links the act with its significant. The important difference to the distinction on the basis of the respective levels of analysis mentioned earlier, is that the actual function of the gesture content in the communicative act is implicitly taken into account. That is, while Nespoulous and Lecours (1986) do not yet turn to the signification of the "money" gesture (e.g. the financial situation of a nation's population), Ekman and Friesen (1969) integrate this final meaning inside a given context into the analysis of the relationship between the form and the content of the gesture. It is this latter approach that was picked up by Nehaniv and colleagues (Nehaniv et al., 2005) in their discussion on gesture types for incorporating gesture understanding in human-robot interaction:

> Note that the degree of arbitrariness in such gestures may vary: The form of the gesture may be an arbitrary conventional sign (such as a holding up two fingers with palm forwards to mean peace, or the use of semaphores for alphabetic letters). On the other hand, a symbolic gesture may resemble to a lesser or greater extent iconically or, in ritualized form, a referent or activity. (2005, p.3)

Iconic coding is further specified by Ekman and Friesen (1969) through an outline of the different types of "visual relationships" that can exist between the nonverbal act and its significant. Box 6 gives an account of these relationships. The authors then distinguish five categories of nonverbal behavior which are among the most widely cited in the literature (Kendon, 2004, p.96):

- Emblems
- Illustrators
- Affect displays
- Regulators
- Adaptors

Visual relationships according to Ekman & Friesen (1969)

Pictorial (drawing a picture of an event, object or person)

Spatial (distance between people, objects or ideas)

Rythmic (indicating flow of idea, accents of a phrase or rate of an activity)

Kinetic (executing an action performance which signifies or is itself the meaning)

Box 6. Type of visual relationships of iconic acts (Ekman & Friesen, 1969)

It is the first category that is closest to the kind of gestures considered in the current study and the class of semaphoric gestures described by Quek et al. (2002) as widely used for computer input. Ekman and Friesen (1969) characterize Emblems as having a direct translation because they refer to a shared definition within a language group. Moreover, their use is "usually an intentional, deliberate effort to communicate" (1969, p.63), a characteristic that is echoed in the use of this kind of gestures in the context of interaction with computers or machines (input entails the intentional communication of commands to the machine and this function remains fixed between situations and persons; see also definition of Semaphoric gestures by Quek et al., 2002).

Unlike for their category of Illustrators ("which are directly tied to speech, serving to illustrate what is being said", 1969, p.68), Ekman and Friesen do not explicitly differentiate further types of Emblems. Nonetheless, the authors acknowledge that "Illustrators can include the use of an emblem" (ibid.), which then is included under the respective label. The following types of Illustrators have been proposed by the authors:

- Batons
- Ideographs
- Deictic movements
- Spatial movements
- Kinetographs
- Pictographs

As can be observed, the scheme of visual relationships as listed in Box 6 is on the base of this distinction. In order to isolate distinguishing features of gesture acts, the adoption of a similar scheme may serve as a first step for the creation of guidelines on the semiotic analysis of gestures in HCI contexts. In fact, the final distinction that led to the creation of two gesture sets to be evaluated in a usability test setting for the current study, was partly based on this scheme.

In his discussion of the ”categories fonctionnelles de la mimogestualité”, Jaques Cosnier (1977) reports his attempt at creating an extensive typological framework for gesture classification. The main categories are the following (translations from Nespoulous and Lecours, 1986):

- Quasilinguistic gestures
- Expressive gestures
- Regulatory gestures
- Phatic gestures
- Metacommunicative gestures
- Coverbal gestures
- Extracommunicative gestures

In this framework, only the Quasilinguistic and Coverbal categories may be of concern in the current context. In fact, Cosnier explicitly mentions his understanding of Quasilinguistic gestures as “quasi linguistic naturelle” (1977, p.2038), as opposed to the made up “dialectes” for particular professions and usage contexts. Nevertheless, the authors’ further differentiated types of “signes quasi linguistiques” are likewise possible to be on the base of such purposefully created gesture dialects. Cosnier proposes the following types (translations by the author):

- Deictics
- Iconics
- Connotatives
- Arbitrary gestures

Similarly to Ekman and Friesen’s (1969) list of Illustrators, Cosnier (1977) then describes the following types of illustrative gestures inside his Coverbal category:

- Deictic
- Spatiographic
- Kinemimic
- Pictomimic

As can be observed, the difference in types between these two categories is limited, and most distinctions among types can be accounted for either by analyzing the degree of arbitrariness (see above Nespoulous & Lecours, 1986; Ekman & Friesen, 1969), or the type of visual relationship (see Box 6). In fact, Cosnier mentions that "Regarding their form and their nature, the Illustratives […] have multiple things in common with the Quasilinguistics mentioned above, one could consider them "Coverbal Quasilinguistics"." (Cosnier, 1977, translation by the author).

Adam Kendon distinguishes the following kinds of gestures (Kendon, 1988a cited by McNeill, 2005):

- Gesticulation
- Speech-linked gestures
- Emblems
- Pantomime
- Signs

McNeill (2005) then arranged these along different continua with each based on a particular dimension (see Box 2). Note that the latter author speaks of gesture "kinds" (2005, p.5) when contrasting his research emphasis, gesticulations, from the other categories. Upon introducing his approach to a classification of gestures to account for their diversity in semiotic properties, he talks of gesture "types" (2005, p.38). In both examinations, McNeill considers "dimensions" as being his preferred solution to a simple categorization. The dimensions he mentions for aligning the different kinds of gestures have been presented in Box 2, his proposed "dimensional framework" (ibid.) is explored in the following paragraph.

3.3 A dimensional approach for classification of semiotic properties

Similar to the frameworks introduced above, McNeill (2005) differentiates gesture types in four categories:

- Iconic (gestures convey images of concrete entities or actions)
- Metaphoric (gestures with abstract use of form or space)
- Deictic (gestures locating entities and actions in space)
- Beats (temporal highlighting of discourse aspects)

These form the basis of the author's subsequent step in overcoming the common notion of categories in most classification schemes. He proposes using dimensions such as *iconicity, metaphoricity, deixis, temporal highlighting,* and *social interactivity*:

> The essential clue that these semiotic properties are dimensional and not categorical is that we often find iconicity, deixis, and other features mixing in the same gesture. Falling under multiple headings is not impossible in a system of categories, but simultaneous categories implies a hierarchical arrangement. We cannot define such a hierarchy because we cannot say in general which categories are dominant and which are subordinate. [...] Because a multiplicity of semiotic dimensions is an almost universal occurrence in gesture, it makes sense to shift from categories to dimensions (2005, pp.41/42).

A dimensional approach of this kind could inform the analysis of gestures in HCI context by allowing the scoring of each element in a set of gestures according to what McNeill calls their "loading" on each dimension (2005, p.42). Although the author arrived at these conclusions inside the context of his focus on gesticulations that occur naturally during speech, he does not deny that the dimensions mentioned above could likewise be at the base of a classification scheme for emblems. Moreover, his proposed dimensions do not present a comprehensive set and, depending on the aim of the analysis and the context, the inclusion of further dimensions could potentially increase the instrumentality of the respective framework. This is maybe the most important insight from McNeill's work.

The discussion of the different schemes and approaches to the classification problem in this chapter contains some of the most representative and widely cited proposals in the literature from different disciplines. Although it does not present a complete survey of all frameworks introduced by researchers in these domains, many of those not mentioned are redundant in their proposed categories. It must be noted that for a thorough analysis of the physical aspects of gestures, further instruments in the form of coding schemes and models have been proposed which are not covered in the mainly semiotic discussions of the abovementioned frameworks.

3.3.1 Extending Karam's framework

Transferring the theoretical findings explored in the preceding paragraphs to a framework for the design and research of gesture-based human-computer interactions requires embedding them into the guidelines that could inform such endeavours. As outlined in the description of Karam's (2006) framework above, its category of "Gestures" lacks advice on any form of physical or semiotic examination. It is hence the first major contribution of this thesis to augment a number of subcategories and parameters and thus extend the prescriptive and descriptive power of the framework. Although all currents in development and research in gesture interaction cannot be comprehensively covered by this updated framework, it should give a first idea of how it could serve to increase the agreement of studies concerning definitions of gestures and methods of analysis. The respective section of the first update of the framework can be seen in Figure 3.

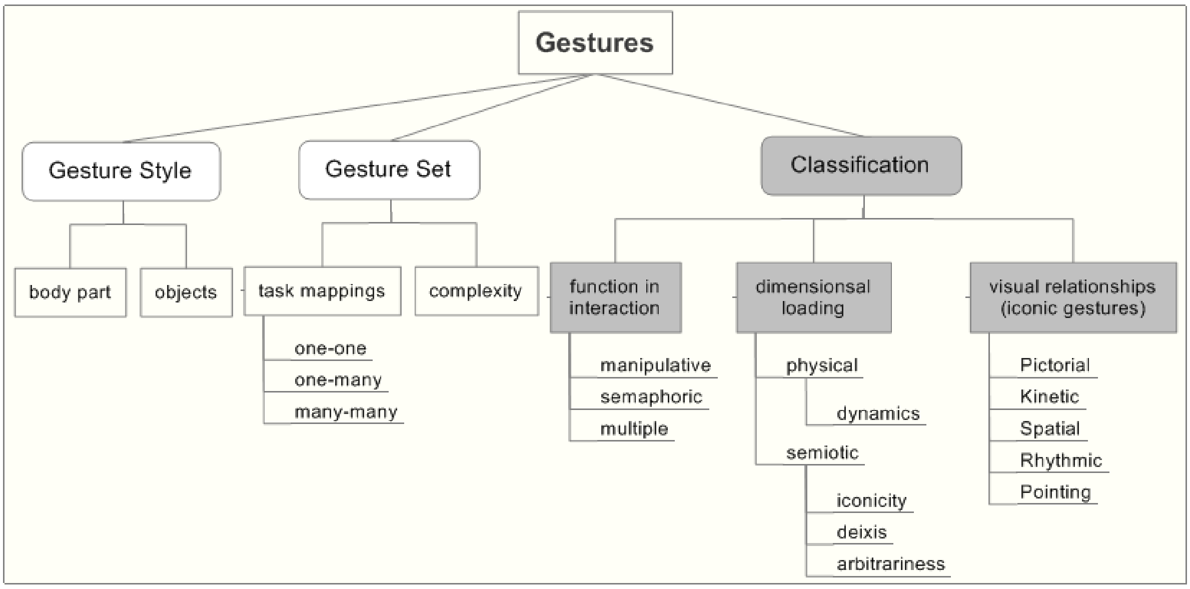


Figure 3. Extended *Gestures* category of Karam's (2006) framework.

Applying the elements of this category of the framework to the kind of gestures used in the current interaction context would initially yield the following results:

- Gesture style
  - Body part: single hand
  - Objects: yes
- Classification
  - Function: semaphoric

The following chapter summarizes the procedure and results of a series of user interviews performed to collect gesture proposals from users for the scenario of a mind mapping session. The characteristics of the gesture set as well as the individual gestures are then determined. In order to harness the power of the framework's guidelines to inform potential differences in usability and learnability between different groups of gestures, the proposed gestures were divided along two types of visual relationships and compared in a formal usability test.

4 Obtaining gestures

The goal of the pre-study was to collect feasible gestures for the given application, a mind mapping tool developed by Mindjet (2007), and the suitability of using gestures for the interaction in general. Moreover, an eventual creation of a gesture vocabulary from the collected instances was pursued.

4.1 Assumptions

As an initial investigation, the assumptions of this pre-study were kept at a general level, as no systematic accounts of semiotic or physical analyses were found in the literature, that are applicable to the kind of gesture possible with the handheld motion-sensing device used for the study. Following Kela et al.'s (2006) observation of questionnaire results concerning gestures with a similar device (SoapBox), the basic hypothesis rested on the difference in spatial dimensions that participants may use to perform their ideas of feasible gestures. It was assumed that gestures which remained on a 2D x-y plane (see Kela et al., 2006, p.288), would be more numerous overall and exhibit a higher variation among users than gestures performed in a 2D y-z plane or all three dimensions (see Figure 4 for a sketch of movement dimensions).

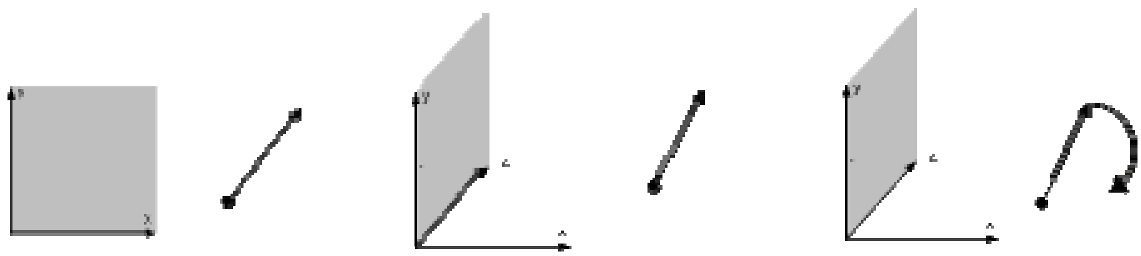


Figure 4. Spatial planes of gesture movements possible with handheld motion-sensing device (notation adapted from Kela et al., 2006). From left to right: 2D x-y, 2D y-z, 3D.

4.2 Pre-study method

4.2.1 Semi-structured interviews

In the technical report on "Usability methods supporting human-centered design" (ISO/TR 16982:2002(E)), section 6.3.4 recommends methods involving direct user feedback when early information in the development process is required. Particularly user interviews are suggested, although no specific information is given on the specific form of interview in question. Nonetheless, section 5.1.6 mentions potential forms ranging from "highly structured to very open-ended" (2002, p.9) and describes advantages over questionnaires such as their flexibility concerning follow-up questions. For the current series of user interviews, a semi-structured approach was adopted that provided for a fixed number of questions concerning experience with input devices and software applications, and their respective functions. In addition, the interviewer could rate the users' appraisal of the target application and its presented functions on a fixed scale. Subsequent to the introductory queries, users could perform the gesture that they found to be most suited for the respective command (which would execute the intended function of the application) and either take down a sketch themselves or supervise the sketching by the interviewer.

4.2.1.1 Participants

The user interviews were conducted during a three week period involving six volunteer participants between 16 and 29 years of age and one of over 60. No systematic differences due to age could be observed in either the responses concerning computer experience or the understanding of the mind mapping concept and application. The nationalities of the interviewed were Italian (4) and German (3). All participants were familiar with standard office applications and mouse/keyboard input, five were regular users of web browsers and two reported knowledge of image processing or statistics/ mathematics software. Beside the mouse-keyboard combination, six had used a touchpad as found in recent laptops, and two participants had at least once employed a pointing device such as the WiiMote or a laserpointer before.

4.2.1.2 Apparatus and Materials

The interviews were conducted in front of a standard LCD 19″ screen with an English version of the mind mapping application being controlled by the interviewer. Participants were handed the accelerometer-equipped model of the laserpointer presented by König, Bieg & Reiterer (2007) which was at the time not yet operational in gesture mode, but was nevertheless employed to allow users to come up with feasible gestures while grasping the device and pressing the designated gesture button (see Figure 5).

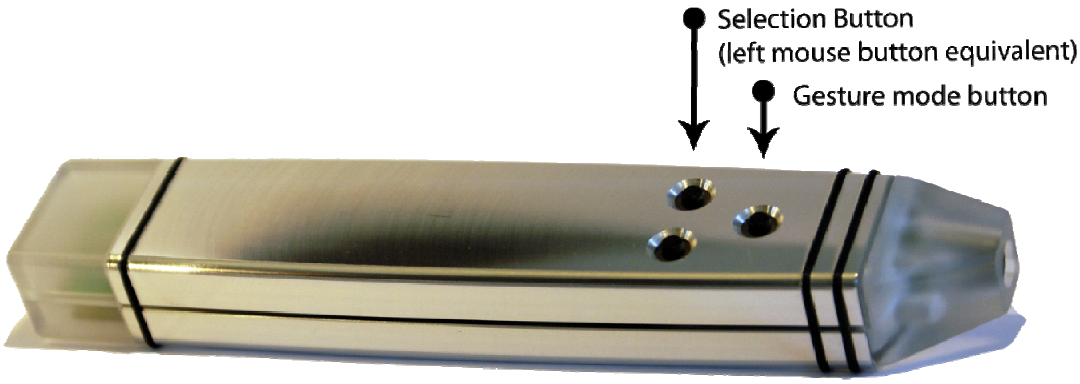


Figure 5. Infrared laserpointer equipped with inertial sensor (from König et al., 2008).

The interview materials were composed by the author both for the Italian and German sessions (the Italian documents were corrected by a native speaker to ensure the correct phrasing of questions, scales and technical jargon). They included guidelines for the interview procedure, a fixed set of questions about the appraisal of the mind mapping software and its functionality, with rating scales to be filled in by the interviewer according to the overall response by the participant. Moreover, each of the functions of the mind mapping application for which a gesture idea needed to be found, was listed on separate sheets with enough space for sketching these gestures while they were performed by the participants. Althoug the actual mind mapping software was in English in both the German and Italian interviews, a program version in each respective language was consulted beforehand, to include the intended translations of concepts and functions in the interview guidelines.

4.2.1.3 Functions

The following basic functions of the MindManager application (MindJet, 2007) were introduced to the participants:

- Navigation/visualization
    - zoom in
    - zoom out
    - split map view vertically
    - remove view split
- Editing
    - add sibling topic
    - add subtopic
    - add relationship line
    - add boundary
    - remove selected topic
    - balance map

4.2.1.4 Procedure

After a brief introduction to the purpose of the interview some of the possible applications of the electronic mind mapping tool were presented (planning a meeting, creating a decision tree, conducting a brainstorming session) and each of the tioned basic functions was demonstrated in the program by the interviewer. Before advancing to the practical part of the interview, the participants were queried about:

1. the most practical/useful aspects of the program (and their general inclination)
2. their understanding of the individual functions

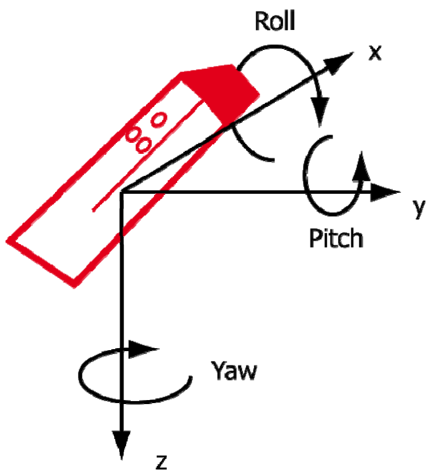


Figure 6. Movement axes of laser-pointer device.

The means to execute a gesture was then shown to the participant in a short demonstration of possible movements that could be performed with the device (up-down, left-right, forward-backward). Although roll (see Figure 6) was not explicitly excluded from the potential movements, no demonstration was given and none of the participants included this movement in any of their proposed gestures.

Standing in front of the screen, the user was then asked to imagine using the laserpointer for executing the respective command for the program functions running on a large display. The interviewer encouraged the actual use of the device for trying out ideas for appropriate gestures, and gave feedback once a suggestion for a discrete gesture was produced by the participant. After sketching down the movement, the interviewer then showed the drawing to the participant, and prompted for the reason why the particular gesture occurred to the participant as an appropriate representation of the given function of the program. Finally, depending on the number of dimensions used, the participant was invited to think of a further gesture for the respective function performed either in a 2D plane or using all three dimensions. The interview ended with three open questions concerning the "feeling" of the gestures and if there could be improvements to the interaction concept. A last question was included to determine the potential use of vibrational feedback by the device.

4.2.1.5 Data Analysis

The participants' responses were noted down by the interviewer in the respective language and later translated into English. The interviewer ratings on the respective scales were encoded together with the occurrence of all instances of proposed gestures in an Excel spreadsheet. All sketches of these gesture proposals were marked with the code from Excel sheet. In case that more than one participant proposed the same gesture, the number of occurrences of this individual gesture was calculated and put in relation to the overall number of proposed gestures for the respective function to obtain the ratio of agreement.

4.3 Results of the pre-study

Nobody of the participating individuals had problems grasping the purpose and underlying concepts of the mind mapping program. Three participants each reported either a readiness to try out the main functions or their willingness to undertake a serious try with the application. One participant expressed his intent to realize a project using the program in the near future. Concerning the understanding of the basic functionality, all participants reported a good general comprehension of the concepts and the respective commands for both navigation/visualization and editing functions, with only one person expressing a good general understanding of the concepts but a few difficulties with the navigation/visualization functions.

Overall, 72 2D and 58 3D gestures were collected during the interview session. This amounts to a response rate of over 100% (some participants suggested two gestures) for 2D and 82,86% for 3D. For 3D gestures only, in 6 of the 10 cases (program functions) some participants failed to come up with ideas. After revising the collection by grouping similar gestures, two preliminary sets were compiled with 56 gestures performed on the 2D x-y plane and 36 gestures performed in 3D or on a 2D y-z plane. The reason for the eventual reassigning of the latter instances was the common observation that the participants did not strictly remain on the 2D y-z plane when performing the movement. Hence, although these gestures were strictly speaking 2D gestures, they were assigned to the 3D group in the revision. Appendix B shows the resulting compilation after the interviews. The agreement among participants concerning individual gestures was higher for 3D although only for three of the ten functions. Figure 7 shows the percentage of users that proposed the same or an equivalent gesture for five program functions.

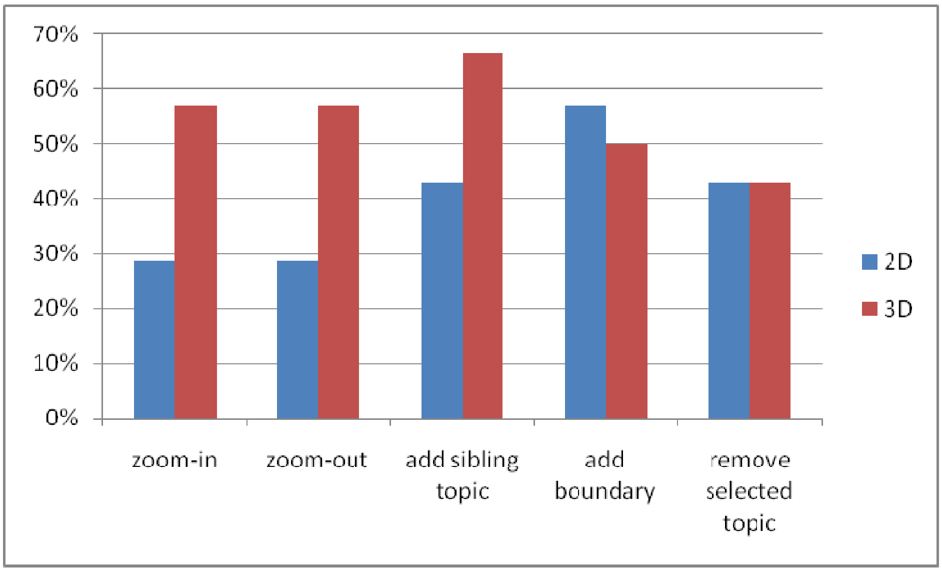


Figure 7.Agreement among participants for gestures proposed for the respective program function.

Finally, six of the seven participants would consider a vibrational feedback as useful for conveying information about the status of the gesture recognition by the computer or for other stages during the interaction.

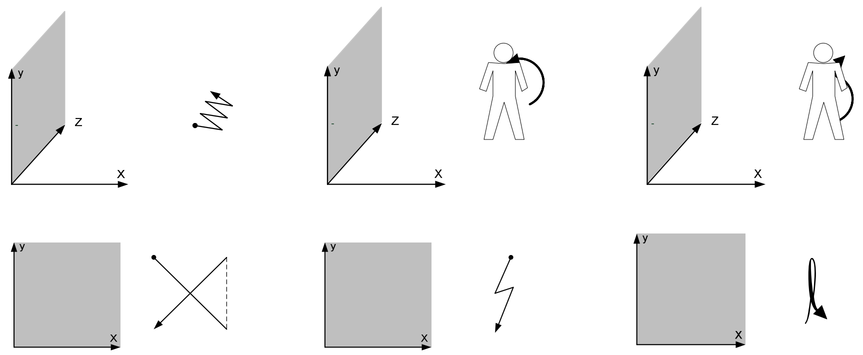


Figure 8. Sample of 3D (above) and 2D (below) gestures proposed for the "remove topic" task

4.4 Discussion and conclusions of the pre-study

The overall response rate in the gesture interview reflects Kela et al.'s (2006) observation that users tend towards gestures on a 2D x-y plane. The basic hypothesis for this pre-study was hence supported. While all functions received at least one proposed gesture command in 2D the following cases failed to do so in 3D from one or more participant (indicated in brackets):

- split map view vertically (1)
- remove view split (2)
- add sibling topic (1)
- add subtopic (2)
- add relationship line (2)
- add boundary (3)

This finding of course leads to the question for which reason participants in the interviews conceived more proposals for gestures remaining on a two-dimensional plane. Some comments suggest that the required use of force by the arm and wrist hindered the creative process for trying out the gestures. Further clues might be gained from participants rating the appropriateness of the number of dimensions used in the gesture as depending on the function at hand. One participant explicitly mentioned that he had adopted a pen-metaphor while handling the laserpointer which implicitly entails a two-dimensional approach to the execution of gestures.

This difference in overall number of proposed gestures notwithstanding, the result concerning the agreement among participants' ideas of gesture commands for individual program functions reveals a further aspect of the difference between the two gesture groups. That is, although in six of ten cases one or more participants failed to come up with a suggestion, the agreement in three of ten cases was higher for gestures using all dimensions (3D). Although neither of these ratios can itself suffice to justify a distinction of gestures based on the employed dimensions, the larger agreement for at least three program functions may be an indication that there simply are fewer alternatives that participants might have deemed appropriate as respective 3D gesture commands. This may be due to the mapping of the gesture command on the program function which is possibly more direct than for the other functions and gestures or when compared to bidimensional gestures. Also, a shared mental model could have played a role in bringing about this result.

In general, the new gesture interaction concept was well received. One remark indicated the wish for the gestures to remain unconspicuous in front of an audience while another expected a presentation to be rendered "più vivace" (more lively) because less time would be spent to recall commands. A further interesting comment related the ease of remembering to the closeness of the gesture movement to something familiar (giving the example of wiping for the remove-function).

The interviews in the pre-study supported the basic assumption that users would find more 2D gestures overall. A distinction on the basis of spatial dimensions was however deemed less meaningful than initially assumed because with a handheld motion-sensing input device, the performing of the gesture movements should not be constrained to a certain plane or number of number of dimensions. For the current analysis, the initial focus on dimensions was on the one hand due to the orientation on existing findings in the literature and because it facilitated notation. Also, prior to conducting the pre-study interviews, it could not be foreseen which strategy participants would adopt when generating ideas for gestures. In other words, no assumptions were made concerning potential properties of the context (mind mapping program and functions, screen, input device) or the cognitive processes of the users, that could have influenced the nature of the proposed gestures. Moreover, the differences in agreement for specific gesture commands that were found for the respective program functions lend a different weight to the results. The preliminary compilation of two sets of gestures according to the spatial dimensions of their movement trajectories is hence mainly followed up in order to create a working basis for a deeper examination of the included gestures. In the following chapter, a semiotic analysis is conducted to determine potential features that are meaningful in an HCI context and would allow the creation of gesture sets to be evaluated and compared with the working laserpointer system inside a usability test.

5 Establishing gesture sets

The preceding discussion of the characteristics exhibited by the gesture proposals given during the interviews remained largely on a physical level. Following Nespoulous and Lecours' (1986) considerations on levels of analysis, an exploration of defining features based on the spatial dimensions is restricted to the substance of expression level. While such a discussion might suffice for other kinds of gestures in HCI (e.g. pen-stroke gestures) it should be more promising to include analyses of how instances of gestures are distinct in their form (form of gestural expression), what each form is presenting (form of gestural content) and what the final signification is (substance of gestural content). The focus would hence be shifted to a semiotic line of investigation taking into account:

- how the gestural unit (the signifier) is “formally” distinct from others
- what is contained in the gesture (the signified)
- how and what meaning is established

All the presented steps towards the discovery of meaningful distinctions to be applied to the obtained gestures hereafter involve an underlying working hypothesis defining the level of analysis and one or more proposed differentiating features. The final aim is to arrive at two or more distinctive sets of gestures that could be scrutinized as part of a user test concerning the overall usability (through measures of efficiency and satisfaction) and learnability/memorability (through measures of effectiveness and satisfaction) of gestures as an input modality (see ISO 9241-11).

5.1 The arbitrariness distinction

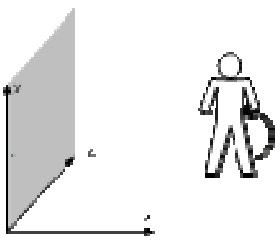


Figure 10. Proposed gesture for "add sibling topic".

The first working assumption is that the gestures in the present collection can be differentiated according the connection between their form of expression (the signifier) and form of content (the signified). That is, how does the physical gesture unit (the distinct movement episode) relate to what it presents (the content). It should be noted that the final signification is not yet touched (here: the actual intended command for a program function). Nespoulous and Lecours (1986) suggest a distinction based on the degree of arbitrariness. Arbitrary gestures are here distinguished from mimetic gestures because understanding the link between the gestural unit and its referent requires prior learning (as in the

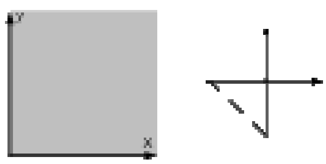


Figure 9. Proposed gesture for "zoom-

"money" example mentioned above). Mimetic gestures on the other hand rely on the transparency of the link as presented in the example of the imitating gesture for running (shuffling of both arms) or the outlining of a metropolis' skyline in the air (1986, p.57). The two latter examples are most comparable to the gestures performed by the users in the pre-study. As can be seen in Figure 10, one gesture proposed for the "add sibling topic" operation shows a mimetic relationship to a "putting-something-in-front" movement by way of imitation. Likewise, the movements sketched in Figure 9 exhibit a mimetic link with a conventional symbol for "PLUS". It can thus be safely assumed that all gestures included in the revised collection are mimetic either because they entail drawing something "in the air" or imitation of an action. This has been confirmed by Nespoulous (personal communication). All instances are characterized by what Nespoulous and Lecours (1986) call their "iconicity" and the particular meaning of what is presented must then be determined through further examinations. Such an analysis could be conducted on the level of "substance of content" because one is concerned with the connection between the form of content (e.g. the "PLUS" symbol) and its signification (the command for the "zoom-in" program function). Despite failing to refer to a level of analysis, it is at this level where most classification attempts in the literature appear to have been made and where the next step in the current semiotic analysis will take place.

5.2 The iconicity dimension

Given the strong iconic character of the gestures elicited with the current motion-sensing laserpointer, the loading on this dimension can be assumed to be very high. Following McNeill's (2005) suggestion to think in terms of dimensions rather than kinds, the current step involves the scrutinization of the gestures in the revised collection for the degree of such loadings. The identification of the dimension(s) proceeds along the following considerations:

- the laserpointer device affords two independent uses: pointing and motion-sensing
- pointing allows interaction through direct manipulation: strong deictic character of the user action
- motion-sensing allows interaction through gestures
- gestures serve the function to communicate commands to the computer: strong iconic character of the user action (results from gesture proposals)

Further examinations concerning the gesture interaction should thus be conducted on the dimension of iconicity. Though there are indications of deixis present in some of the collected gestures (the movement in Figure 10 proceeds from the side to the front center), this deixis dimensions does not play a role in the communication with the machine. Nonetheless, the directional information conveyed with deixis may play a role in the connection between the form (form of content) and the meaning (substance of content), independently of its importance for the interaction process. The average loadings exhibited by the collected gestures were determined by the author on the basis of McNeill's (2005) definitions and are shown in Figure 11.

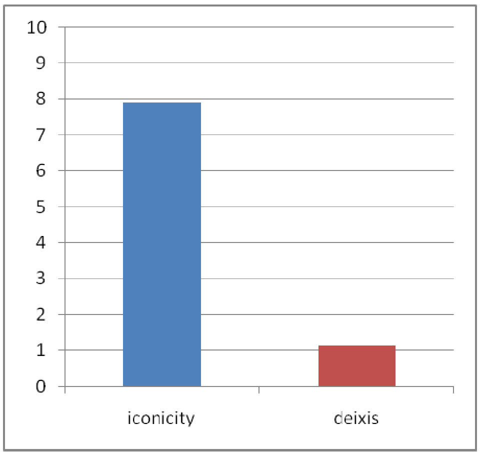


Figure 11. Average dimensional loading for collected gestures (loading range from 0-

The subsequent step concentrates then on how this iconicity exhibited by the gestures characterizes the establishing of the final meaning which in the current context are the respective program functions of the mind mapping application. The aim is to further examine what aspects of this dimension can be differentiated that could potentially make a difference for the gestures' impact on the usability of the interaction process. Following the visual relationships (see Box 6) put forward by Ekman & Friesen's

(1969), the visual character inherent in iconicity (from Greek εκών: *image*) can be determined as follows for the collected gestures:

- pictorial (the content of the gesture presents a picture)
- kinetic (the content of the gesture presents an action)

It must be noted that it is likewise possible to create dimensions for these visual relationships, just as deixis refers to nothing less than the pointing relationship described by Ekman and Friesen (1969). However, in order to not contribute to the already highly redundant categories in the existing classification systems, the current differentiation is oriented on the strong iconic quality of the gestures obtained in the pre-study.

In order to further reflect on what is presented in these gestures, the final distinction that is applied to the revised set eventually used for the test, is the following:

- symbol/object iconic (the content of the gesture presents a symbol or an object)
- movement-iconic (the content of the gesture presents a movement)

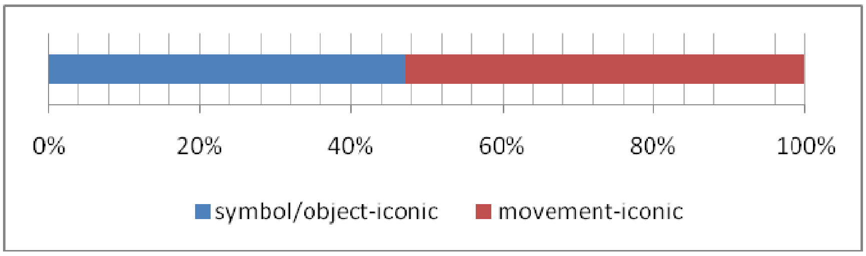


Figure 12. Ratio of iconicity-specific gestures in the set used for the user test. (symbol/object-iconic: N=25 movement-iconic: N=28

This is echoed by McNeill when he defines “imagery in gesture [as being] actional as well as visuospatial” (2007, p.2). Moreover, this distinction also manifested itself in the different levels of dynamic execution across the proposed gestures during the interviews. It refers to the observation of how elaborate users’ movements were for each gesture type. This property was however not measured quantitatively and remains thus a subjective characteristic.

5.3 Creating the final gesture set

The final revision of the gesture collection included the cancellation of all gestures that were proposed for the following program functions:

- split map view vertically
- remove view split
- add subtopic
- balance map

Instead, the combination of copy/paste was added to the list of functions to be employed during the user tests. On the one hand, the revision reflects the concern that the final number of gestures that were to be remembered by participants in the user tests would be too high. In comparison with the number of gestures included by Schlömer et al. (2008; 4 gestures) and Kela et al. (2006; 8 gestures) the current number of individual gesture commands would have amounted to ten, respectively 12, if all functions including copy/paste were accommodated. The final two sets contained eight gestures each, covering the following program functions:

- zoom-in
- zoom-out
- copy
- paste
- add sibling topic
- add relationship-line
- add boundary
- remove topic

The individual gestures are presented in the following chapter. Since there had been no gesture proposals made for the copy/paste-commands, the author included his own ideas in the respective sets (see Table 2).

5.4 Individual mental models

The distinction on the iconicity dimension put forward in the semiotic analysis presented above is considered to be made on the level of signification (substance of content). Since the process of generating ideas for gestures during the user interviews initiated with the respective reference by the interviewer to the demonstrated program function, one underlying assumption in the investigation until this point has been that all steps in the analysis of the gestures of the initial collection regarded properties of importance in the participants' mental processing of the program functions, the gestures and their connection. However, due to the limited agreement among the proposed gestures, the final two sets do not necessarily reflect "the best choice" in terms of familiarity and along some of the usability attributes presented below. In the end, the assignment of a particular gesture to be used as a command in the user test fell to the author. As part of an iterative process, the final sets were nonetheless repeatedly scrutinized for their suitability in many discussions before the actual testing began. It is thus acknowledged that although the different types of iconicity could be distinguished based on the proposals of the participants, the *level* of "perceptual similarity" (Roth, 2005) relates much closer to the individual perception of each user. However, investigating the specific mental models each participant created during the interviews (integrating background knowledge with the particular program functions and the way of interaction) would have required a much more time consuming process of questioning and protocolling of the understanding and interpretation exhibited by each individual. Since the global aim of the study is to investigate the suitability of gestures for large-screen interaction purposes, the concentration remained on the nature of the gestures and their practical application. Nevertheless, in generating hypotheses for potential differences that might arise between the two types of gestures, reference is made to the perceptual character of the gesture content and how it may influence the cognitive storage and retrieval processes of users' learning and recall efforts. In particular, a claim is made that for gestures presenting pictorial information, creating the connection to the final meaning (the program function) requires more explicit mental elaboration (creating mnemonic aids) as compared to gestures containing movement imitation. Moreover, the latter imitative character is considered to render the gesture more familiar to the users. These assumptions would have to be assessed more thoroughly in an experimental setup focusing on the cognitive processes involved in learning of gesture commands. First indications though are given in the presentation of the questionnaire results and by some of the comments recorded during the tests.

6 Evaluating gesture interaction

In the preceding chapters, the process of identifying potential differentiating features of gestures on the different levels of analysis is reported. The integration of the insights of this effort into Karam's (2006) framework and their application to the classification of gestures collected during pre-study interviews allowed the creation of two distinct gesture sets differing on the dimension of iconicity. This chapter describes the steps in the evaluation of this final collection of gestures concerning their suitability as an input modality for the interaction with the Powerwall, a large high-resolution display, in the context of an electronic mind mapping session. The assessment regarding the specific usability objectives (as suggested by the ISO 9241-11 standard) were at the core of the investigation. As described by the standard (ISO, 1998), usability refers to "the extent to which a product can be used by specified users to achieve specified goals with effectiveness, efficiency and satisfaction in a specified context of use" (1998, p.2) with this context being comprised of users, tasks, equipment as well the physical and social environments in which the product (hardware, software or materials) is used. Any given appraisal concerning the three yardsticks of usability can employ various measures depending on the particular objective of the investigation. In general, effectiveness is defined as the "accuracy and completeness with which users achieve specified goals", efficiency describes the "resources expended in relation to the accuracy and completeness with which users achieve goals" and satisfaction is characterized by the "freedom from discomfort, and positive attitudes towards the use of the product" (all quotes from ISO 9241-11, 1998). In what follows, the conditions given for the current evaluation are specified inside the extended framework for the design and research of gesture-based interaction introduced in the first part of this thesis. Prior to that, the adopted objectives and resulting hypotheses concerning the research questions are presented.

6.1 Hypotheses

Two major usability objectives relevant to the research questions are of importance, desired overall usability and learnability:

- Learnability
    - how many of the gestures are learned in a given amount of time (effectiveness measure)
    - Memorability – how many of the gestures are remembered given a fixed time lag (effectiveness measure)
    - how easy is it to learn the gestures (satisfaction measure)
- Overall usability
    - time to complete a task list (efficiency measure)
    - joy-of-use (satisfaction measure)

In addition, assessing the need for support should indicate the level of independence with which users could use the gestures.

- support requirements
    - number of calls for help (effectiveness measure)

The operationalizations for the above constructs are presented in the method section. The main hypotheses had their focus on the success with which the users could recall and apply the gestures in general, and if compared between the types of gesture (sets).

- [h1]: There will only be a slight decrease in the overall number of successfully recalled gestures at the beginning of session 2 when compared to session 1. This outcome is also expected for the number of correctly used gestures during the task list phase.

This hypothesis concerns the general degree of learnability and memorability of the gestures, which was expected to be high due to the combination of implicit and explicit learning by the users. The learning process in the current setup is assumed to involve explicit information of the gesture content and the mapping onto the program function, as well as implicit knowledge through practicing the movement sequences. As suggested for example by Gentile (1998), these processes may act in parallel though no claims could be found about potential advantages as formulated for the gestures in the current tests.

- [h2]: In general, more movement-iconic gestures than symbol/object-iconic gestures will be recalled during the prompting interviews.

Movement-iconic gestures are expected to exhibit an advantage in initial learning as well as retention over a one-week time lag. A conjecture is made concerning the underlying reason by referring to users forming a more direct connection to the respective program function when learning this type of gesture.

- [h3] More movement-iconic gestures than symbol/object-iconic gestures will be performed correctly in the task list phase of both sessions and across scenarios.

This gesture type is expected to be superior concerning the correct use by the users. A larger degree of familiarity that facilitates execution is suggested as an influence.

6.2 Method

First, relevant aspects for the description of the experimental method are presented in terms of the categories provided by the extended framework. The scheme lists these categories (e.g. “Application domain”), subcategories (“Interaction context”) and parameters (“Physical requirements”) with the respective values if applicable. For example, if “discrepancies with parameters” has the value “no”, it is implied that user goals cannot change some of the other parameters. In other words, even if a user should alter his or her goal of what to achieve in the software application, other parameters such as “complexity” remain the same since the overall characteristics of the tasks do not become less or more complex as a result of this goal change. For detailed descriptions of the building blocks of the framework, see Karam (2006).

The evaluation was conducted as a formal usability test similar to lab testing as e.g. described by Barnum (2002), involving supervision and direct observation by the author and video-recording of the sessions. The room at the University of Konstanz housing the Powerwall was used as the test-site. Apart from the camera equipment, the conditions during the test were intended to be as close as possible to potential usage situations. This was supported by distributing the testing phases over two sessions thus separating the learning and initial use of the gestures from their independent application by the participants during the scenarios.

6.2.1 Usability Test

6.2.1.1 Participants

The target users for the gesture interaction inside an electronic mind mapping context are students and teaching personal alike. Assuming that creativity and planning tools of this kind could potentially be employed in regular as well as irregular occasions such as meetings, presentations and seminars, it should not be necessary to have specific skills or knowledge of the system to be used for productive purposes. The sample for the usability tests thus included six male and seven female right-handed undergraduate students from non-computer science subjects with a mean age of 23,3 years. Five participants indicated a regular use of their computer of 2-3 hours a day, while five reported a daily use of more than three hours. Half of all participating users had used a laserpointer at least once before, though not interactively as a means of input. All were familiar with the standard mouse/keyboard combination with six persons having tried out a WiiMote and one a presenter before. Of the seven individuals having used a form of gesture input before, three indicated pen or mouse gestures and six game gestures realized on a Wii console. Finally, none made use of any form of mind mapping regularly or occasionally, with four participants reporting no knowledge of the technique at all and six claiming to be familiar with the concept. Only two had ever tried an electronic mind mapping tool before. All participants had to sign a consent form (see Appendix D) prior to the test to allow the videotaping of the session for backup reasons and were paid €10, - in compensation after the second session. Participants were assigned randomly to the two conditions with one group being demonstrated the symbol/object-iconic gesture set and the other the movement-iconic gesture set.

6.2.1.2 Equipment

The experiments were conducted in the premises of the Powerwall room of the University of Konstanz, Faculty of Information Engineering. Participants were standing up in front of the 5.20 × 2.15 m (17.06 × 7.05 ft) display at a distance of three metres indicated by a centered floor marker orthogonal to the display. The laserpointer could be handled freely to allow maximum flexibility for the interaction during the tasks. However, since the continuous wireless transmission of the motion-sensor signals could not be insured by the device at that stage of development, a cable needed to remain attached during all sessions.

Other equipment included a control workstation placed on the left side of the room levelling the standing user while a 1.10 m lecture desk with a wireless mouse/keyboard combination connected to the display remained at the left edge of the display to allow test preparation by the supervisor. The video camera remained fixed in a position of 1,50 meters distance in an angle of approximately 110° from the user facing the display. Likewise, the supervisor seat was chosen on the left side of the user in a position levelling the tripod. This constellation was chosen to allow participants to be aware of either while not obstructing their field of view or action. As a consequence, the information extractable from the video recordings was limited as the focus was on the users' arm movements and the coarsely visible display actions. This will be noted later when discussing the system's recognition accuracy.

- *Application domain*
  - *Interaction context: LDG*
    - *Physical requirements: medium*
    - *Cognitive requirements:medium*
  - *Task characterstics*
    - *Criticality: medium*
    - *Complexity: low*
  - *User goals*
    - *Discrepancies with parameters: no*
- *Enabling technologies*
  - *Input device: motion-sensing laserpointer*
    - *Interaction zone: 0*
    - *Mobility: .5*
  - *System performance*
    - *System accuracy: medium*
    - *System response time: high*
- *Gestures*
  - *Gesture style*
    - *Body part: hand*
    - *Objects: yes*
  - *Gesture set: 8 gestures*
    - *Complexity: medium*
    - *Task mapping: one-one*
  - *Classification*
    - *Function: semaphoric*
    - *Dimensional loading*
      - *Physical: dynamic*
      - *Semiotic: iconicity, deixis*
    - *Visual relationships: pictorial, kinetic*
- *System response*
  - *Modality: visual*
- *Feedback*
  - *Reflexive: no*
  - *Recognition: yes (tactile)*
  - *Response: no*

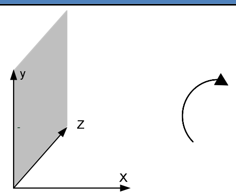
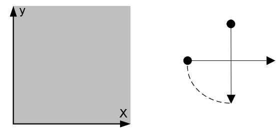
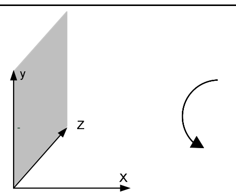
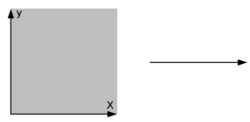
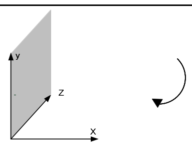
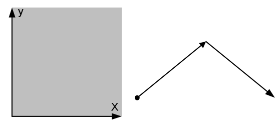
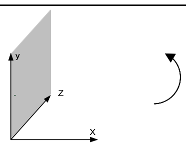
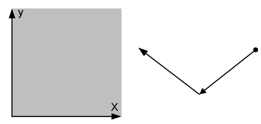
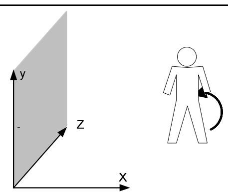
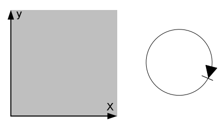
Table 1. Application of the extended framework to the evaluation of the gesture interaction using a motion-sensing laserpointer (based on Karam, 2006).

Direct pointing was enabled by having the reflection of the infrared laser beam on the display tracked using an arrangement of infrared cameras. The cursor thus followed any pointing movements by the users in an immediate fashion. The gesture interaction was realized by feeding the acceleration data from the motion sensors into a recognition algorithm based on Hidden Markov Models (HMM) as introduced by Schlömer et al. (2008). The original configuration reported by these authors achieved an average recognition rate between 84% and 94,3% with four gestures. The rate reported by Kela et al. (2006) for a set of with eight gestures lay between 81,2% and 98,9% depending on the optimization efforts on the underlying models. For the current study, a similar elaborate process of model revision to increase the recognition rate was not possible. The algorithm applied to the motion information in the current system was thus deemed sufficient to enable the use of an eight-gesture set for the interaction. After a series of pre-tests, a recognition rate of 88% was observed when using movement-iconic gestures. No such information is available for the symbol/object-iconic set because of missing recordings of the respective pre-tests sessions.

The MindManager 2007 software (MindJet, 2007) was chosen as the mind mapping application and ran on the main computer connected to the Powerwall display with the maximum resolution of 4640 × 1920 px. Demographic data, computer experience and ratings of the various aspects of the interaction were obtained with the help of pre- and post-test questionnaires. Participants were provided a clipboard to fill out the questionnaire while seated. Together with documents containing a general introduction, the manual for using the application and the gestures and the task instructions, all participants received the same verbal and written information in both sessions. Procedure manuals for the supervisor were devised to allow each session to be conducted identically. Finally, protocol sheets were used to keep track of task completion and explicit recall of the gestures.

6.2.1.3 *Tasks and scenarios*

The task list and scenario elements of the user test included the following program functions to be executed using the respective gesture commands:

function	gesture command (movement-iconic)	gesture command (symbol/object-iconic)
zoom-in		
zoom-out		
copy selected		
paste selected		
add sibling topic		

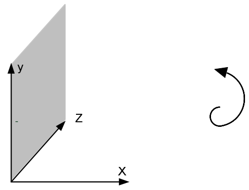
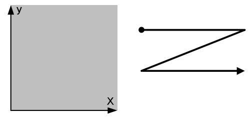
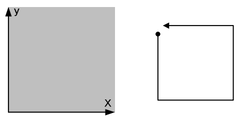
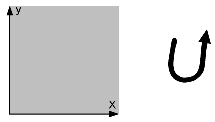
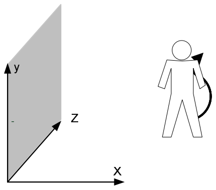
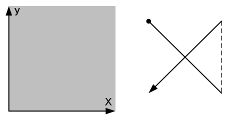
add relationship line		
add boundary		
remove selected topic		

Table 2. Program functions and their gesture commands used in the user test.

The *task list* to be completed in both sessions contained eight individual instructions for performing different commands on a basic mind map with four existing elements distributed around a central topic.

1. Beginning from the central main object (“main topic“), create four new objects connect to this central element.
   Add to each of these new objects two further directly connected objects.
2. Connect any two objects with a line to create a thematic relation.
3. Select an object and copy it.
   Insert the copied object at any given location in the map.
4. Zoom at will into the map to gain a more detailed view of the elements.
5. Select an element of your choice and delete it.
6. Zoom out of the map to obtain an overview of your constructed map.

The two *scenarios* at the end of session 2 offered the free application of gesture commands by the participants. The goal of each scenario was to adjust the displayed map according to the differences in a target map printed on the instruction sheet. Participants were told to have reached the goal when all necessary operations to arrive at the target map were performed, even if the adjusted map was not an exact duplicate.

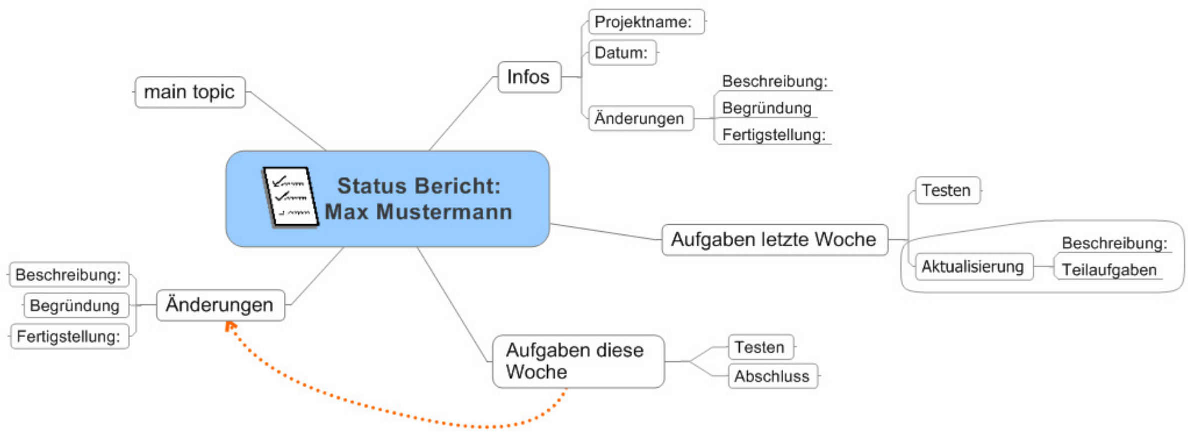


Figure 13. Target map for scenario 1.

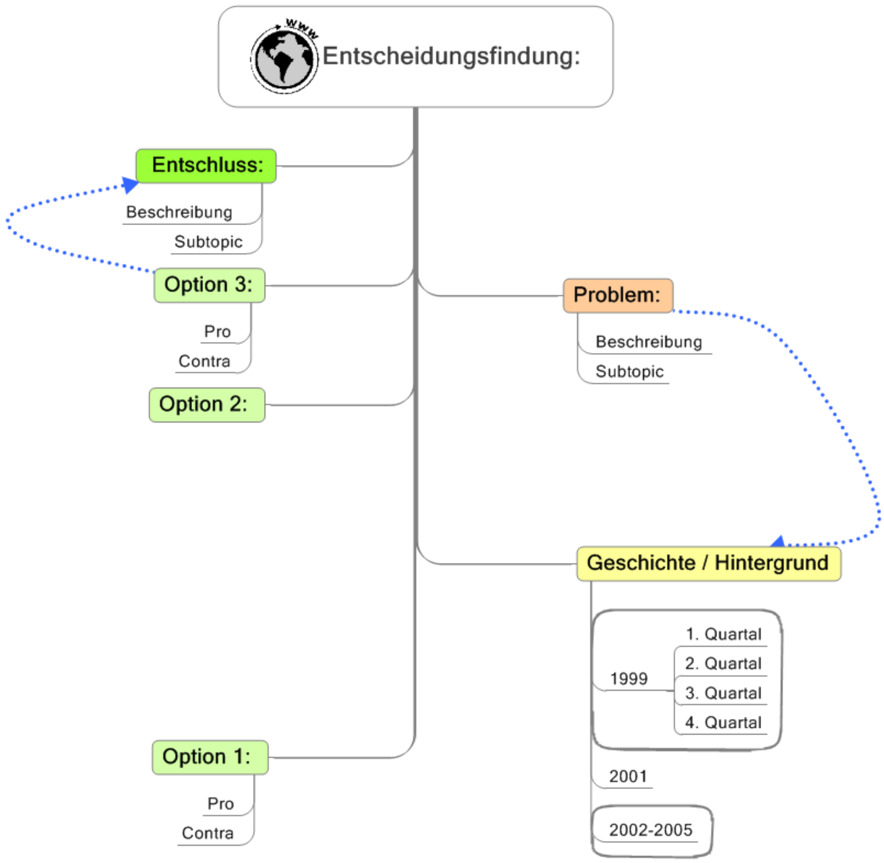


Figure 14. Target map for scenario 2.

6.2.1.4 Procedure

At the beginning of the first session the participant was welcomed in the Powerwall room and given a short description of the display. After reading the introduction sheet, he/she filled out the personal information and computer experience questionnaire. Before being asked to sign the consent form, the video recording was explained in order to reduce possible apprehensions of being videotaped. The user was then lead to a chair facing the display (though behind the marker) and read the first two pages of the test manual. These contained short descriptions of the interaction modes (direct pointing and gestures) as well as an overview of the basic program functions (see Appendix C). The latter were subsequently demonstrated in a simple map by the supervisor (not yet employing the laserpointer). The participant then read the description of how to execute gestures with the laserpointer. This was then again demonstrated and the participant was given the chance to try out direct pointing by handling the device for a maximum of one minute. Prior to starting the training phase, the user read the instructions on how training would proceed and which functions were being covered by gestures.

In the context of machine learning, "training" usually refers to the feeding of data to an algorithm and the subsequent creation of a model. Apart from the training of the system with the specific properties of each gesture's movements by the different participants (one training set was created for each user and reused in the second session), this initial part of the current test also served as a learning and practice phase. It was conducted through the supervisor naming the function and demonstrating the respective gesture. In order to avoid any leading information concerning the content (form) of the gesture during this learning phase when participants were forming individual mnemonic aids, the supervisor made sure furthermore that no description was given of what the gesture was presenting or implying, and responded with a nod or simple verbal affirmation to any conjectures uttered by the user.

Each gesture had to be consecutively performed 15 times before the movement data was transmitted to the model by the supervisor. The order of the gestures was counterbalanced inside each group of participants to avoid sequence effects in practice (Reese, 1997). After each gesture, the supervisor switched the device from training to recognition to allow users to try out the gesture and observe its effects on the display.

Before continuing to the task list phase, a document (see Appendix C) was handed out containing the instructions for completing the task list. Users then indicated when they were ready and were asked to stand at the marked position in front of the display. Each individual task was performed in a sequential manner with the aim of creating a first map. Participants were instructed to work through the list on their own and only turn to the supervisor in case of system failure or when they did not recall the respective gesture at all. Nevertheless, reduced verbal feedback was given to not shut down the communication channel between supervisor and participant and create a potentially discomforting artificial situation (Boren & Ramey, 2000).

User performance was logged manually using a scoring sheet to indicate the level of effectiveness in executing the correct gesture according to the following criteria:

- Easy: gesture used correctly on first try
- Medium: gesture used correctly on second or third try with apparent difficulties
- Hard: gesture used correctly on third or fourth try with expressed difficulties

Only when assistance by the supervisor was necessary did the specific task not receive a “pass” rating but was considered “failed”. In addition, three levels of assistance were distinguished:

- L1: user is asked to reflect once more
- L2: user is asked to try again
- L3: user receives a demonstration of the correct gesture

In case the system failed or wrongly recognized a correctly executed gesture after more than three tries, the user was instructed to continue to the subsequent task.

After completing the task list, the participant filled out the rating questionnaire. Following this, he/she was prompted to demonstrate the respective gesture of the program function named by the supervisor. The order of functions during these prompting interviews was also counterbalanced to avoid recency effects from the prior gesture performance of the task list phase. The participant was then thanked and reminded of the ensuing test appointment in one weeks' time.

Figure 15. Test phases of the two sessions separated by one week. D=Demonstration & Training, TL=task list, Pr=prompting interviews, Sc=Scenario.

Figure 16. Using the motion-sensing laserpointer (wireless version) for gesture interaction and direct pointing with an electronic mind mapping application on the POWERWALL large high-resolution display

The second session began with a prompting interview conducted in the same fashion and again counterbalanced according to a digram-balanced latin square (Wagenaar, 1969). Each participant was then given once more the sheet containing the task list instructions. After completing the task list, the supervisor explained the scenario phase and handed over the document showing the first target map. The user adjusted the map on the display employing the respective gesture commands in free order. As in the task list phase, the supervisor logged the success in gesture execution. Expected problems during this logging like confounding the user's intended and actual performed gesture did not emerge in any of the scenarios across all participants.

6.2.1.1 Measures

The respective operationalizations and measures for the constructs of the usability objectives are listed in Table 3. Since every user participated in two sessions with a one-week lag, the points of measurement are included in the list to illustrate on what data the measure is based. Moreover, the particular instrument employed is indicated as well. The dependent variables measured by the instruments are the following:

- Learnability
    - Recall prompting (interview): number of gestures recalled
    - Tasks (task list/scenario): number of correct gestures executed without help
    - score on items (questionnaire): g2, SD3
- Overall usability
    - task list/scenario: overall completion times
    - score on items (questionnaire): g4
- support requirements
    - tasks (task list/scenario): number of "assist" scores

The independent variables include:

- Between-subjects factor
    - type of gesture: symbol/object-iconic vs. movement-iconic
- Within-subjects factor:
    - session : session 1 vs. session 2

construct	operationalization	instrument	possible scores
Learnability		*session 1*	
	how many of the gestures are learned in a given amount of time	recall prompting interviews	remembered / not-remembered
		task list	performed correctly on 1st try / failed to perform correctly on 1st try
		session 1 vs. session 2	
	Memorability – how many of the gestures are remembered given a fixed time lag	recall prompting interviews	remembered / not-remembered
		task list	performed correctly on 1st try / failed to perform correctly on 1st try
		session 1	
	how easy is it to learn the gestures	questionnaire	rating score
Overall usability		*session 1 vs. session 2*	
	time to complete a task list (effi-ciency)	task list	time
		session 2	
	time to complete a task scenario (efficiency)	scenario	time
		session 1 vs. session 2	
	joy-of-use	questionnaire	rating score
support re-quirements		*session 1 vs. session 2*	
	number of calls for help (effective-ness)	task list, scenario	L3 scores

Table 3. Constructs included in the usability objectives under investigation in the user tests and their respective operationalizations and instruments.

6.2.1.2 *Data Analysis*

The experiment was a 2x2 mixed-model, repeated measures factorial design with one between-subjects factor (type of gesture) and one within-subjects factor (session). User performance for the scenario phase entailed only type of gesture as the between-subjects factor while the order of gesture demonstration during the training phase was counterbalanced for each group but nevertheless included as a between-subjects factor (sequence) to be analyzed for potential effects. With regard to the relatively low recognition accuracy exhibited by the system during the tests (63% overall as opposed to 85% reported by the authors of the algorithm), data concerning subjective ratings and completion times are assumed to be more affected than scores of recall and successful execution which were administered by the supervisor.

6.3 Results of user test

The number of gestures that could be successfully recalled during the prompting interviews was generally very high with 93,75% correct responses at the end of session 1 and 84,38% at the beginning of session 2.

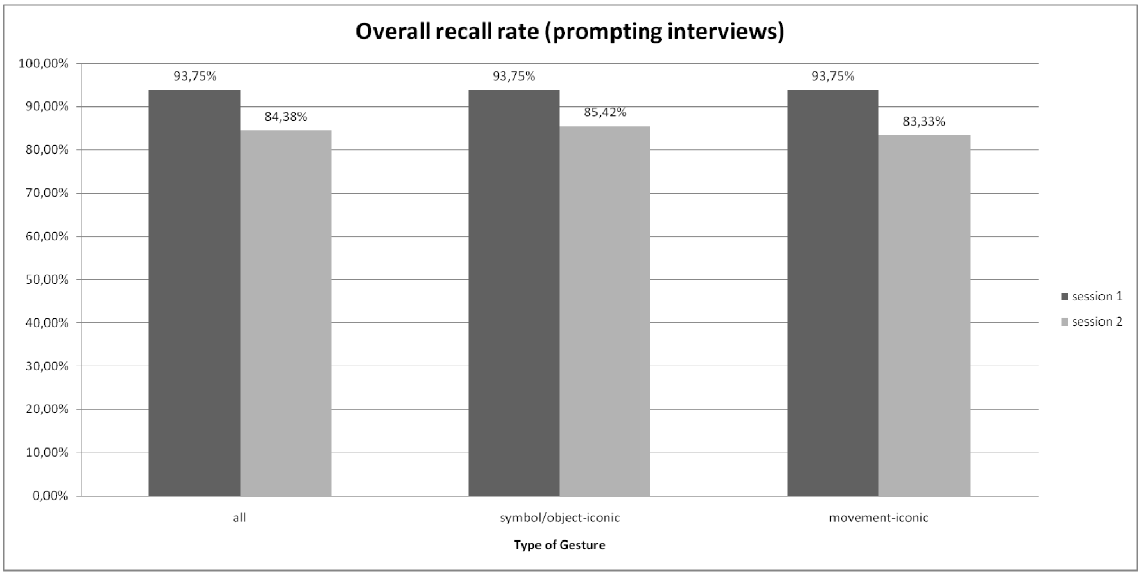


Figure 17. Percentage of correct responses during recall prompting in session 1 and 2: a) across all gestures (N=96) b) for symbol/object-iconic gestures (N=48) c) for movement-iconic gestures (N=48).

The number of recalled gestures did not differ significantly by gesture type neither for session 1 ($\chi^2(1, N = 96) = 1, p > .05$) nor for session 2 ($\chi^2(1, N = 96) = 0{,}779, p > .05$). 41 out of 48 symbol/object-iconic gestures were recalled successfully at the beginning of session 2 with 40 out of 48 correct responses for movement-iconic gestures. Users were able to execute the correct gesture without assistance at an overall rate of 93,52% when completing the task list in session 1 and 95,37% in session 2. In session 1, out of an overall number of 54 gestures in the task list phase for each group, 50 were performed without assistance by the participants in the symbol/object-iconic group and 51 by the participants using the movement-iconic gesture set.

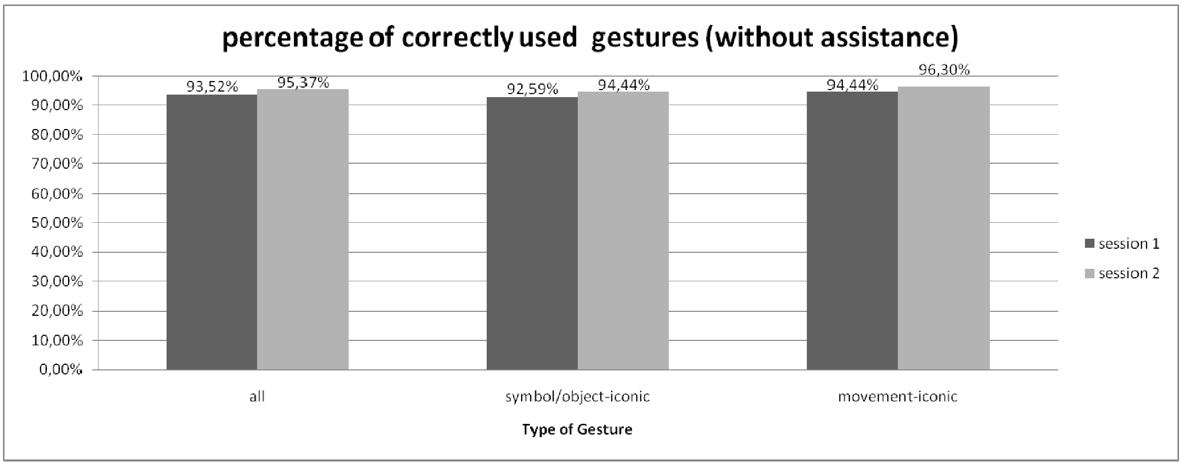


Figure 18. Percentage of correctly performed gestures during the task list phase in session 1 and 2 a) across all gestures (N=108) b) for symbol/object-iconic gestures (N=54) c) for movement-iconic gestures (N=54).

For session 2, these numbers were minimally better with 51 out of 54 performed without assistance by the symbol/object-iconic gestures group and 52 for those using the movement-iconic gestures. As expected after observing these data, the number of correctly used gestures did not differ significantly by gesture type neither for the session 1 task list ($\chi^2(1, N = 108) = 0{,}696$, $p > .05$) nor for the session 2 task list ($\chi^2(1, N = 108) = 0{,}647$, $p > .05$). Results from the scenarios show that users in the respective group were able to perform the movement-iconic gestures without assistance for all necessary operations required to arrive at both target maps (48 out of 48 successful in scenario 1; 90 out of 90 successful in scenario 2).

Participants using symbol/object-iconic gestures successfully performed 46 out of 48 in scenario 1 and 89 out of 90 in scenario 2. Together with the results from the task list phase across both sessions, the difference in the number of successfully employed gestures between the two gestures sets is minimal if not negligible. Comparing the effectiveness ratings for each task on the list it can be observed that between sessions, the percentage of gestures performed by the users with apparent difficulties ("medium" rat-

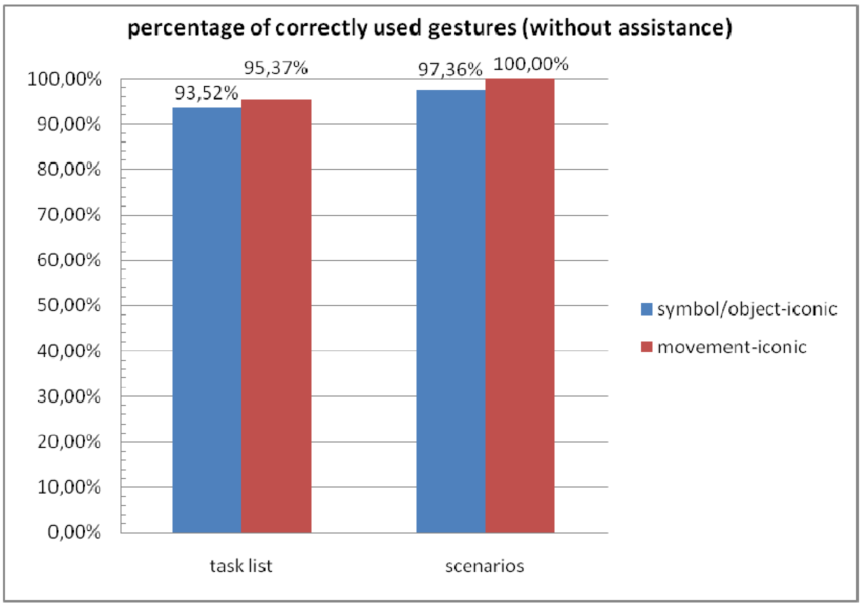


Figure 19. Percentage of correctly performed gestures during the task list and scenario phases.

ing) increased slightly for the symbol/object-iconic group but actually decreased for the movement-iconic group. Likewise, the percentage of gestures that were performed correctly on first try ("easy" rating) turned out to be higher in session 2 but only for users in the movement-iconic gestures group.

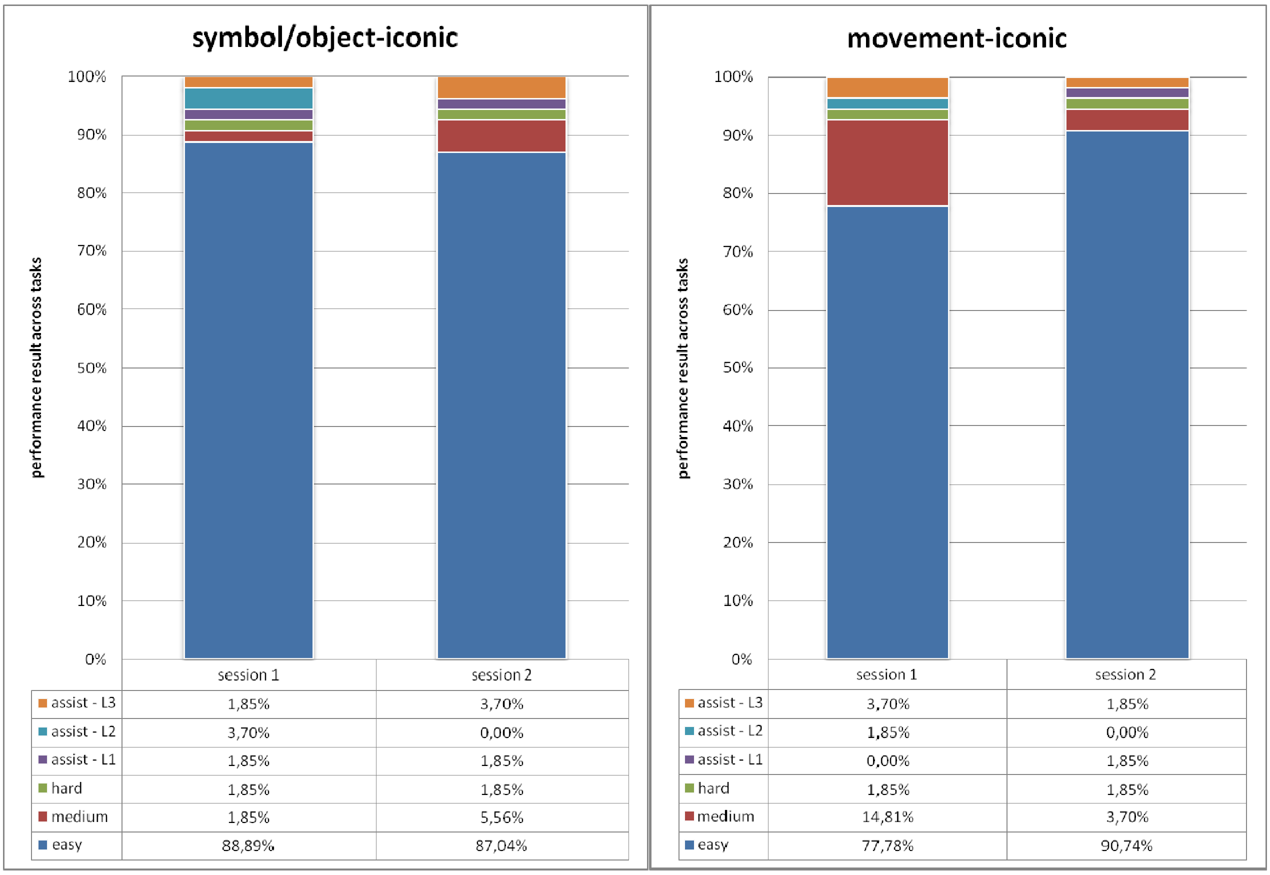


Figure 20. Differences in effectiveness ratings for task list (session 1/2): a) symbol/object-iconic gestures b) movement-iconic gestures

A series of mixed-model repeated-measures analyses of variance was ran to check for significance of the effects of time of measure (session) within participants and of gesture type (set) between participants. Concerning the number of correctly recalled gestures in the prompting interviews, the main effect for gesture type was not significant $F(1,10) = .034$, $p > .05$. Thus, there was no overall difference for participants in the symbol/object-iconic gesture group (mean number of recalled gestures per interview: 7,17 / 8) compared to participants in the movement-iconic gesture group (7,08 / 8). No significant main effect for time of measurement was obtained, $F(1,10) = 2,58$, $p > .05$.

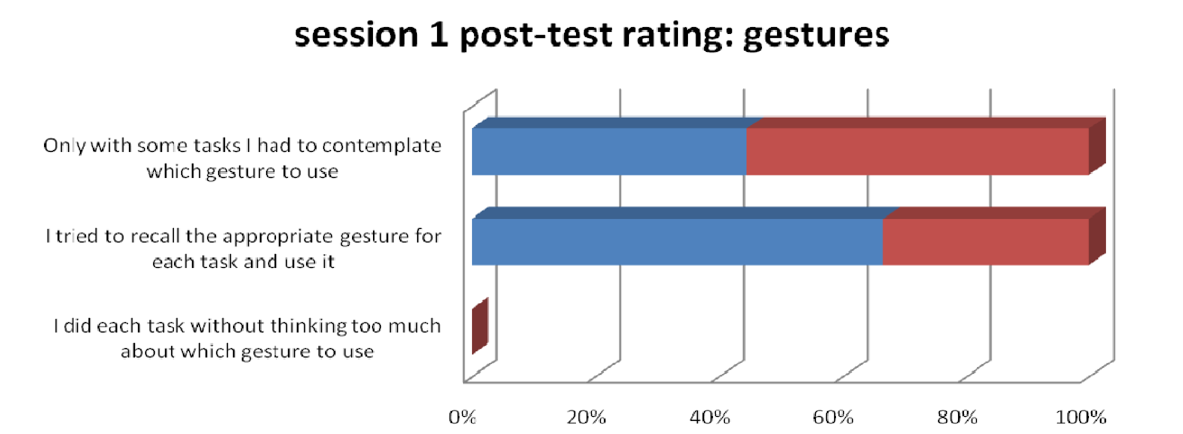
The mean number of recalled gestures per interview was not significantly higher for session 1 (7,5 / 8) compared to session 2 (6,75 / 8). No significant session x gesture type effect could be obtained, $F(1,10) = .032$, $p > .05$. A main effect was neither found for gesture type, $F(1,10) = .345$, $p > .05$, nor time of measurement, $F(1,10) = .588$, $p > .05$, regarding the number of correctly performed gestures in the task list phase. The same holds true for any interaction effect of the two factors.

When looking at the time each participant spent at completing the task list, a significant main effect could not be found neither for the between-subjects, $F(1,9) = .508$, $p > .05$, nor the within-subjects factor, $F(1,9) = 2.01$, $p > .05$, and an interaction effect did not emerge, $F(1,9) = .122$, $p > .05$. Each participant in the symbol/object-iconic gesture group spent a mean time of 5m 45s to complete the task list in session 1 and 5m 3s in session 2. Participants using the movement-iconic gestures took an average 5m 14s to finish the task list in session 1 and 4m 48s in session 2.

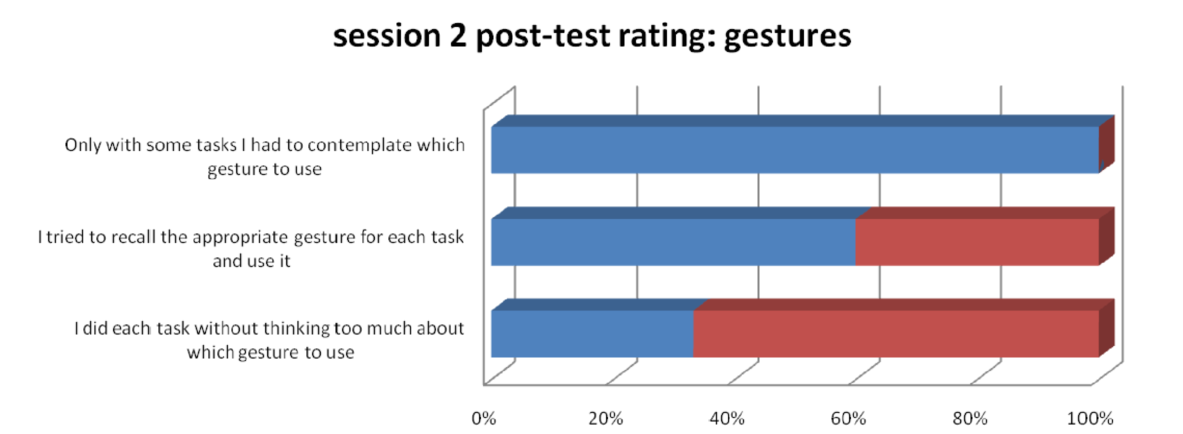
Analysis of completion time for the scenarios paint a similar picture, although due to their less sequential character, the data are to be interpreted more carefully. An independent samples t-test showed no significant difference in completion time between the two groups for scenario 1, $t(10) = .767$, $p = .46$, M (symbol/object-iconic group) = 00:04:03, M (movement-iconic group) = 00:03:24. For scenario 2, a significant difference emerged between the two groups $t(10) = 3.31$, $p = .01$. That is, the average completion time of participants using movement-iconic gestures (M = 04:16, SD = 00:33) was significantly different from that of the symbol/object-iconic group (M = 06:36, SD = 01:27). The number of correctly used gestures did not differ across groups neither for scenario 1, $t(5) = -1.59$, $p = .18$, nor scenario 2, $t(10) = -1$, $p = .34$.

As a further validation of the data, potential effects of learning order were checked with a further mixed-model, repeated-measures analysis of variance showing no significant effect for the sequence of gesture demonstration counterbalanced in each group, $F(1,5) = 1.2$, $p > .05$, and no Type of Gesture x Sequence effect, $F(1,5) = 1.07$, $p > .05$.

Results from the questionnaire items on ease-of-learning show that by session 2, participants reported being quite comfortable in using the gestures and that they did not have to reflect too much during the interaction. Especially users in the movement-iconic gestures group ticked the respective answer in 66,67% of cases while only 33,33% of users from the other group indicated that they “did each task without thinking too much about which gesture to use” (see Figure 21). Mean ratings on a 7-point Likert scale concerning various aspects of the gestures and the interaction are presented in Figures Figure 22 and Figure 23 and in Appendix F.



	I did each task without thinking too much about which gesture to use	I tried to recall the appropriate gesture for each task and use it	Only with some tasks I had to contemplate which gesture to use
G1 (symbol/object-iconic)	0,00%	33,33%	66,67%
G2 (movement-iconic)	0,00%	16,67%	83,33%



	I did each task without thinking too much about which gesture to use	I tried to recall the appropriate gesture for each task and use it	Only with some tasks I had to contemplate which gesture to use
G1 (symbol/object-iconic)	33,33%	50,00%	16,67%
G2 (movement-iconic)	66,67%	33,33%	0,00%

Figure 21. Questionnaire item regarding approach to using the gestures. Percentages indicating number of users that indicated respective answer.

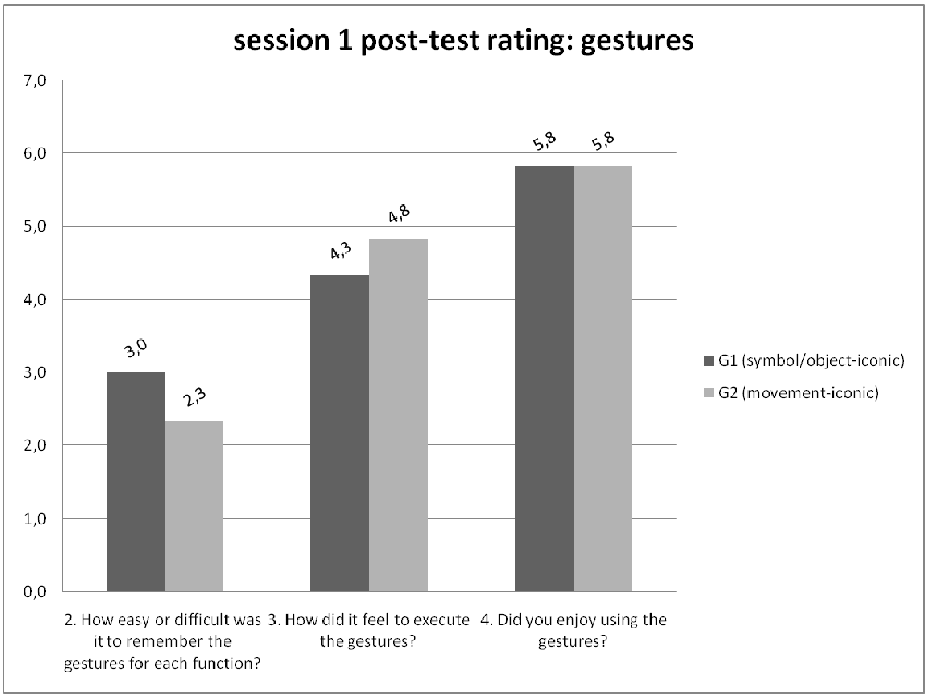


Figure 22. Mean ratings for the two groups of gestures. 2. memorability (1=very easy, 7=very difficult) 3.feeling (1=not at all natural, 7= very natural), 4.joy-of-use

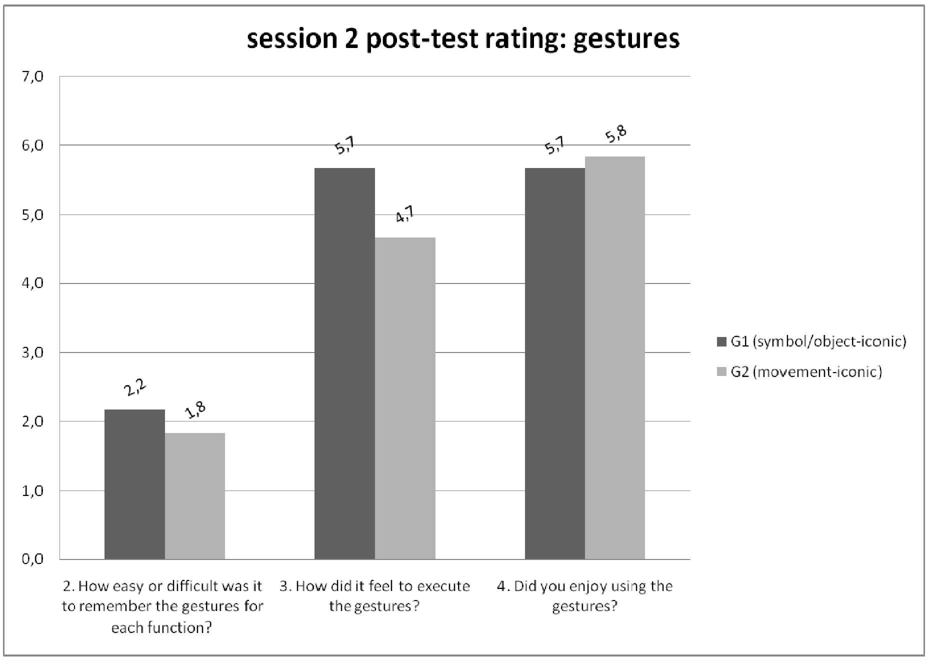


Figure 23. Mean ratings for the two groups of gestures. 2. memorability (1=very easy, 7=very difficult) 3.feeling (1=not at all natural, 7= very natural), 4.joy-of-use

6.4 Discussion of the user test

In reviewing the results from the usability test, a first general observation is the general success of the users in learning the eight gestures for the respective basic functions and their retention over a one-week time lag. Learnability, as reflected by the percentage of gestures users were able to correctly employ (93,52%) and explicitly recall (93,75%) after a short demonstration and training phase, appears indeed quite remarkable. Given that even after seven days of non-use, participants were able to explicitly recall 84,38% of the gestures and correctly used them in 95,37% (actually an improvement if minuscule) of cases during the task list, the gestures chosen for the current interaction purpose seem to be highly effective concerning their learnability and memorability. In addition, users regarded the gestures as easy to learn (score between 1,8 and 3,0 on a scale from 1, "very easy", to 7, "very difficult") in the questionnaire and rated ease-of-learning on average between 5,8 and 6,3 (on a scale from 1, "very difficult- " to 7, "very easy to learn in the given timeframe"). The implications provided by the data are presented with regard to the hypotheses that were postulated prior to testing.

- [h1]: *There will be only slight differences between the overall number of successfully recalled gestures at the end of session 1 and the beginning of session 2. This outcome is also expected for the number of correctly used gestures during the task list phase.*

Both assumptions are supported by the results as the number of correctly recalled gestures decreased only by four and five cases for the respective gesture group when comparing the overall number from the prompting interviews conducted in session 1 and session 2. Curiously however, the number of correctly executed gestures actually increased, if slightly, from 101 overall to 103 for the task list phase. These outcomes suggest the potential of gestures as an input means that is easy to learn and suited for occasional use.

- [h2]: *In general, more movement-iconic gestures than symbol/object-iconic gestures will be recalled during the prompting interviews.*

Hypothesis 2 regarded the difference between the two gesture sets in terms of how well users were able to retain the comprised gestures over a period of one week. No differences much less significant effects were found in the data from the prompting interviews. In fact, users in the symbol/object-iconic gestures group and the movement-iconic gestures group successfully recalled an overall 41 and 40 gestures, respectively. No further conjecturing was hence done for the underpinnings of a potential advantage of the movement-iconic gestures in terms of memorability.

- [h3] *More movement-iconic gestures than symbol/object-iconic gestures will be performed correctly in the task list phase of both sessions and across scenarios.*

The data from the task list phase indicate initial support for this hypothesis with movement-iconic gestures being performed at 95,37% correctly in the task list and at a perfect 100% in the scenarios. As no significant difference emerged concerning the overall numbers of correctly executed gestures in either the task list phase or the scenarios, no confirmation can be established for the hypothesis.

Furthermore, the data revealed no significant main effects or interaction effects for the factor of gesture type or session in any of the measures. While this could simply be due to similar learnability and memorability characteristics of both gesture types, a post-hoc analysis showed the low achieved power (0.07, two-tailed) in the t-test for finding the low effect of gesture type on the number of correctly used gestures in the task list phase of session 1 ($d = .24$) and session 2 ($d = .32$). Since increasing the number of participants for the usability test across multiple sessions might be difficult to achieve in terms of time and resources, revisiting the measurement approach could lead to more reliable instruments for assessing indicators of learnability and memorability for gestures used as a means of input.

6.4.1 Recognition accuracy of the system

Assessing overall usability in terms of efficiency indicated by completion times and joy-of-use resulted in more difficult to interpret data since the amount of time spent on the task list or scenarios can only be analyzed in relation to results obtained in a different session or by a different participant. It can be observed however, that although the overall steps required in scenario 2 were twice as many as in scenario 1, participants using the movement-iconic gestures required only a few seconds more on average than users from the other group for scenario 1.

This finding though needs to be regarded with caution as data for completion time for scenario 2 was missing in three cases due to technical reasons. Likewise, joy-of-use ratings need to be interpreted with prudence as do all other items in the question-

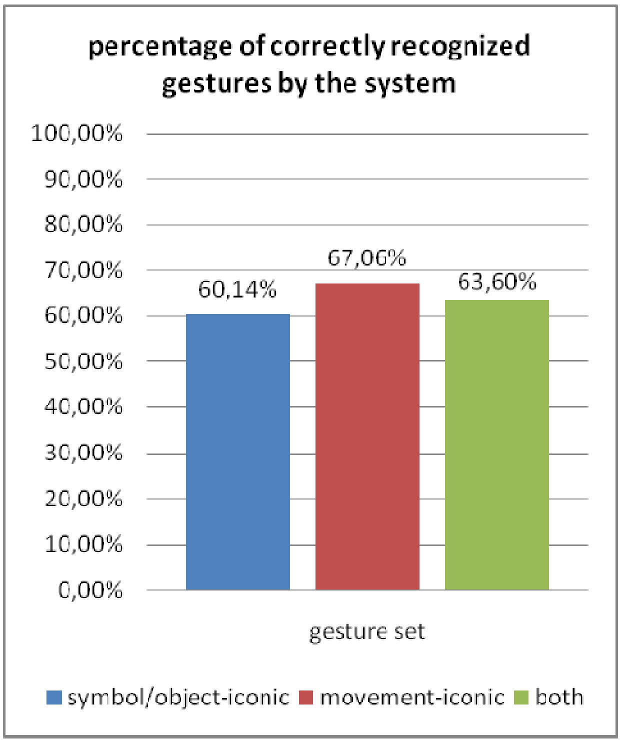


Figure 24. Average recognition accuracy of the algorithm introduced by Schlömer et al. (2008) and incorporated in the current system enabling interaction with the POWERWALL.

naire since the subjective judgment of all aspects of the gestures and the interaction are influenced by the system performance. Especially recognition accuracy can potentially entail a large influence since user frustration is likely to increase with every correctly performed gesture that is wrongly or not recognized by the system. An analysis of the recognition performance was thus performed to assess this potential influence with results showing that unlike reported by Schlömer et al. (2008) who obtained a minimum overall accuracy of 85% with a set of five gestures, the incorporated algorithm in the current system arrived only at an overall recognition rate of 60% (symbol/object-iconic gesture set) and 67% (movement-iconic gesture set). While this was largely unexpected after initial recognition accuracy results of 88% for eight (movement-iconic) gestures during pretesting, it should less influence the learnability and memorability measures than the subjective rating by the participants. Nevertheless, an impact on the number of successfully recalled and correctly used gestures might be due to differences in users' exposure, that is, through the number of actually executed gestures. For instance, although participants were instructed to continue to the next task whenever the system failed to recognize a correctly performed gesture either during the task list or scenario phase, the fact that some users did physically perform the respective gestures more than once might have led to effect of increased learning through practice, which unlike the exposure during the demonstration phase could not be counterbalanced. Since no significant difference in recognition rate was found between the two gesture sets, such an effect was assumed non-existent although it cannot be ruled out that differences might have emerged in some of the measures in case of a higher recognition accuracy by the system. On the other hand however, the relatively low recognition accuracy does indeed have an influence on the completion times for the task list and scenario since every gesture trial by the user extends the time spent on the tasks. Apart from the general concerns regarding completion times as a measure in usability testing (as mentioned e.g. in Stasche, 2005), the respective results presented above need to be cautiously interpreted given the recognition performance exhibited by the current state of the system. In addition, subjective ratings and opinion as measured by the post-test questionnaire are prone to be vulnerable to system performance in general and in the current context of gesture interaction, due to potentially increased frustration on the part of the user.

Despite these limitations regarding the interpretation of some of the data obtained in the usability test, the decision to go ahead with the implementation of a working gesture recognition system on the basis of an open source algorithm (wiigee.sourceforge.net) by Schlömer et al. (2008) is justified by the added value of evaluating gesture interaction and the role of a collection of gestures for it inside a real usage scenario involving a software application running on a large, high-resolution display being controlled by a motion-sensing laserpointer capable of both direct pointing and gesture input through HMM-based recognition. The groundwork is hence laid to orientate the continued development on all aspects of the interaction as outlined by the different categories in the framework for design and research of gestural interactions introduced by Karam (2006).

A final observation showed that user opinion on joy-of-use (average rating of 5,8 out of 7) and "naturalness" (average rating between 4,3 and 5,7) of the gesture execution was generally positive given the prototype status of the gesture recognition system and input device. Together with the very low support requirements (gestures had to be demonstrated to participants again only in three cases during each of the task list phases and only once during the scenarios) these results paint an already highly positive picture of adopting gesture interaction for use contexts involving large displays. The low need for assistance furthermore underlines how easy it was for the users to remember the intended gestures without any visual aid or reminder.

6.5 Conclusion of the user test

The results presented thus far support the general claim that using gestures enabled by motion-sensing handheld input devices as a means of input for HCI purposes is a feasible alternative to standard mouse/keyboard input. Given the fact that the device used in the current study allows two complementary modes of interaction that tap on the familiarity of using the hand for communication purposes, the question remained just how well gesture commands can be learned and memorized and which features of the gestures themselves facilitated the exploitation of this familiarity. Although no statistical significance could be observed for the data obtained in the user tests, the overall number of gestures that could be recalled explicitly by the participants when prompted was very high and users correctly executed the required gestures in over 93% of cases. The near-perfect use of the gesture commands in the scenarios confirms the application of gesture interaction to the mind mapping use case on a large high-resolution display. Participants generally enjoyed interacting with the display and were able to fluently execute gesture commands to create and adjust mind maps.

7 General discussion

In the line of investigations for this study, certain gestures were proposed during user interviews and then scrutinized as to their suitability of being used as a means of communicating commands to a computer inside an electronic mind mapping context. If one is to make a statement about the generalizability of the mappings of particular gestures to the commands applied in the current context of use, it is crucial to at least outline the underlying agreement of each of these commands with the characteristics of general operations found for most computer interfaces. As Jeff Raskin (2000) puts it prior to distinguishing between sets of “elementary actions” and “elementary operations”: “[...] interfaces for various applications are not as different as they might seem [...]. Applications seem more different than they are because you are attending to the content...” (2000, p.101). According to the author, the actions that a user can perform using the input device and interface are relatively limited. Keyboards afford key taps and longer pressing and graphical input devices such as mice, touchpads, trackballs serve to move a cursor on the display with the main button signalling the location of pointing.

In the current context, the direct pointing using a laserpointer can be considered characteristic of graphical input devices and a set of elementary actions would only have to be defined for the gesture modality. Viewed in line with the approach used by Raskin (2000), users can move the laserpointer freely in three dimensions and indicate distinct movement episodes by pressing a button. These units of gestures can serve to signal commands useful in the context.

The elementary operations which, as described by Raskin (2000, p.104), are applied on content and are represented in the set of commands used in the context of the current study using a mind mapping application, are the following:

- Generate: create new content (“modified from empty to nonempty”)
    - *Command: “add new sibling topic”*
- Delete: remove content (“modified from nonempty to empty”)
    - *Command: “remove selected topic”*
- Copy: send and receive of content (“duplicated at a different internal location”)
    - *Command: “copy”/ “paste”*

In addition, the navigational operation of zooming is mapped onto two commands:

- Zooming: change viewpoint on content (“overall metaphor [...] of flying”)
    - *Command: “zoom-in”*
    - *Command: “zoom-out”*

The results from the prompting interviews as well as the task list and scenario phases provide several first insights on which gesture type may be more suited to represent a certain class of elementary operations. Although the differences are low it can be observed that for navigation commands (in this case the “zoom” pair), the respective gestures were both recalled better during the prompting interviews and used correctly more often during the task list and scenario phases by users in the movement-iconic gestures group.

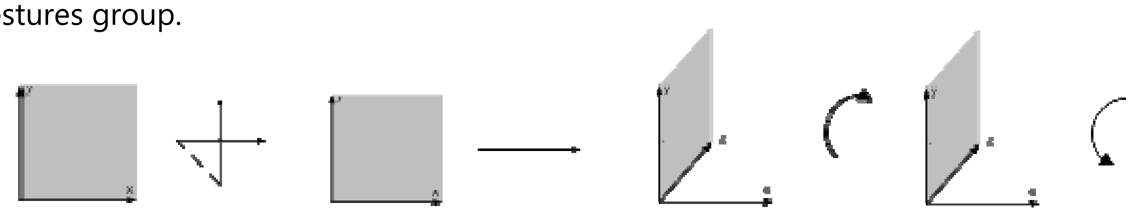


Figure 25. Symbol/object-iconic gestures (1&2) and movement-iconic gestures (3&4) for "zooming" operations.

In addition, the gesture for “delete” was recalled successfully more often during the prompting interviews by users in the symbol/object-iconic group. Referring to the particular gestures that were at the base of this outcome, it remains to be investigated which underlying causes led to these differences and on which level of analysis they are to be found. For example, despite presenting the commonly known symbols ‘PLUS’ and ‘MINUS’, the symbol/object-iconic gestures appeared to be less easily connected to the zoom-in and zoom-out functions than their movement-iconic counterparts. On the other hand, the ‘X’ symbol as found in many other software applications representing delete-operations was more readily recalled explicitly than the respective movement-iconic gesture.

Figure 26. Symbol/object-iconic gesture (left) and movement-iconic gesture (right) for "delete" operation

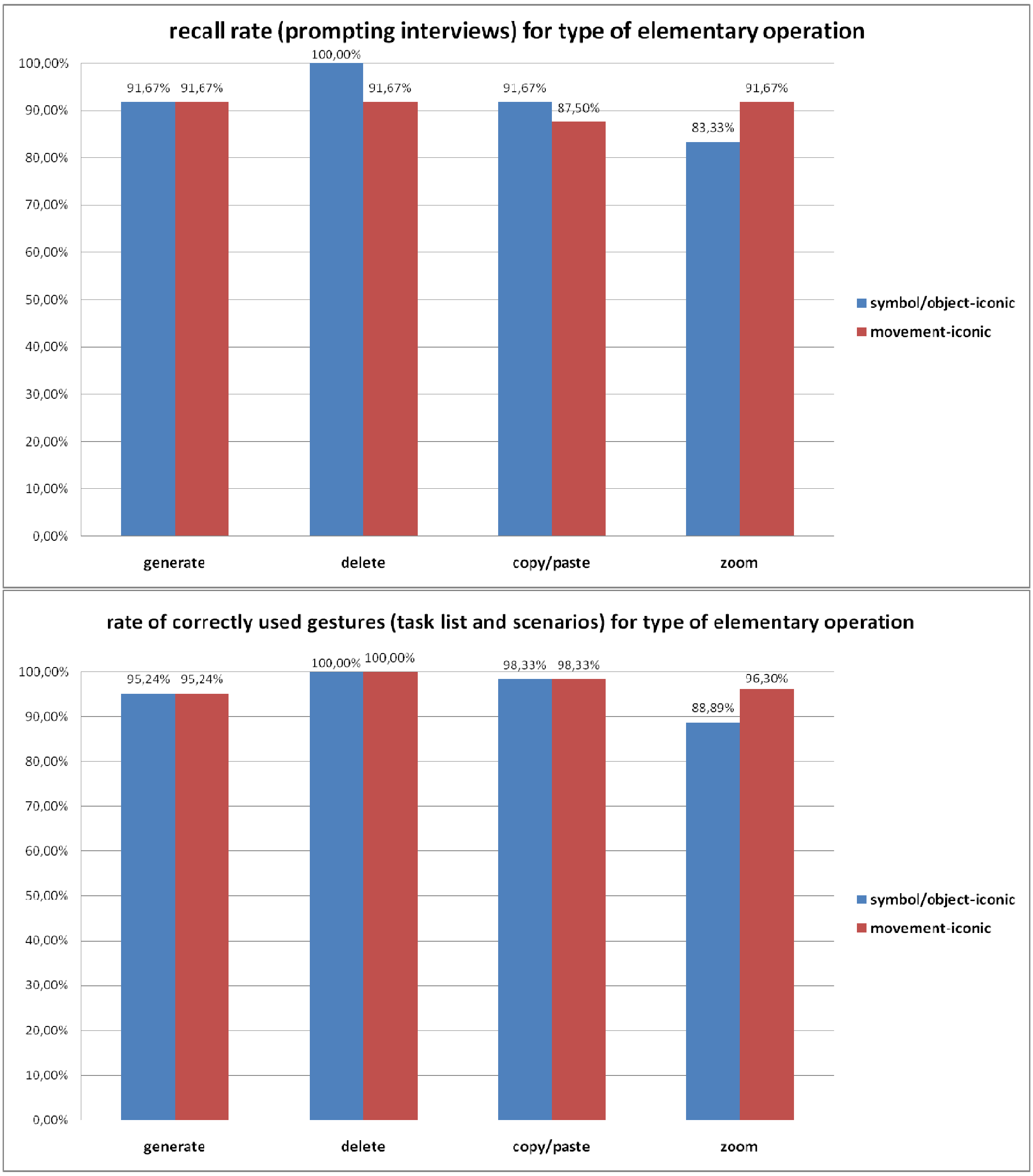


Figure 27. Rate of explicitly recalled gestures (above) and correctly used gestures (below) for elementary operations found in the mind mapping application.

While these results provide an initial clue as to the appropriateness of each of the present gesture types for the elementary operations involving deletion and navigation by zooming, it remains to be tested if such results can be obtained in the context of different on-screen interfaces and usage scenario. Furthermore, the collection of gestures obtained during the pre-study user interviews contains many more gesture proposals for the program functions involving elementary operations. These forms of interviews could hence be conducted with the additional aim of discovering feasible gestures for a number of elementary operations by having participants elicit gesture ideas in other application contexts and analyzing all instances according to their degree of agreement across the different types of operations.

As a general remark on the findings in the pre-study and usability test, it must be noted that the process of obtaining gesture proposals and scrutinizing them according to particular features strongly benefits from the continuous reference to the level of analysis. Especially the steps for arriving at distinct sets of gestures require that changes in the level must be indicated to allow the justification for making reference to existing theories. Concerning the external validity of the pre-study, the differences in the interview context, when compared to the intended context of use in front of a large high-resolution display, must be taken into account as they may limit the degree to which the obtained gestures qualify for transfer to an evaluation scenario. Likewise, a usability test of gesture input should reduce potential sources of user frustration as much as possible when the actual gestures *per se* or the interaction technique are under focus. The fact that the results from the current usability evaluation reveal a largely favourable attitude towards the use of gestures as a complementary input means for large display interaction indicates how much the gestures as such were suited for this purpose despite the lower-than-expected system performance for recognition. Adjustments to the testing method that should increase internal validity might involve relying on more objective data such as system logs of users' gesture execution, devising a standardized questionnaire and establish suitable constructs for gesture attributes, dividing the scoring of user performance among two supervisors and increasing the amount of information that can be extracted from video recordings.

8 General conclusion and outlook

The research endeavour presented in this Master's thesis began with few pretences to the eventual impact on both the body of knowledge concerning human gestures and their use as an input modality for human-computer interaction purposes. Given the lack of a unified terminology or conceptual framework it was necessary for the current line of investigation to conduct a large amount of groundwork prior to offering a potential path for transferring insights from various disciplines into a framework for the design and research of gesture-based interaction introduced by Maria Karam (2006). Considering that in the HCI field an interdisciplinary character of the research is often the rule, elements from psychology, linguistics and anthropology turned out to be rare in this existing framework which led the author to adopt the initial goal of contributing to the theoretical background that informs the advice on design and research provided by the various categories proposed by Karam. Similar to the subsequent pre-study conducted for obtaining gesture ideas in the context of an electronic mind mapping session and employing a custom-built laserpointer with both motion-sensing and direct pointing capabilities, researchers and developers could scrutinize gestures realized by their specific input device solution and usage contexts according to the parameters proposed in this thesis or likewise extend the respective categories as has been demonstrated above. Furthermore, the current work then presented a first attempt of integrating a reconciled approach to gesture interaction research into a standardized formal evaluation of the the overall usability and learnability of a collection of gestures and their use as a means of controlling a software application on a large, high-resolution display. In view of the varied strategies by researchers in the fields of HCI, ubiquitous computing, machine learning, robotics and computer graphics for investigating gestures and their application, it is the authors thorough conviction that the utilization and further elaboration of a framework is a fruitful activity that will contribute to the understanding of the value of the human gestural modality for the communication and interaction with machines, computers, and interfaces. The results from the evaluation presented in this thesis provide a positive indication of this value for using handheld device gestures as a means of control and more research will undoubtedly promise valid support for establishing the use of gestures as an input modality for human-computer interaction. Apart from investigations into the nature of the gestures employed, future work could study the individual mental strategies that users rely on when adopting gestures as representations of specific interactions. In a similar vein, researchers should look into the role of visual aids or feedback in this learning process. It may for example be interesting to explore how gestures could be demonstrated coherently to users in the initial learning phase by displaying an image or animation and how exactly the gestures' dimensions could be represented. Furthermore, the interaction using gestures as introduced in the current work may exhibit different characteristics concerning its suitability if the context of use comprises a varying on-screen interface, functionality and environmental features such as screen size, resolution and distance as well as group configurations and individual computer experience.

9 References

Ballagas, R., Borchers, J., Rohs, M., & Sheridan, J. G. (2006). The Smart Phone: A Ubiquitous Input Device. *Pervasive Computing, IEEE, 5*(1), 70-77.

Barnum, C. M., & Dragga, S. (2001). *Usability Testing and Research*: Allyn & Bacon, Inc. Needham Heights, MA, USA.

Belgardt, S., Schwan, N., & Reichardt, D. M. Towards a Presentation Mirror: First Steps in Using the BlueWand Technology for Gesture Analysis. *ITCS 2005*.

Birnholtz, J. P., Grossman, T., Mak, C., & Balakrishnan, R. (2007). An exploratory study of input configuration and group process in a negotiation task using a large display. *Proceedings of the SIGCHI conference on Human Factors in computing systems*, 91-100.

Bolt, R. A. (1980). “Put-that-there”: Voice and gesture at the graphics interface. *Proceedings of the 7th annual conference on Computer graphics and interactive techniques*, 262-270.

Boren, T., & Ramey, J. (2000). Thinking aloud: reconciling theory and practice. *Professional Communication, IEEE Transactions on, 43*(3), 261-278.

Brewster, S., Lumsden, J., Bell, M., Hall, M., & Tasker, S. (2003). Multimodal'eyes-free'interaction techniques for wearable devices. *Proceedings of the conference on Human factors in computing systems*, 473-480.

Buzan, T., & Buzan, B. (1996). *The Mind Map Book: How to Use Radiant Thinking to Maximize Your Brain's Untapped Potential*. Plume.

Cadoz, C. (1994). Le geste canal de communication homme/machine: la communication «instrumentale». *TSI. Technique et science informatiques, 13*(1), 31-61.

Cao, X., & Balakrishnan, R. (2003). VisionWand: interaction techniques for large displays using a passive wand tracked in 3D. *Proceedings of the 16th annual ACM symposium on User interface software and technology*, 173-182.

Cao, X., & Zhai, S. (2007). Modeling human performance of pen stroke gestures. *Proceedings of the SIGCHI conference on Human Factors in computing systems*, 1495-1504.

Chen, X., & Davis, J. (2002). LumiPoint: Multi-User Laser-Based Interaction on Large Tiled Displays. *Displays, 23*(5), 205-211.

Cosnier, J. (1977). Communication non verbale et langage. *Psychologie Médicale, 9*(11), 2033-2049.

Efron, D. (1941). Gesture and Environment. *New York: King's Crown*.

Efron, D. (1972). *Gesture, race and culture; a tentative study of the spatio-temporal and" linguistic" aspects of the gestural behavior of eastern Jews and southern Italians in New York City, living under similar as well as different environmental conditions*. The Hague: Mouton.

Ekman, P., & Friesen, W. V. (1969). The repertoire of nonverbal behavior. *Semiotica, 1*(49-98).

Fitzmaurice, G. W., Ishii, H., & Buxton, W. A. S. (1995). Bricks: laying the foundations for graspable user interfaces. *Proceedings of the SIGCHI conference on Human Factors in computing systems*, 442-449.

Foehrenbach, S., König, W. A., Gerken, J., & Reiterer, H. (2008). *Natural Interaction with Hand Gestures and Tactile Feedback for large, high-res Displays*. Paper presented at the MITH'08: Workshop on Multimodal Interaction Through Haptic Feedback; held in conjunction with AVI'08: International Working Conference on Advanced Visual Interfaces.

Gardener, B., & Gardener, P. (1969). Teaching sign language to a chimpanzee. *Science, 165*, 664-672.

Gentile, A. M. (1998). Movement Science: Implicit and Explicit Processes during Acquisition of Functional Skills. *Scandinavian Journal of Occupational Therapy, 5*(1), 7-16.

Hewes, G. W. (1973). Primate Communication and the Gestural Origin of Language. *Current Anthropology, 14*(1/2), 5.

Hilliges, O., Terrenghi, L., Boring, S., Kim, D., Richter, H., & Butz, A. (2007). Designing for collaborative creative problem solving. *Proceedings of the 6th ACM SIGCHI conference on Creativity & cognition*, 137-146.

Huang, E. M. (2006). *The Design and Analysis of Large Display Groupware Applications.* Unpublished PhD thesis, Georgia Institute of Technology.

International Organization for Standardization. (1998). ISO 9241: Ergonomic requirements for office work with visual display terminals (VDTs) - *Part 11: Guidance on usability*.

International Organization for Standardization. (2002). ISO/TR 16982: Ergonomics of human-system interaction - Usability methods supporting human-centered design.

iPhone. (2007). [Mobile phone]: Apple.

Karam, M. (2006). *A framework for research and design of gesture-based human computer interactions.* Unpublished PhD thesis, University of Southampton.

Karam, M., & schraefel, M. C. (2005). A Taxonomy of Gestures in Human Computer Interaction. *ACM Transactions on Computer-Human Interactions*.

Kela, J., Korpipää, P., Mäntyjärvi, J., Kallio, S., Savino, G., & Jozzo, L. (2006). Accelerometer-based gesture control for a design environment. *Personal and Ubiquitous Computing, 10*(5), 285-299.

Kendon, A. (2004). *Gesture: Visible Action as Utterance*: Cambridge University Press.

Kettebekov, S., & Sharma, R. (2000). Understanding Gestures in Multimodal Human Computer Interaction. *International Journal on Artificial Intelligence Tools, 9*(2), 205-223.

König, W. A., Bieg, H. J., & Reiterer, H. (2007). Laserpointer-Interaktion für große, hochauflösende Displays. In T. Gross (Ed.), *Mensch & Computer 2007* (pp. 69-78). München: Oldenbourg.

König, W. A., Bieg, H. J., Schmidt, T., & Reiterer, H. (2007). Position-independent interaction for large high-resolution displays. *Proc. IHCI 2007*, 117-125.

König, W. A., Böttger, J., Völzow, N., & Reiterer, H. (2008). *Laserpointer-Interaction between Art and Science*. Paper presented at the IUI'08: Proceedings of the 13th international conference on Intelligent User Interfaces.

Kratz, L., Smith, M., & Lee, F. J. (2007). Wiizards: 3D gesture recognition for game play input. *Proceedings of the 2007 conference on Future Play*, 209-212.

Kwon, D. Y., Wuermlin, S., & Gross, M. (2007). *mCube - Towards a Versatile Gesture Input Device for Ubiquitous Computing Environments.* Paper presented at the UCS.

Latoschik, M. E., & Wachsmuth, I. (1998). Exploiting distant pointing gestures for object selection in a virtual environment. *Gesture and Sign Language in Human-Computer Interaction*, 185-196.

Lingrand, D., Meunier, S., Renevier, P., Pinna-Dery, A. M., Soula, J., Riveill, M., et al. (2006). Gestaction 3D: combiner commandes gestuelles et interactions 3D.

Mäntyjärvi, J., Kela, J., Korpipää, P., & Kallio, S. (2004). Enabling fast and effortless customisation in accelerometer based gesture interaction. *Proceedings of the 3rd international conference on Mobile and ubiquitous multimedia*, 25-31.

Mäntylä, V. M., Mäntyjärvi, J., Seppänen, T., & Tuulari, E. (2000). Hand gesture recognition of a mobile device user. *Multimedia and Expo, 2000. ICME 2000. 2000 IEEE International Conference.*

Martell, C. (2005). *FORM: An Experiment in the Annotation of the Kinematics of Gesture.* Unpublished PhD thesis, University of Pennsylvania.

McNeill, D. (2005). *Gesture and Thought*: University Of Chicago Press.

McNeill, D. (2007). Gesture and Thought. In A. Esposito (Ed.), *Fundamentals of verbal and nonverbal communication and the biometric issue. Proceedings of the NATO Advanced Study Institute on the Fundamentals of Verbal and Nonverbal Communication and the Biometric Issue (2006 : Vietri sul Mare, Italy).* Amsterdam: IOS.

MindManager. (2007). [Software]: MindJet.

Mozilla Firefox. (2007). [Software].

N95, N80. (2006). [Mobile phone]: Nokia.

Neale, D. C., Carroll, J. M., & Rosson, M. B. (2004). Evaluating computer-supported cooperative work: models and frameworks. *Proceedings of the 2004 ACM conference on Computer supported cooperative work*, 112-121.

Nehaniv, C. L., Dautenhahn, K., Kubacki, J., Haegele, M., Parlitz, C., & Alami, R. (2005). A methodological approach relating the classification of gesture to identification of human intent in the context of human-robot interaction. *Robot and Human Interactive Communication, 2005. ROMAN 2005. IEEE International Workshop on*, 371-377.

Nespoulous, J. L., & Lecours, A. R. (1986). Gestures: Nature and Function. In J. L. Nespoulous, P. Perron & A. R. Lecours (Eds.), *The Biological Foundations of Gestures: Motor and Semiotic Aspects*. Hillsdale, NJ: Lawrence Erlbaum Associates.

Nintendo. (2007). Wii controllers. Retrieved May 12, 2008, from http://www.nintendo.com/wii/what/controllers

Opera browser. (2007). [Software].

Patel, S. N., Pierce, J. S., & Abowd, G. D. (2004). A gesture-based authentication scheme for untrusted public terminals. *Proceedings of the 17th annual ACM symposium on User interface software and technology*, 157-160.

Payne, J., Keir, P., Elgoyhen, J., McLundie, M., Naef, M., Horner, M., et al. (2006). Gameplay issues in the design of spatial 3D gestures for video games. *Conference on Human Factors in Computing Systems*, 1217-1222.

Pirhonen, A., Brewster, S., & Holguin, C. (2002). Gestural and audio metaphors as a means of control for mobile devices. *Proceedings of the SIGCHI conference on Human factors in computing systems: Changing our world, changing ourselves*, 291-298.

Quek, F., McNeill, D., Bryll, R., Duncan, S., Ma, X., Kirbas, C., et al. (2002). Multimodal Human Discourse: Gesture and Speech. *ACM Transactions on Computer-Human Interaction, 9*(3), 171-193.

Raskin, J. (2000). *The Humane Interface: New Directions for Designing Interactive Systems*: Addison-Wesley Professional.

Reese, H. W. (1997). Counterbalancing and Other Uses of Repeated-Measures Latin-Square Designs: Analyses and Interpretations. *Journal of Experimental Child Psychology, 64*(1), 137-158.

Rekimoto, J. (2001). GestureWrist and GesturePad: Unobtrusive Wearable Interaction Devices. *Proc. ISWC'01, Fifth Intl Symposium on Wearable Computers*, 87-94.

Roth, W. M. (2005). Talking Science: Language and Learning in Science Classrooms. *Rowman & Littlefield Publishers, Inc.*, 256.

Rubine, D. (1991). Specifying gestures by example. *Proceedings of the 18th annual conference on Computer graphics and interactive techniques*, 329-337.

Schlömer, T., Poppinga, B., Henze, N., & Boll, S. (2008). Gesture recognition with a Wii controller. *Proceedings of the 2nd international conference on Tangible and embedded interaction*, 11-14.

semiosis. (2008). *Merriam-Webster Online Dictionary.* Retrieved July 5, 2008, from http://www.merriam-webster.com/dictionary/semiosis

Smith, J., White, T., Dodge, C., Paradiso, J., Gershenfeld, N., & Allport, D. (1998). Electric field sensing for graphical interfaces. *Computer Graphics and Applications, IEEE, 18*(3), 54-60.

Sreedharan, S., Zurita, E. S., & Plimmer, B. (2007). 3D input for 3D worlds. *Proceedings of the 2007 conference of the computer-human interaction special interest group (CHISIG) of Australia on Computer-human interaction: design: activities, artifacts and environments*, 227-230.

Stasche, A. (2005). *Interference with task performance using the think-aloud technique in Usability Testing*. Unpublished BSc thesis, University of Osnabrueck.

Streitz, N. A., Geißler, J., Holmer, T., Müller-Tomfelde, C., Reischl, W., Rexroth, P., et al. (1999). i-LAND: an interactive landscape for creativity and innovation. *Proceedings of the SIGCHI conference on Human factors in computing systems: the CHI is the limit*, 120-127.

Tsukada, K., & Yasumura, M. (2002). Ubi-Finger: Gesture Input Device for Mobile Use. *Proceedings of APCHI, 1*, 388-400.

Tuulari, E., & Ylisaukko-oja, A. (2002). SoapBox: A Platform for Ubiquitous Computing Research and Applications. *Proceedings of the First International Conference on Pervasive Computing*, 125-138.

Vogt, F., Wong, J., Po, B. A., Argue, R., Fels, S. S., & Booth, K. S. (2004). Exploring collaboration with group pointer interaction. *Computer Graphics International, 2004. Proceedings*, 636-639.

Wagenaar, W. A. (1969). Note on the construction of digram-balanced Latin squares. *Psychological Bulletin, 72*(6), 384-386.

Wexelblat, A. (1995). An approach to natural gesture in virtual environments. *ACM Trans. Comput.-Hum. Interact., 2*(3), 179-200.

Wexelblat, A. (1998). Research Challenges in Gesture: Open Issues and Unsolved Problems. *Gesture and Sign Language in Human-Computer Interaction: International Gesture Workshop, Bielefeld, Germany, September 1997: Proceedings*.

Wilson, A., & Shafer, S. (2003). XWand: UI for intelligent spaces. *Proceedings of the conference on Human factors in computing systems*, 545-552.

Wobbrock, J. O., Wilson, A. D., & Li, Y. (2007). Gestures without libraries, toolkits or training: a $1 recognizer for user interface prototypes. *Proceedings of the 20th annual ACM symposium on User interface software and technology*, 159-168.

Wu, M., Shen, C., Ryall, K., Forlines, C., & Balakrishnan, R. (2006). Gesture Registration, Relaxation, and Reuse for Multi-Point Direct-Touch Surfaces. *Proceedings of IEEE TableTop-the International Workshop on Horizontal Interactive Human Computer Systems*, 183-190.

Wundt, W. (1904, 2. Aufl.). *Völkerpsychologie. T. 1, Die Sprache*. Leipzig: Engelmann.

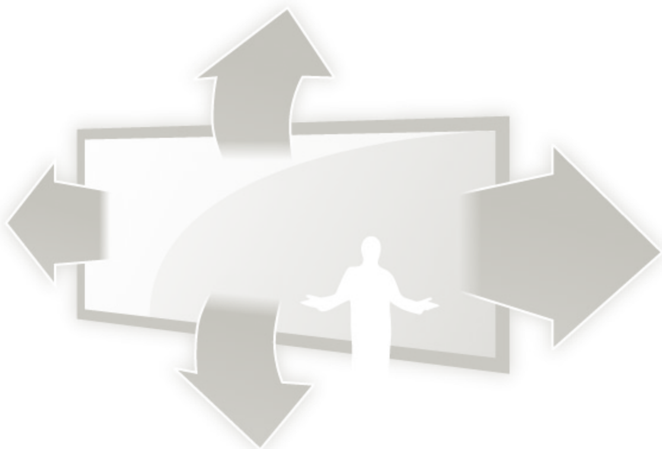
Wundt, W. M. (1912). *Elemente der Völkerpsychologie: Grundlinien einer psychologischen Entwicklungsgeschichte der Menschheit*. Leipzig: A. Kröner.

Zimmerman, T. G., Lanier, J., Blanchard, C., Bryson, S., & Harvill, Y. (1986). A hand gesture interface device. *Proceedings of the SIGCHI/GI conference on Human factors in computing systems and graphics interface*, 189-192.

10 Appendices

10.1 Appendix A

10.1.1 Interview materials (German)



INTERVIEW ZUR GESTENINTERAKTION

NUMMER ____

ALTER ____

GESCHLECHT ☐ männlich ☐ weiblich

ORT ____

HALBSTRUKTURIERTES INTERVIEW

Datum Durchführung __/__/__ Koordinator____________________ N. Nutzer_____

Es handelt sich um ein halbstrukturiertes Interview: der Interviewer richtet sich nach einer Liste von Informationen, die erfragt werden sollen, das Gespräch wird aber frei geführt. Der Interviewee kann sein Meinungen und Wünsche frei äußern ohne eine bestimmte Reihenfolge einzuhalten. Der Leiter überprüft, ob alle notwendigen Informationen erfasst wurden und kann zur Klarstellung nachfragen, um die wichtigsten Inhalte zu vertiefen.

Computererfahrung

Computernutzung
- *Kenntnisse von Eingabegeräten*
- *EDV-Kreativwerkzeuge*
- *Weitere Softwareerfahrung*

Kennen Sie ...?

Welche Programme nutzen Sie hauptsächlich? *Spiele ausgeschlossen*

Welche Funktionen sind für Sie am relevantesten? *Welche im selbigen Programm genutzt werden*

Warum genau? ____

____

____

BEWERTUNG VON SOFTWARE FÜR MIND MAPS

Was finden Sie praktisch/nützlich an diesem Programm?

____

____

____

____

ERGEBNIS:

☐ *Keinerlei Verständnis für das Programm oder das Konzept*

☐ *Bereitschaft die Hauptfunktionen auszuprobieren*

☐ *Entschlossen das Programm ernsthaft auszuprobieren*

☐ *Offensichtlich gewillt ein Projekt mithilfe des Programms anzugehen*

VERSTÄNDNIS DER GRUNDFUNKTIONEN

1. *Darstellungsfunktionen*
2. *Funktionen für die Erstellung von mind maps*

Interaktion mit den Befehlen:

- ***Darstellung:***
  - *Vergrößern* **STRG + '+'**
  - *Verkleinern* **STRG + '-'**
  - *Map teilen* **STRG + ALT + V**
  - *Map vereinen* **STRG + ALT + C**

Finden Sie Funktion X verständlich?

ERGEBNIS:

- ☐ *Allgemeine Verständnisprobleme*
- ☐ *Einzelne Verständnisschwierigkeiten bei den Funktionen aber hohe Wahrscheinlichkeit des Erlernens*
- ☐ *Gutes Verständnis der Konzepte mit wenigen Schwierigkeiten bei der entsprechenden Funktion*
- ☐ *Gutes allgemeines Verständnis der Konzepte und der entsprechenden Funktionen*

- ***Hinzufügen von neuen Objekten***
  Unterzweig hinzufügen **EING**
  Nebenzweig hinzufügen **EINFG**
  Verbindungslinie hinzufügeni **ALT + H + R**
  Umrandung hinzufügen **STRG + UMSCH + B**
  ausgewählten Zweig löschen **STRG + UMSCH + ENTF**
  Map ausgleichen **STRG + ALT + B**

Finden Sie Funktion X verständlich?

ERGEBNIS:

- ☐ *Allgemeine Verständnisprobleme*
- ☐ *Einzelne Verständnisschwierigkeiten bei den Funktionen aber hohe Wahrscheinlichkeit des Erlernens*
- ☐ *Gutes Verständnis der Konzepte mit wenigen Schwierigkeiten bei der entsprechenden Funktion*
- ☐ *Gutes allgemeines Verständnis der Konzepte und der entsprechenden Funktionen*

ZEICHNUNG GEEIGNETER GESTEN

Darstellungsfunktionen:

- *Vergrößern* ***STRG + '+'***
- *Verkleinern* ***STRG + '-'***
- *Map teilen* ***STRG + ALT + V***
- *Map vereinen* ***STRG + ALT + C***

Nun möchte ich Sie bitten sich Gesten zu überlegen, die Sie als geeignet ansehen, die in der Tabelle genannten Funktionen auszuführen. Die Gesten können zweierlei Art sein: zweidimensional und dreidimensional. Zweidimensionale Gesten werden wie auf eine unsichtbare Wand gezeichnet ausgeführt, die Linie kann sowohl gerade als auch kurvig sein und es können ebenso geometrische Figuren gezeichnet werden. Es ist somit möglich eine beliebige Form oder Linie in eine beliebige Richtung in der Plane zu zeichnen, die die entsprechende Funktion des Programms repräsentieren soll. Dennoch soll die Geste als eine Bewegung (während des Drückens des oberen Buttons) ausgeführt werden.
Für die dreidimensionalen Gesten können Sie sich noch mehr Freiheit nehmen, passende Gesten zu überlegen. Jede Bewegung die Ihnen als Ganzes natürlich und gewohnt erscheint kann somit eine der Funktionen des Programms repräsentieren. Auch hier soll die Geste während des Drückens des oberen Buttons ausgeführt werden. Sie können die Geste auch selber in die entsprechende Spalte der Tabelle eintragen.

Gestentyp/ Befehl	Gesten (zweidimensional)	Gesten (dreidimensional)
Vergrößern		

Verkleinern		
Map teilen		
Map vereinen		

Kommentare und Ideen

- *Durchführbarkeit*
- *Komplexität*
- *nötige Anstrengung im Gedächtnis zu behalten*
- *nötige Kraftanstrengung*

Wie empfanden Sie die Geste für den Befehl X ... ?

____

____

Wie könnte Sie verbessert werden?

____

____

____

____

____

____

Bearbeitungsfunktionen:

- *Unterzweig hinzufügen* ***EING***
- *Nebenzweig hinzufügen* ***EINFG***
- *Verbindungslinie hinzufügen* ***ALT + H + R***
- *Umrandung hinzufügen* ***STRG + UMSCH + B***
- *ausgewählten Zweig löschen* ***STRG + UMSCH + ENTF***
- *Map ausgleichen* ***STRG + ALT + B***

Gestentyp/ Befehl	Gesten (zweidimensional)	Gesten (dreidimensional)
Unterzweig hinzufügen		

Nebenzweig hinzufügen		
Verbind-ungslinie hinzufügen		

Gestentyp/ Befehl	Gesten (zweidimensional)	Gesten (dreidimensional)
Umrandung hinzufügen		

ausgewählt-en Zweig löschen		
Map ausgleichen		

Kommentare und Ideen

- *Durchführbarkeit*
- *Komplexität*
- *nötige Anstrengung im Gedächtnis zu behalten*
- *nötige Kraftanstrengung*

Wie empfanden Sie die Geste für den Befehl X … ?

Wie könnte Sie verbessert werden?

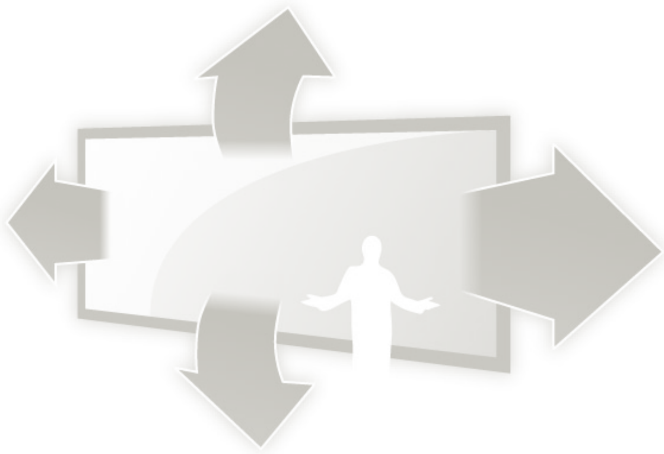
VORSCHLÄGE ZUM FEEDBACK ***(feedback: Gerät gibt eine Rückmeldung per Vibration)***

Auf welche Weise und zu welchem Zeitpunkt könnte das Vibrationsfeedback nützlich sein?

ERGEBNIS:

- ☐ *feedback könnte in keinerlei Weise nützlich sein*
- ☐ *feedback könnte den Modus des Eingabegerätes signalisieren*
- ☐ *feedback könnte nützlich sein, den Status der Gestenerkennung anzuzeigen*
- ☐ *feedback könnte für verschiedene Funktionen des Eingabegerätes nützlich sein*

10.1.2 Interview materials (Italian)



INTERVISTA

NUMERO ____

ETÀ ____

SESSO ☐ maschile ☐ femminile

PROVENIENZA ____

INTERVISTA SEMISTRUTTURATA

Data compilazione __/__/__ Coordinatore____________________ N. prog. utente_____

Si tratta di un'intervista semi-strutturata: l'intervistatore si avvale di una lista di informazioni da raccogliere, ma il colloquio viene gestito in maniera libera. L'intervistato può raccontare le sue opinioni e le sue aspirazioni senza un ordine preciso. L'operatore verifica di aver annotato tutte le informazioni necessarie e può porre domande di chiarimento, valutando quali contenuti meritano un maggiore approfondimento.

ESPERIENZE INFORMATICHE

Competenze dell' uso computer

- *conoscenza dei dispositivi*
- *programmi di progettazione creativa*
- *esperienze di altri programmi*

Lei conosce …?

____

____

____

Quale programma usa (Lei) principalmente? *I giochi sono esclusi*

____

____

Quali sono le funzioni più rilevanti per Lei? *Quello che si può fare con quel programma*

____

____

Perché? ____

____

____

____

VALUTAZIONE SOFTWARE PER LE MAPPE MENTALI

Cosa trova funzionale nel programma?

____

____

____

____

SINTESI:

☐ *Nessun apprezzamento del programma o del concetto*

☐ *Disposizione a provare le funzioni principali*

☐ *Buona disponibilità ad intraprendere una prova seria*

☐ *Volontà evidente di realizzare un progetto usando il programma*

COMPRENSIONE FUNZIONI DI BASE

1. *funzioni di visualizzazione*
2. *funzioni di modificazione delle mappe*

Interazione usando comandi:

- ***gestione visualizzazione***
  ingrandire **STRG + '+'**
  rimpicciolire **STRG + '-'**
  tagliare mappa **STRG + ALT + V**
  riunire mappa **STRG + ALT + C**

Lei trova comprensibile la funzione ...?

____

____

____

____

SINTESI:

☐ *Generale difficoltà di capire*

☐ *Alcune difficoltà capire le funzioni ma alta probabilità di imparare velocemente*

☐ *Buona comprensione dei concetti con poche difficoltà eseguire le funzione rispettive*

☐ *Buona comprensione generale dei concetti e delle rispettive funzioni*

- ***creazione di nuovi ogetti***
  inserire nuovo argomento principale **EING**
  inserire nuovo sottoargomento **EINFG**
  inserire relazioni rami **ALT + H + R**
  creare nuovo contorno **STRG + UMSCH + B**
  rimuovere selezionato **STRG + UMSCH + ENTF**
  bilanciare mappa **STRG + ALT + B**

Lei trova comprensibile la funzione ...?

____

____

____

SINTESI:

☐ *Generale difficoltà di capire i concetti alla base delle funzioni*

☐ *Alcune difficoltà capire le funzioni ma alta probabilità di imparare velocemente*

☐ *Buona comprensione dei concetti con poche difficoltà eseguire le funzioni rispettive*

☐ *Buona comprensione generale dei concetti e delle rispettive funzioni*

DISEGNARE GESTI FATTIBILI

Funzioni di visualizzazione:
- *ingrandire*
- *rimpicciolire*
- *tagliare mappa*
- *riunire mappa*

Adesso Le chiedo di pensare a dei gesti che ritiene appropriati per eseguire le funzioni indicate nelle colonne della tabella sottostante. I gesti possono essere di due tipi: bidimensionali e tridimensionali. Nel primo caso (gesti bidimensionali), immagini di tracciare una linea su una parete immaginaria, la linea può essere sia dritta che curva (es. linea singola dall'alto verso il basso, forme geometriche, etc.). È possibile tracciare qualsiasi forma in qualsiasi direzione per creare gesti che rappresentino il rispettivo comando. Tuttavia, i gesti devono esprimere un solo movimento (es. tracciare una linea o).
Nel secondo caso (gesti tridimensionali), è libero di scegliere i gesti utilizzando le tre dimensioni. Ogni singolo movimento che sembri naturale o familiare può rappresentare un comando del programma. Le chiedo inoltre di simulare questi gesti con il dispositivo (laserpointer) tenendo premuto il pulsante superiore per la durata del gesto e di disegnarli nei riquadri della tabella.

tipo gesto/ comando	gesti (bidimensionali)	gesti (tridimensionali)
ingrandire		

rimpicciolire		
tagliare mappa		
riunire mappa		

Commenti e idee

- *fattibilità*
- *complessità*
- *livello di impegno per memorizzare*
- *livello di forza necessario*

Il gesto per il comando x Le sembra … ?

Come potrebbe essere migliorato?

Funzioni di modificazione:

- *inserire nuovo argomento principale*
- *inserire nuovo sottoargomento*
- *inserire relazioni ram*
- *creare nuovo contorno*
- *rimuovere selezionato*
- *bilanciare mappa*

tipo gesto/ comando	gesti (bidimensionali)	gesti (tridimensionali)
inserire nuovo argomento principale		
nuovo sotto-argomento		
inserire relazioni rami		

tipo gesto/ comando	gesti (bidimensionali)	gesti (tridimensionali)
creare nuovo contorno		
rieliminare selezionato		
bilanciare mappa		

Commenti e idee
- *fattibilità*
- *complessità*
- *livello di impegno per memorizzare*
- *livello di forza necessario*

Il gesto per il comando x Le sembra ... ?

____

____

Come potrebbe essere migliorato?

____

____

PROPOSTE PER IL FEEDBACK ***(feedback: l'apparecchio ti segnala di aver capito il comando o che sta eseguendo un comando)***

In che modo e quando potrebbe essere utile il feedback di vibrazione ...?

____

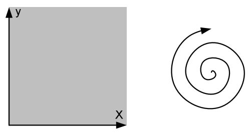
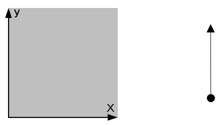
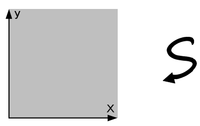
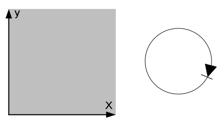
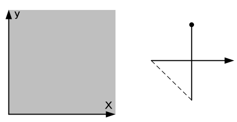
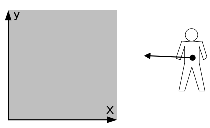
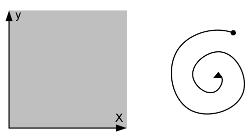
____

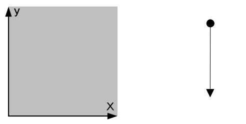
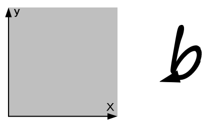
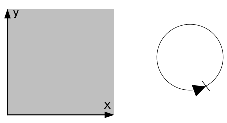
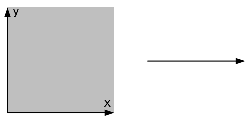
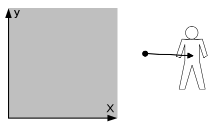
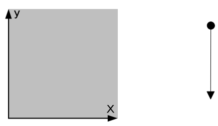
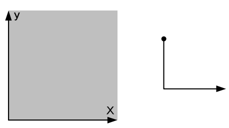
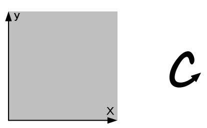
SINTESI:

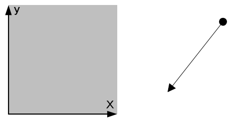
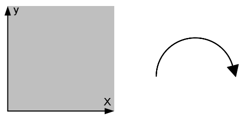
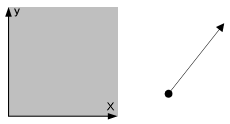
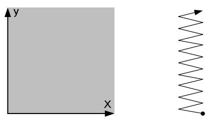
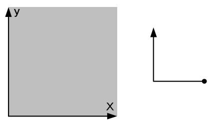
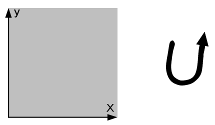
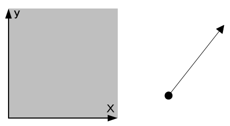
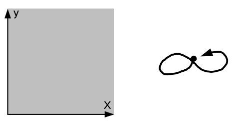
- [ ] *Il feedback non servirebbe da nessuna parte*
- [ ] *Il feedback potrebbe essere usato per segnalare il modo del dispositivo*
- [ ] *Il feedback potrebbe servire a segnalare lo stato del riconoscimento dei gesti*
- [ ] *Il feedback potrebbe essere utile per diverse funzioni del dispositivo*

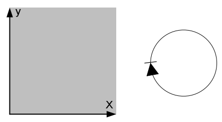
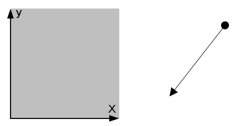
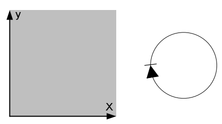
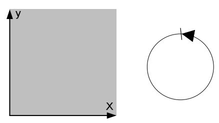
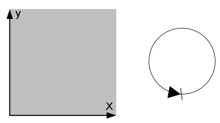
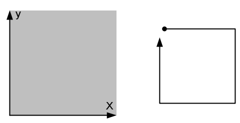
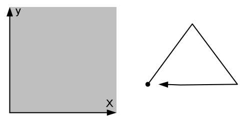
10.2 Appendix B

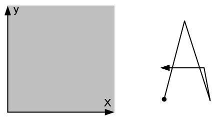
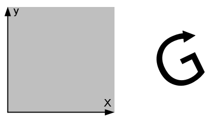
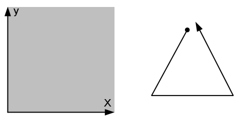
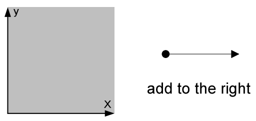
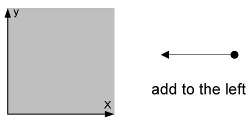
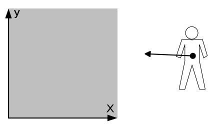
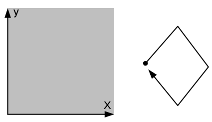
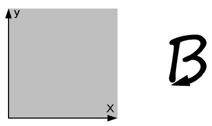
Pre-study gesture collection (after revision).

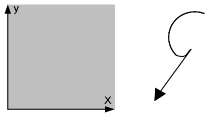
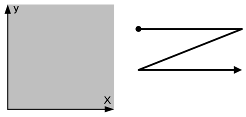
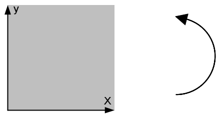
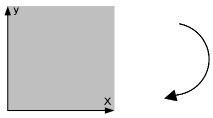
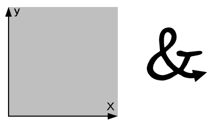
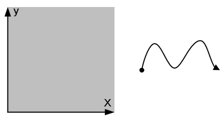
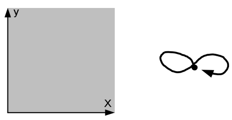
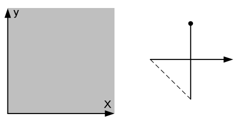
function	gestures (2D)	Participants	Code
zoom-in		1	SZI1
		2, 7	SZI2
		3	SZI3
		4, 7	SZI4
		5	SZI5
		6	SZI6
zoom-out		1	SZO1

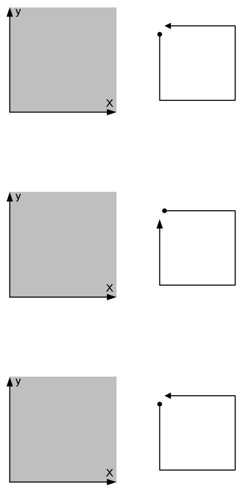
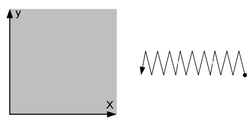
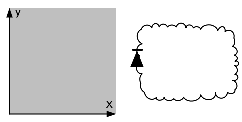
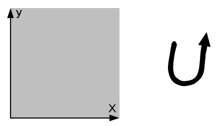
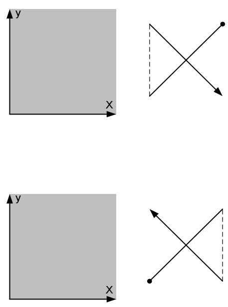
		2, 7	SZO2
		3	SZO3
		4, 7	SZO4
		5	SZO5
		6	SZO6
split map vertically		1, 6	SSM1
		2	SSM2
		3	SSM3

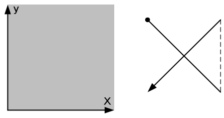
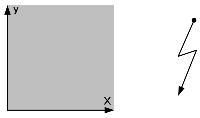
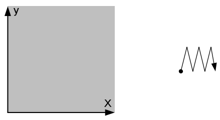
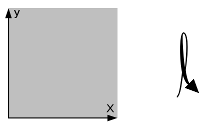
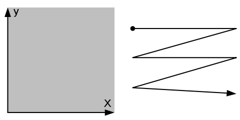
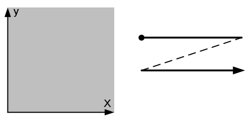
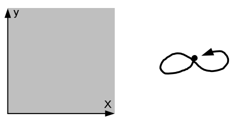
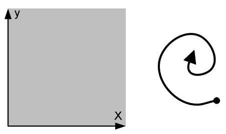
		4	SSM4
		5	SSM5
		7	SSM6
remove split		1	SRS1
		2	SRS2
		3	SRS3
	 	4	SRS4 SRS5

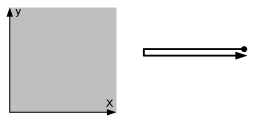
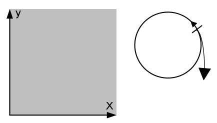
		5	SRS6
		6	SRS7
		7	SRS8
add sibling topic	  	1,2,6	SAT1
		3	SAT3
		4	SAT4

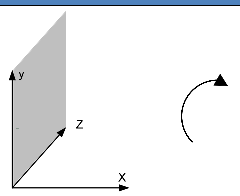
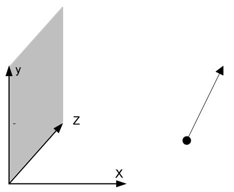
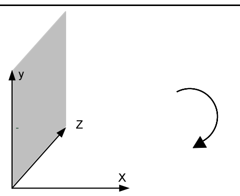
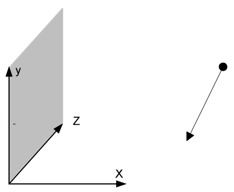
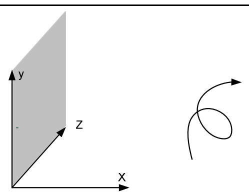
		5	SAT5
		7	SAT7
add subtopic		1	SST1
	  	2,6	SST2
		4	SST4
		5	SST5

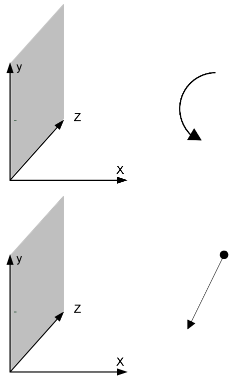
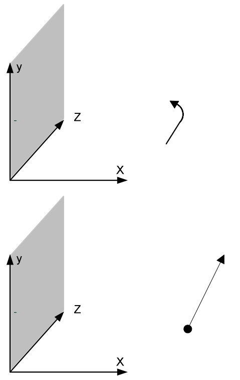
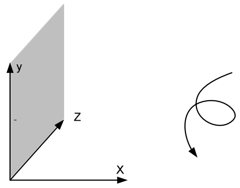
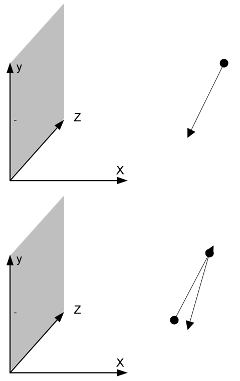
		7	SST7
add relationship line		1, 6	SRL1
		2	SRL2
		3	SRL3
		4	SRL4
		5	SRL5
		6	SRL6
		7	SRL7

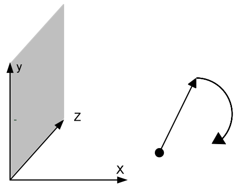
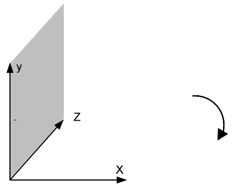
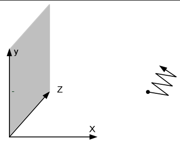
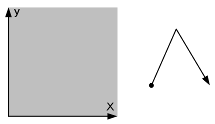
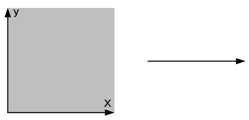
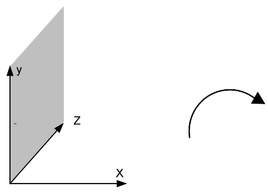
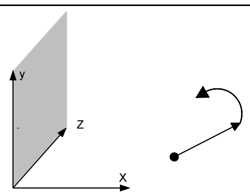
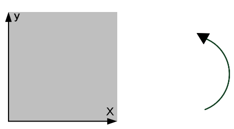
add boundary		1,2,4,6	SAB1
		3	SAB3
		5	SAB4
		7	SAB6
remove only selected topic		1,2,4	SRT1

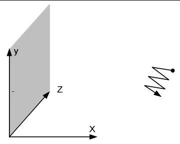
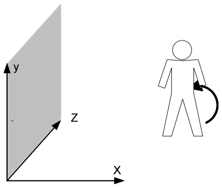
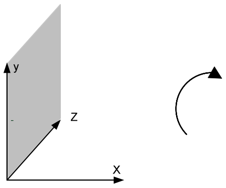
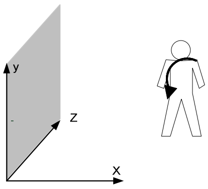
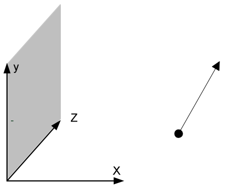
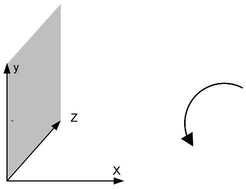
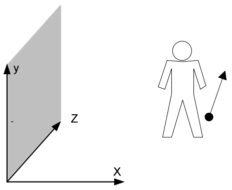
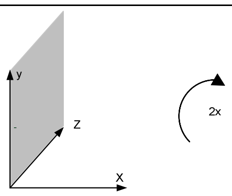
	 !		
		5	SRT4
		6	SRT5
		7	SRT6
balance map	 	1,2,5	SBM1
		3	SBM2
		4	SBM3

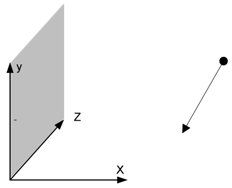
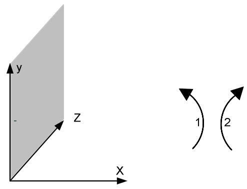
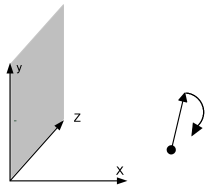
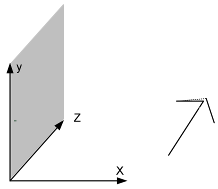
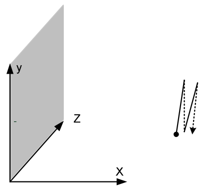
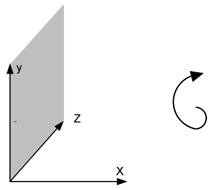
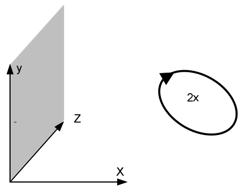
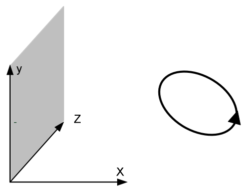
		6	SBM5
		7	SBM6

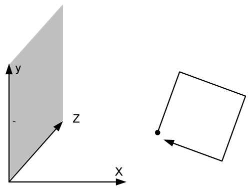
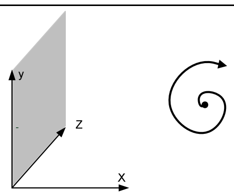
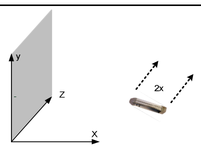
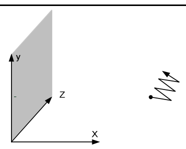
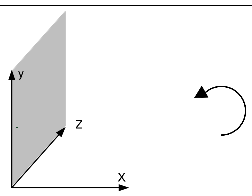
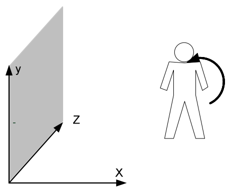
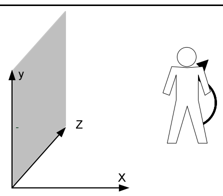
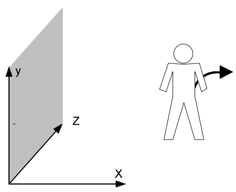
function	gestures (3D)	Participants	
zoom-in	 	1,2,3,5	MZI1
	 	4,6	MZI3
		7	MZI5

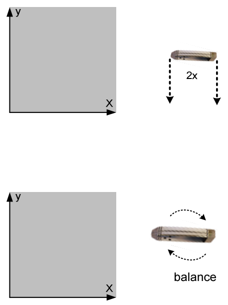
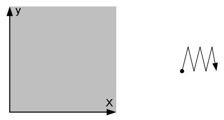
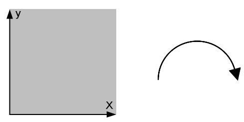
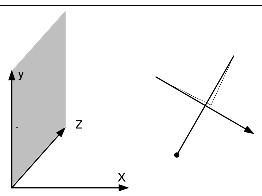
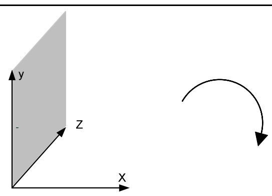
zoom-out		1,2,3,5	MZO1
		4,6	MZO3
		7	MZO5
split map vertically		1,5	MSM1

	 	2,7	MSM2
		3	MSM3
		4	MSM4
remove split	 	1,6	MRS1
	 	2,7	MRS2

		3	MRS3
add sibling topic		1	MAT1
	  	2,4,5,7	MAT2
		6	MAT5
add subtopic		1	MST1
		2, 7	MST2

		5	MST3
add relationship line		1	MRL1
		2	MRL2
		4	MRL3
		6	MRL4
		7	MRL5
add boundary		1,2	MAB1

		4	MAB3
		7	MAB4
remove selected topic		1	MRS1
		2	MRS2
	 	3,4	MRS3
	 	5,6,7	MRS5

balance map		1,3,5	MBM1
		2	MBM2
		4	MBM3
		6	MBM5
		7	MBM6

10.3 Appendix C

User test materials

Herzlich Willkommen

Zunächst möchten wir uns bei Ihnen bedanken, dass Sie sich bereit erklärt haben, an unserer Untersuchung teilzunehmen. Bevor es nun gleich losgeht, wollen wir Ihnen mit Hilfe dieser kurzen Einführung vermitteln, um was es uns bei dieser Untersuchung überhaupt geht und welche Rolle Sie dabei spielen.

Große, hochauflösende Displays, sind eine nicht zu unterschätzende Alternative zu Projektoren (Beamer) für die Darstellung und effektvolle Präsentation komplexer Sachverhalte und großer Datenmengen. Der Vorteil bei einem Gerät wie der POWERWALL der Universität Konstanz ist vor allem die enorme effektive Auflösung auf einer Fläche von 5x2 Metern. Nachteile bestehen (außer in den noch sehr hohen Anschaffungs- und Betriebskosten) in dem Problem der Steuerung von Präsentationen und Programmen, die bisher nur mithilfe einer (semi-)stationären Lösung mit einer einfachen Computermaus bewerkstelligt wurde.

In unserer Studie wollen wir nun eine alternative Steuerungsmöglichkeit auf ihre Zweckmäßigkeit und Nützlichkeit untersuchen um eine Voraussetzung bereitzustellen die Vorteile der POWERWALL besser nutzen zu können und die Arbeit unkomplizierter und angenehmer zu gestalten. Hierfür besteht neben der direkten Steuerung des Mauszeigers durch Zeigebewegungen des Laserpointers auch die Möglichkeit, dem Computer einzelne Befehle über Gesten zu übermitteln. Ziel ist es herauszufinden, ob dieser zusätzliche Modus des Laserpointers sich für die Steuerung von z.B. MindMapping-Programmen eignet. Und an dieser Stelle kommen Sie ins Spiel, denn der beste Weg für uns, dies herauszufinden, besteht darin, dem Benutzer direkt bei der Interaktion mit der POWERWALL zu zuschauen. Wir werden Sie also im Laufe der Untersuchung bitten, bestimmte Aufgaben im MindMapping-Programm durchzuführen und anschließend Ihre Meinung zu der Benutzung mit Hilfe von verschiedenen Fragebögen kundzutun.

Die Steuerungsgeräte stehen also bei dieser Untersuchung auf dem Prüfstand und nicht Sie als Benutzer. Sie sind vielmehr in der Rolle des Prüfers, welcher uns die Möglichkeit gibt, Benutzungsprobleme mit den Geräten und dem Display zu erkennen und letztendlich zu beseitigen.

Für die Auswertung der gewonnenen Daten ist es notwendig, dass wir den Test auf Video aufzeichnen können. Hierfür benötigen wir allerdings Ihr Einverständnis, wobei wir uns im Gegenzug verpflichten, das Videomaterial anonymisiert und lediglich zu Auswertungszwecken zu verwenden. In diesem Zusammenhang haben wir ein separates Dokument vorbereitet, welches Sie auf der nächsten Seite finden.

Abschließend wünschen wir Ihnen viel Spaß und möchten uns noch einmal für Ihre Teilnahme bedanken!

Bedienung des Programms:

Bedienung des Programms – Direkte Steuerung

Während des Tests werden Sie mit einem speziellen Laserpointer arbeiten. Diesen können Sie zum Zeigen und Selektieren von Objekten im MindMapping-Programm benutzen. Der linke Knopf funktioniert hierbei ähnlich wie die linke Maustaste (Klicken).

Bedienung des Programms – Ausführen von Befehlen mit Gesten

Unabhängig davon wird Ihnen die Ausführung der auf der folgenden Seite vorgestellten Funktionen (Befehle) des Programms mithilfe einzelner Gesten möglich sein, die durch einen Bewegungssensor im Gerät aufgezeichnet werden:

Sagen Sie nun bitte dem Untersuchungsleiter, dass Sie bis zu dieser Stelle gelesen haben!

Funktionen des Programms:

Funktionen des Programms – Direkte Steuerung

Im Programm können Sie die dargestellten Elemente durch Klicken **selektieren**. **Mehrere Objekte selektieren** Sie indem Sie auf eine leere Stelle klicken und durch Halten des Knopfes das Selektionsquadrat um die gewünschten Objekte ziehen. Klicken Sie auf ein Objekt und halten Sie den linken Knopf um es zu **verschieben**. Die Menüs und Buttons im Programm sind für den Test nicht relevant.

Funktionen des Programms –Befehle ausgeführt durch Gesten
Durch die Ausführung einzelner Gesten können Sie die folgenden Programmbefehle ausführen:

Vergrößern (**zoom-in**)	-	Vergrößern der Ansicht der map
Verkleinern (**zoom-out**)	-	Verkleinern der Ansicht der map
Objekt kopieren (**copy**)	-	Kopieren von einem oder mehreren vorher selektierten Objekten der map
Objekt einfügen (**paste**)	-	Einfügen von kopierten Objekten an der vorher selektierten Stelle auf der map
neues **Objekt hinzufügen**	**-**	Erstellen eines neuen Objektes ("topic")
Verbindungslinie hinzufügen	**-**	Erstellen einer Linie zwischen zwei Objekten der map (Moduswechsel)
Umrandung hinzufügen	**-**	Erstellen eines Rahmens um ein vorher selektiertes Objekt und aller untergeordneten
Objekt löschen	**-**	Löschen des vorher selektierten Objekts und evtl. untergeordneten Objekten

Gesten:

Jeder dieser Befehle kann durch eine bestimmte Geste ausgeführt werden. Die einzelnen Gesten werden Ihnen durch den Testleiter im Folgenden demonstriert. Üben Sie jede einfach 10x, bevor wir sie dem System beibringen.

Ausführen der Gesten (Trainieren des Systems):

Bevor Sie die Gesten im Programm benutzen können, müssen diese im System trainiert werden. Um dem System zu signalisieren, dass Sie eine Geste durchführen, drücken Sie den mittleren Knopf des Laserpointers. Führen Sie die Geste 15x aus. Wichtig hierbei ist, dass der mittlere Knopf während der gesamten Bewegung gedrückt wird, da hierdurch dem System die Informationen Ihrer Bewegung übermittelt wird. Der Testleiter wird Sie nach Ausführen der Trainingsbewegungen auf das weitere Vorgehen hinweisen.

Zusammenfassung des Ablaufs – 1.Teil:

Für jeden Befehl/Geste läuft die Einführung inkl. des Trainings also wie folgt ab:

- Demonstration der Programmfunktion (durch den Testleiter)
- Vorführen der Gestenbewegung (durch den Testleiter)
- Üben der Geste durch den Benutzer (Sie!): 10x Ausführen der Bewegung
- Beginn des Trainings: 15x Ausführen der Geste, jeweils durch Halten des mittleren Knopfes von Anfang bis Ende der Bewegung
- Ende des Trainings und Ausprobieren des Befehls anhand der Einzelaufgabe (siehe nächste Seite)

Einzelaufgaben:

Jede so durch das System gespeicherte Geste können Sie im Anschluss an das Training jeder Geste im Programm ausprobieren. Hierfür führen Sie nach Aufforderung des Testleiters die Geste an einer Beispiel-map im Programm aus.

Zum Ausführen einer Geste drücken Sie bitte den mittleren Knopf des Laserpointers und halten Sie diesen gedrückt bis sie die Geste vollständig ausgeführt haben. Sollte das System die Geste nicht erkannt haben, wird Ihnen dies anhand eines kurzen Vibrations-feedback durch den Laserpointer mitgeteilt. Sollte das System eine andere als die gewollte Geste ausführen, versuchen Sie es erneut.

Die Einzelaufgaben sind nummeriert und bestehen aus den folgenden Aktionen:

1. Vergrößern und Verkleinern - **Zoomen** Sie jeweils abwechselnd in die map **hinein** und **hinaus**. Die jeweilige Geste sollte ingesamt 3x ausgeführt werden

2. Objekt kopieren und einfügen - Selektieren Sie ein Objekt der map durch Klicken des linken Knopfes und führen die **copy**-Geste zum Kopieren aus. Fügen Sie es an einer beliebigen Stelle in der map mithilfe der **paste**-Geste wieder ein.

   Neues Objekt hinzufügen - Fügen Sie mithilfe der **Objekt hinzufügen-**Geste der map ein neues Objekt ("topic") hinzu.

3. Verbindungslinie hinzufügen - Wechseln Sie mithilfe der entsprechenden Geste in den **Verbindungslinien-Modus.** Verbinden Sie nun Verbinden Sie nun 2 Objekte Ihrer Wahl durch Klicks.

4. Umrandung hinzufügen - Erstellen Sie mithilfe der **Umrandung-**Geste einen Rahmen um Objekte Ihrer Wahl, indem Sie vorher das „Kopf"-Objekt selektieren, welchem alle weiteren ungeordnet sind.

5. Objekt(e) löschen - Entfernen Sie mithilfe der **Löschen-**Geste vorher selektierte Objekte Ihrer Wahl von der map.

10.4 Appendix D

Einverständniserklärung

Bitte lesen Sie die folgenden Zeilen aufmerksam durch.

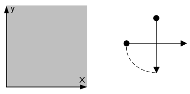
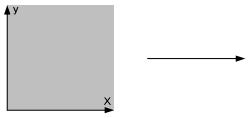
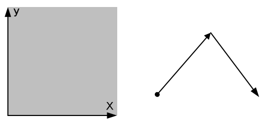
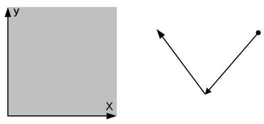
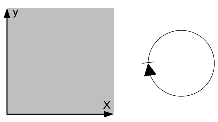
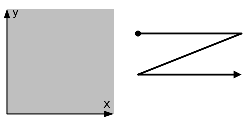
Um eine bessere Auswertung der gewonnenen Daten zu erreichen, werden wir eine Videoaufzeichnung des Tests vornehmen. Durch die Unterzeichnung dieses Formulars erklären Sie sich damit einverstanden. Im Gegenzug garantieren wir Ihnen, die Aufzeichnung anonymisiert und lediglich zu Auswertungszwecken zu verwenden.

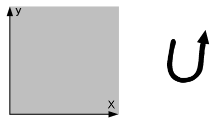
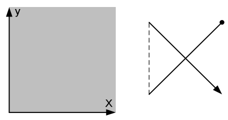
Hiermit erkläre ich mich mit den oben genannten Punkten einverstanden:
Name, Vorname ______________________________
Unterschrift ______________________________
Datum ______________________________

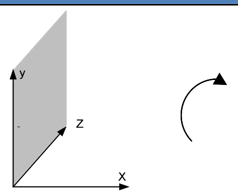
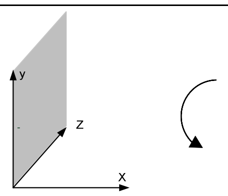
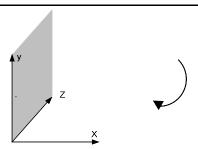
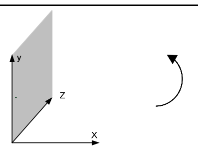
Hiermit verpflichtet sich die Untersuchungsleitung, die Videoaufzeichnung sowie sämtliche sonstigen gewonnenen Daten lediglich zu Auswertungszwecken im Rahmen dieser Untersuchung zu verwenden:
Name, Vorname ______________________________
Unterschrift ______________________________
Datum ______________________________

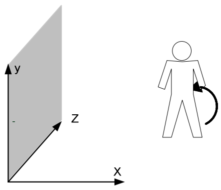
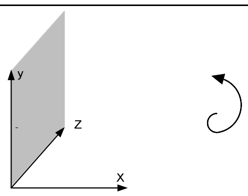
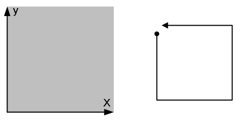
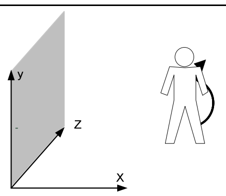
10.5 Appendix E

Gesture sets of the user tests

function	gestures (symbol/object-iconic)	participants
zoom-in		2, 7
zoom-out		5
copy selected		
paste selected		
add sibling topic		1,2,6
add relationship line		1, 6

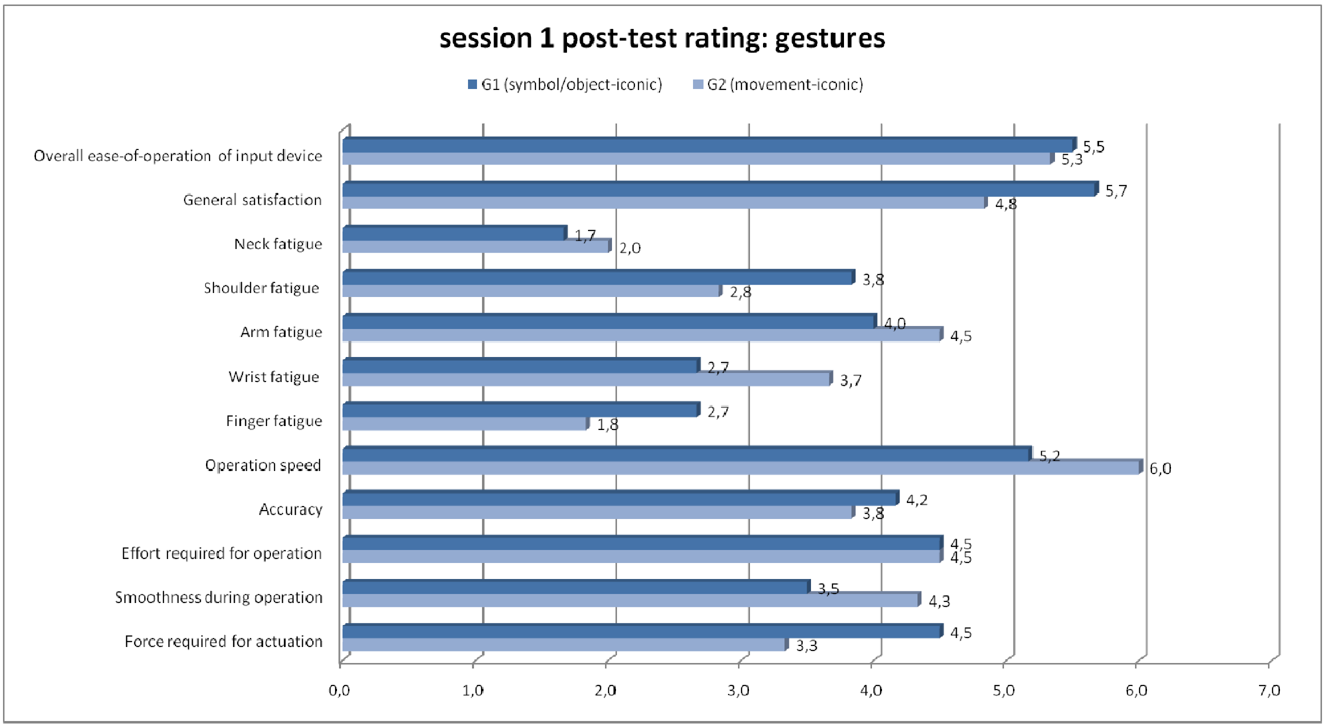
add boundary		7
remove selected topic		1,2,4

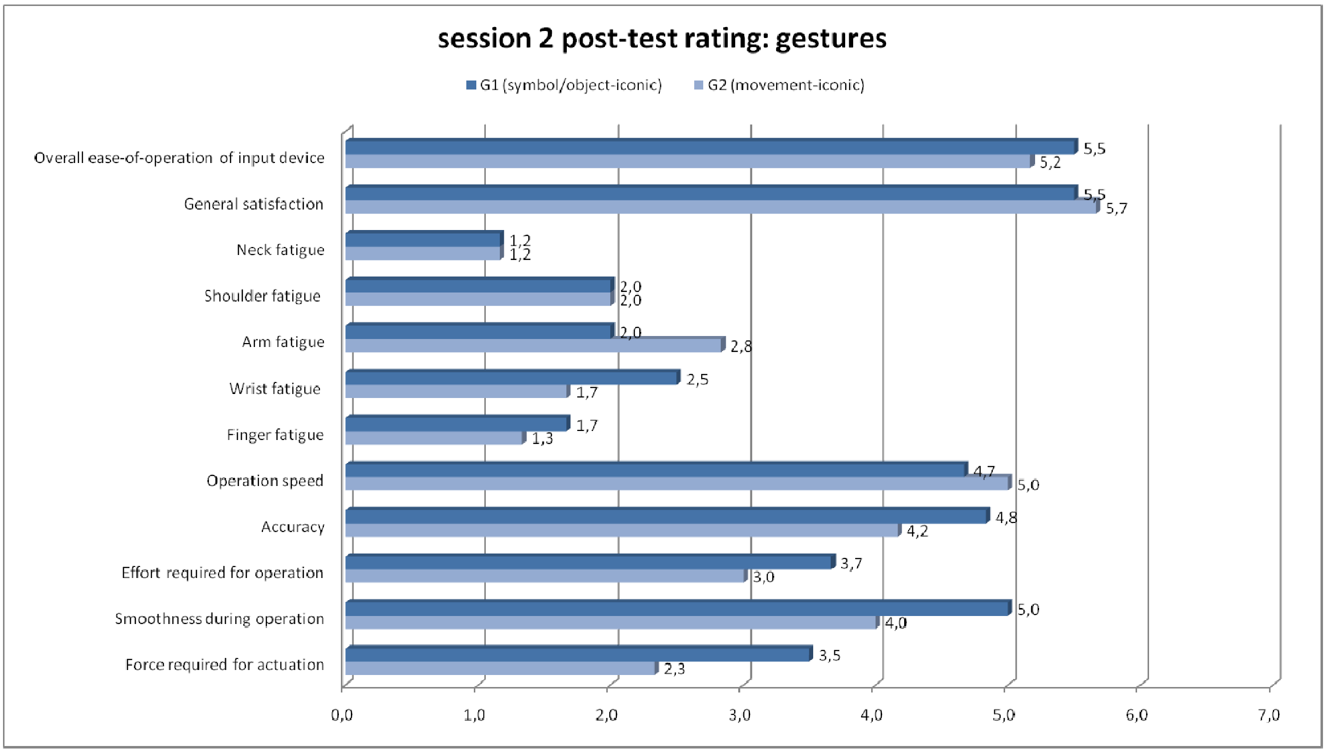
function	gestures (movement-iconic)	participants
zoom-in		1,2,3,5
zoom-out		1,2,3,5
copy selected		
paste selected		

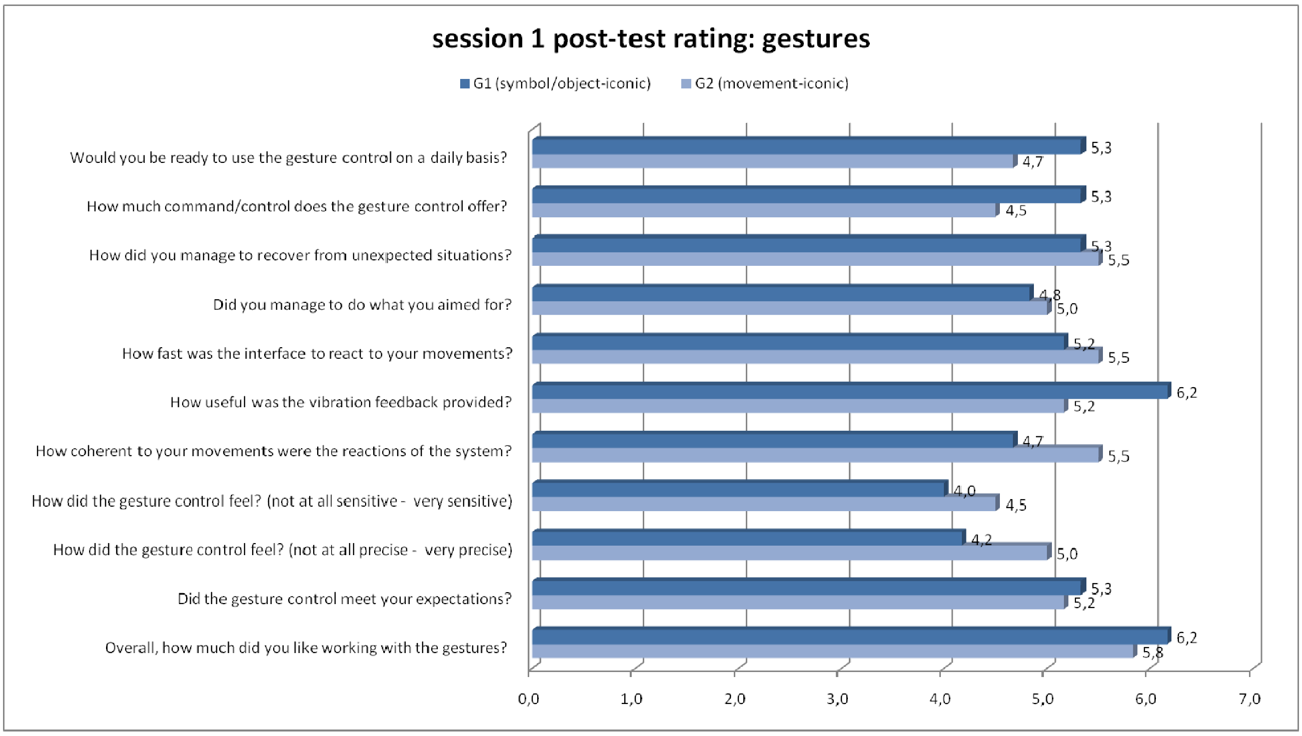
add sibling topic		1
add relationship line		7
add boundary		1,2,4,6
remove selected topic		5,6,7

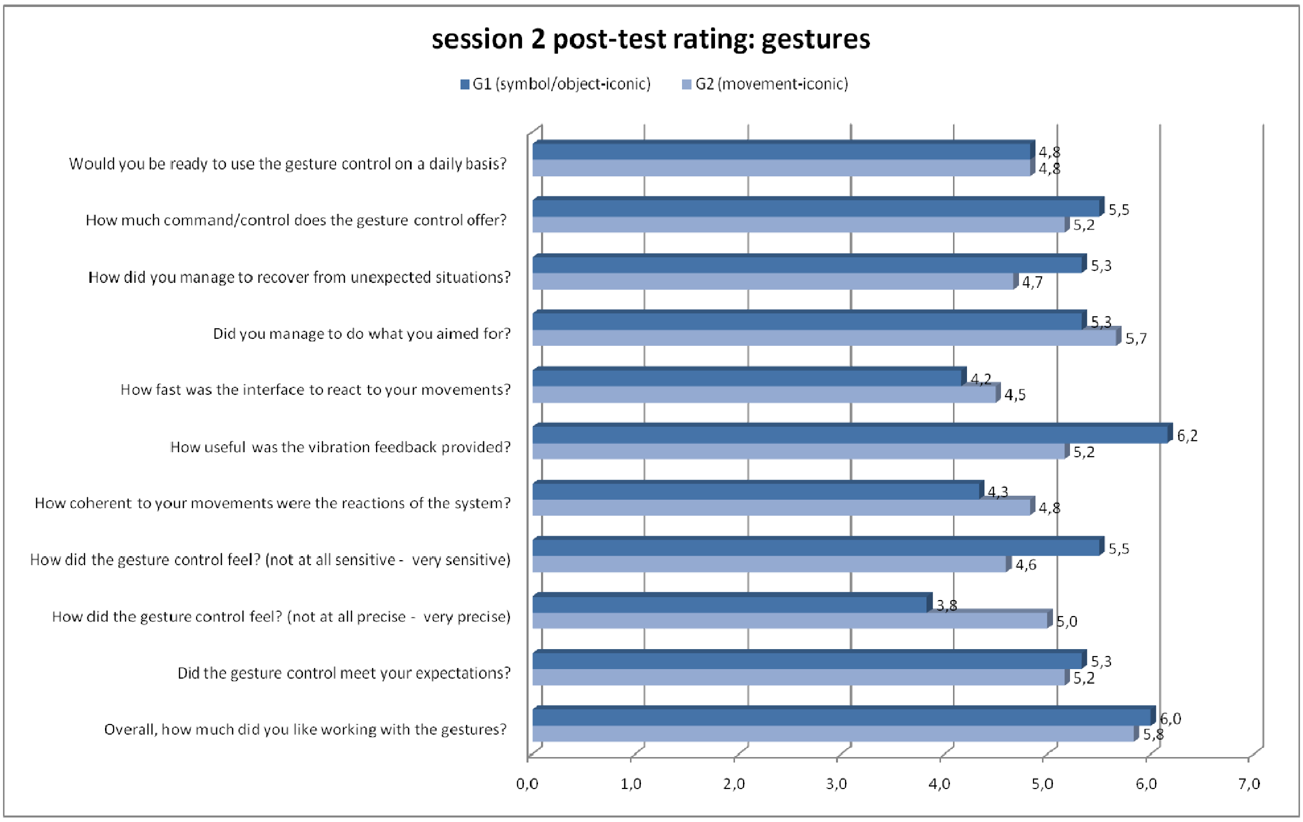
10.6 Appendix F

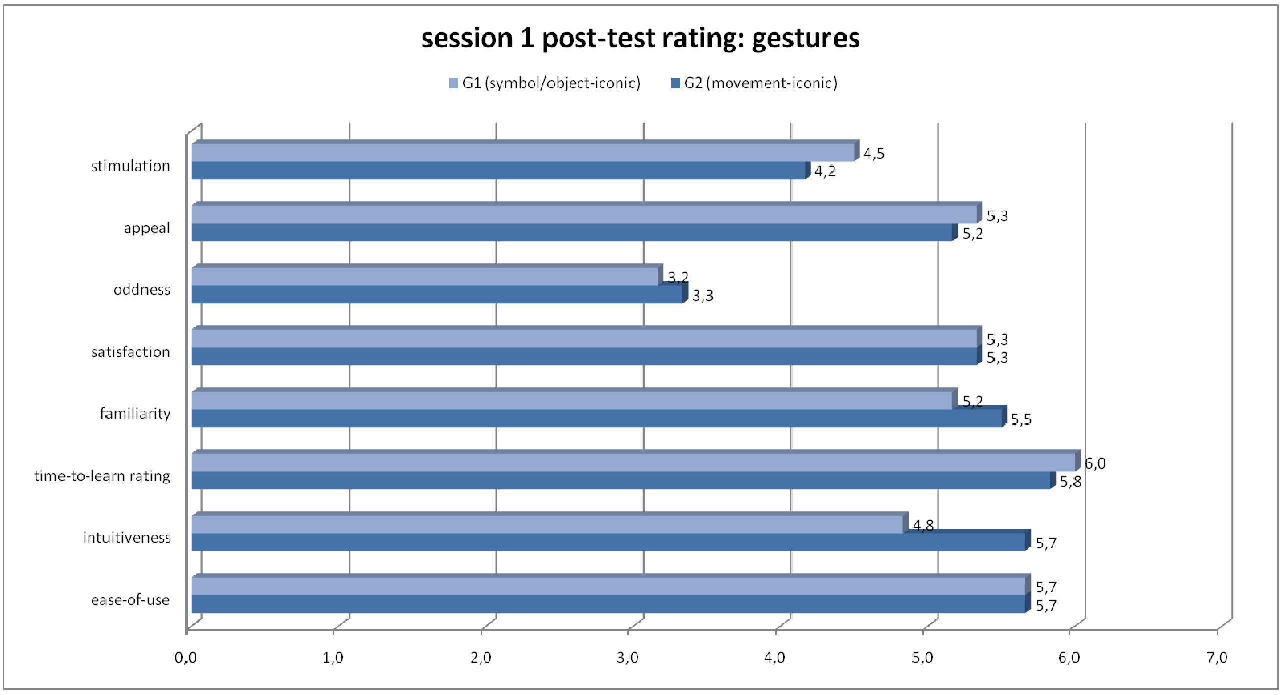
Further results from questionnaire ratings

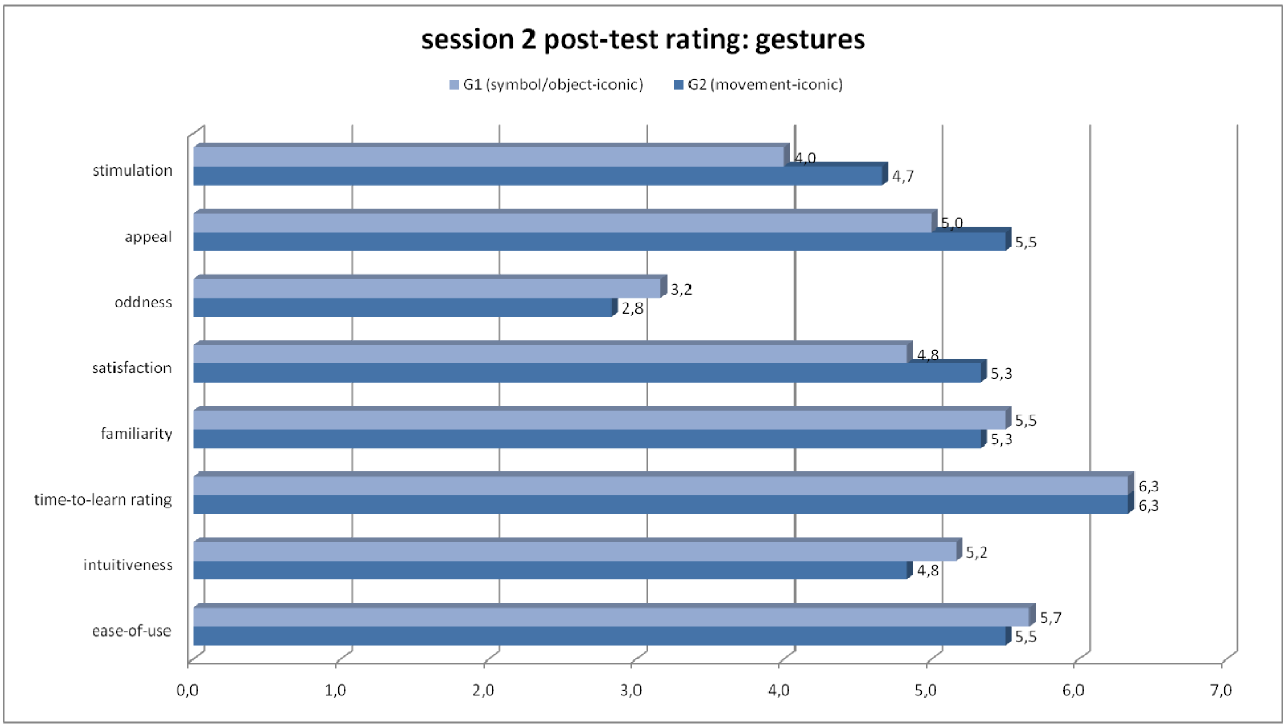





session 1 post-test rating: gestures
G1 (symbol/object-iconic)
G2 (movement-iconic)
Would you be ready to use the gesture control on a daily basis?
How much command/control does the gesture control offer?
How did you manage to recover from unexpected situations?
Did you manage to do what you aimed for?
How fast was the interface to react to your movements?
How useful was the vibration feedback provided?
How coherent to your movements were the reactions of the system?
How did the gesture control feel? (not at all sensitive - very sensitive)
How did the gesture control feel? (not at all precise - very precise)
Did the gesture control meet your expectations?
Overall, how much did you like working with the gestures?
5,3
4,7
5,3
4,5
5,3
5,5
4,8
5,0
5,2
5,5
6,2
5,2
4,7
5,5
4,0
4,5
4,2
5,0
5,3
5,2
6,2
5,8
0,0
1,0
2,0
3,0
4,0
5,0
6,0
7,0


session 2 post-test rating: gestures
G1 (symbol/object-iconic)
G2 (movement-iconic)
Would you be ready to use the gesture control on a daily basis?
How much command/control does the gesture control offer?
How did you manage to recover from unexpected situations?
Did you manage to do what you aimed for?
How fast was the interface to react to your movements?
How useful was the vibration feedback provided?
How coherent to your movements were the reactions of the system?
How did the gesture control feel? (not at all sensitive - very sensitive)
How did the gesture control feel? (not at all precise - very precise)
Did the gesture control meet your expectations?
Overall, how much did you like working with the gestures?
4,8
4,8
5,5
5,2
5,3
4,7
5,3
5,7
4,2
4,5
6,2
5,2
4,3
4,8
5,5
4,6
3,8
5,0
5,3
5,2
6,0
5,8
0,0
1,0
2,0
3,0
4,0
5,0
6,0
7,0


session 1 post-test rating: gestures
G1 (symbol/object-iconic)
G2 (movement-iconic)
stimulation
appeal
oddness
satisfaction
familiarity
time-to-learn rating
intuitiveness
ease-of-use
4,5
4,2
5,3
5,2
3,2
3,3
5,3
5,3
5,2
5,5
6,0
5,8
4,8
5,7
5,7
5,7
0,0
1,0
2,0
3,0
4,0
5,0
6,0
7,0


session 2 post-test rating: gestures
G1 (symbol/object-iconic)
G2 (movement-iconic)
stimulation
appeal
oddness
satisfaction
familiarity
time-to-learn rating
intuitiveness
ease-of-use
4,0
4,7
5,0
5,5
3,2
2,8
4,8
5,3
5,5
5,3
6,3
6,3
5,2
4,8
5,7
5,5
0,0
1,0
2,0
3,0
4,0
5,0
6,0
7,0