Digital Transformatives

Activating User Intrinsic Potentials

Von der Fakultät für Mathematik, Informatik und Naturwissenschaften der RWTH Aachen University zur Erlangung des akademischen Grades eines Doktors der Ingenieurwissenschaften genehmigte Dissertation

vorgelegt von

Diplom-Ingenieur
Hagen Buchholz

aus Wiesbaden

Berichter(innen): Univ.-Prof. Dr. rer. pol. Matthias Jarke

Prof. Wolfgang Prinz, PhD

Prof. Irene Mittelberg, PhD

Tag der mündlichen Prüfung: 08. August 2014

Diese Dissertation ist auf den Internetseiten der Hochschulbibliothek online verfügbar.

Contents

1 INTRODUCTION 1

1.1 BASIC IDEA 3

1.2 RESEARCH QUESTIONS 4

1.3 RESEARCH METHODOLOGY 4
1.3.1 Hypothesis-Feature Graph 6

1.4 OUTLINE 7

1.5 BASIC SCHEMA OF DIGITAL TRANSFORMATIVES 7
1.5.1 Performance and context 7
1.5.2 Basic Schema of Digital Transformatives 8

2 RELEVANT COGNITIVE AND PRACTICAL BACKGROUND 13

2.1 RELEVANT COGNITIVE MECHANISMS 14
2.1.1 Mnemonics Devices 14
Cognitive Evidence based Research on Mnemonics Devices 17
Conclusions 19
2.1.2 Human Communication as Transformative 19
2.1.3 Conceptual Metaphors 21
Conclusions 22
2.1.4 Concepts, Prototypes, and Categories 23
Similarity Measures for Prototype Categorization 32
Conclusions 34
2.1.5 Performance, proceduralization, and category prototypes 34
From controlled to automated processes to habituation 38
Conclusions 43
2.1.6 Pervasiveness of Similarity Comparisons 44
Conclusions 45
2.1.7 Cognitive Efficiency Catalysts in Communication 46
Conclusions 49
2.1.8 Efficiency in Semiotics 50
Conclusions 51
2.1.9 Summary of Human Cognition in Digital Transformatives 52

2.2 RETROSPECTIVE PROPERTY EXTRACTION OF EXISTING DIGITAL TRANSFORMATIVES **55**
2.2.1 General Property Extraction Procedure 55
Determine Relevant Systems Shifting the User Context 56
Extract the Context Shifts 56
Cognitive Assessment Metrics for Digital Context Shifts 57
Collect characteristics and Find Shared Characteristics 58
2.2.2 Determine Relevant Digital Transformatives 58
2.2.3 Cognitive Domain 59
Digital Mnemonics 59
Conclusions 66
2.2.4 User Interface Metaphors 69
The Desktop Metaphor 69
Spatial and Narrative Metaphors 71
Conclusions - User Interface Metaphors 74
2.2.5 Affective Domain 76
Games as a Metaphor - Gamification and Serious Games 76
Explorative Playful Environments 79
Curiosity Driven Ludic Examples 86
Conclusions 87
2.2.6 Psychomotor Domain 89
Cognitive and Practical Background Constructivistic Learning with Transitional Objects and Digital Manipulatives 89
Tangible User Interfaces 94
2.2.7 Affordances from UI Concepts 99
Conclusions 100
2.2.8 Summary of the Retrospective Property Extraction 101

3 CONCEPT DESIGN OF DIGITAL TRANSFORMATIVES **105**

3.1 CONCEPT OF COGNITIVE EFFICIENCY DRIVERS IN HUMAN COMMUNICATION **105**

3.2 DIGITAL TRANSFORMATIVE CONCEPT - COGNITIVE PROTOTYPE BASED SYSTEM DEVELOPMENT FOR IMPROVED EFFICIENCY IN HUMAN-MACHINE COMMUNICATION **107**

3.3 DIGITAL TRANSFORMATIVE MAIN CHARACTERISTICS **112**

3.4 BASIC CONCEPT VALIDATION **115**
3.4.1 Familiarity as a Driver for Human Efficiency 115
Familiarity Relates to Productivity in Work Environments 115
3.4.2 Memory - Basic Concept Test Case 116
Memory - Setup, Apparatus, and Test Procedure 118
Memory - Evaluation 119
Conclusions 122

3.4.3 Correlation between Familiarity and Performance in Team Sports 122
Measures for performance and familiarity 123
Team Sport Analytical Procedure and Results 124
Conclusion 126
3.4.4 Conclusions of the Basic Concept Validation 128

4 DESIGNING DIGITAL TRANSFORMATIVES 131

4.1 DESIGN METHODOLOGY OF INTERACTIVE SYSTEMS 131
4.1.1 Interactive Systems Development 133
4.1.2 Interactive Systems Design Phase 135
Requirements Engineering, Specification, and Brainstorming 135
Conceptualization and Prototyping 146
Evaluation 148
Conclusions 149

4.2 INTERACTIVE SYSTEM DESIGN METHODOLOGY FOR DIGITAL TRANSFORMATIVES 152
4.2.1 Methods for Assessing Familiarity 154
Heuristic expert estimation 154
Probabilistic Environmental Observation 154
Learning Curve Analysis 162
User Rated Familiarity 163

4.3 SYSTEMATIC METHODOLOGY FOR FINDING A DIGITAL TRANSFORMATIVE CONTEXT 171
4.3.1 Prerequisites 172
4.3.2 DT Design Challenge 172
4.3.3 Salient Super Prototype Identification Approach 175
Similarity Based Cognitive Structures of Prototype Categories 175
User Determined Prototype Concepts 178
Salient Super Prototype Identification Procedure 181
4.3.4 Sub-Action Modeling Approach 183
Sub-Action Modeling Procedure 184

4.4 SUMMARY - OVERALL DIGITAL TRANSFORMATIVES DESIGN PROCESS 186

5 ITERATIVE USE CASE PROTOTYPES 189

5.1 GENERATION 1 – CONCEPT DESIGNS AND SCIENCE CENTER TO GO 190
5.1.1 Science Center To Go 193
Prototype Evolution 193
The Latest Prototype: The Science Center To Go Suitcase 195
The Mini Wing Experiment 196

The Doppler Experiment 197
The Double Slit Experiment 198
The Double Cone Experiment 199
The Boltzmann Experiment 201
Science Center To Go: A Digital Transformative 202
Science Center To Go Evaluation 204

5.2 GENERATION 2 – FIRST APPROACHES BASED ON FAMILIARITY 206
5.2.1 Captology and Persuasive Technologies 206
Conclusion and critical reflection 208
5.2.2 Drinking Garden – a Digital Beverage Coaster 210
Concept design 210
Implementation 211
First Tests 213
Test results 214
Results in Detail 214
Conclusions 216
5.2.3 Assess the Assessable 218
Background 219
System Concept for Releasing User Intrinsic Performance 221
Experimental Methodology 224
Pre-Tests 225
Tests 226
Evaluation Results 228
Discussion and Future Work 237

5.3 GENERATION 3 – COOPERATIVE COGNITIVE PROTOTYPING TOOLS 239
5.3.1 DT Web Module Sub-Action Modelling 240
5.3.2 DT Web Module Salient Super Prototype Modeling 242
Cognitive Prototype Oriented System Design 242
Salient Super Concept Modeling Game Concept 243
Conclusions – DT Mechanisms and Guidelines 246
5.3.3 Reading Speed Test 247
5.3.4 Conclusion – How to Create Tools for Capturing Cognitive Concept Structures 250
Choose Proper Representations of Concepts 250
Determine Relative Salience 250
Determine Relative Similarity Distance 251
Concluding Guidelines 251

6 CONCLUSION, CRITICAL REFLECTION AND OUTLOOK 253

6.1 DIGITAL TRANSFORMATIVES SUMMARY 253
6.1.1 Basic Schema 253

6.1.2 Background and Working Principle 255
Background on cognitive efficiency 255
Retrospective study of existing Systems 255
Model for Cognitive Efficiency in Communication 256
Working Principle of Digital Transformatives 257
6.1.3 Designing Digital Transformatives 258
Super Salient Prototype Identification 260
Sub-Action Modeling Approach 261
Methods for Familiarity Assessment 262
Relations between Familiarity, automatization, and Cognitive Prototypes 265
Integrated System Design Methodology 265
6.1.4 Use Case Prototypes 267
6.1.5 Function Characteristics and Design Guidelines 268
Functional Characteristics 269
Design Guidelines 270

6.2 FUTURE WORK – COGNITIVE PROTOTYPES BEYOND INTERACTIVE SYSTEM DESIGN 273
6.2.1 CPs at Conventions, Standards, Norms, and the Force of Habit 273
6.2.2 CPs in Perception and Recognition 275
6.2.3 CPs Help Promoting an Image 277
6.2.4 CPs on Pleasantness and Affection 279
6.2.5 CPs in Communication and Innovation 281
6.2.6 CPs in social structures 282

6.3 NEW CHANCES TO ANALYZE AND UTILIZE COGNITIVE PROTOTYPES 283

6.4 CRITICAL REVISION 284
6.4.1 DTs might Hinder Personal Development 284
6.4.2 Failures through the Force of Habit 284

6.5 WRAP-UP AND FUTURE WORK 285
6.5.1 Research Questions 285
RQ1. Is it possible to learn from mnemonics in order to improve human machine interfaces? 285
RQ2. Are the key working principles of mnemonic devices applicable through human machine interfaces (HMI) of interactive systems? 285
RQ3. Are there methodologies for systematically applying the working principles, in order to foster the creation of such enhanced systems? 285
6.5.2 Outcomes Beyond Research Questions 285
6.5.3 Outlook 286

APPENDIX A - INVERSE DOT PATTERN ILLUSTRATION 292

APPENDIX B - HYPOTHESES, FEATURES, AND GUIDELINES 294

CONCEPT DESIGN GUIDELINES 299

FORMULAS 302

TABLES 303

FIGURES 304

BIBLIOGRAPHY 314

RELATED PUBLICATIONS 350

CURRICULUM VITAE 352

Abstract

Since the advent of online app-repositories, accessibility to interactive digital systems increased tremendously. Nowadays, users are able to directly download and try several alternative applications. This generally tightens decision characteristics, putting the selling point focus further towards aspects such as "joy of use" or "intuitiveness". But what exactly do users mean when they express such demands? Intuitiveness rather describes a feeling than a measurable feature, making it hard to be addressed by developers. This text provides an understanding and measures for those fuzzy terms.

However, first and foremost, a new class of highly efficient interactive digital systems is being defined and characterized, aiming for their systematic development. Following the idea of cognitive efficiency mechanisms, such as mnemonic devices, it is the goal of those systems to activate hidden user potentials by transforming the original function context into a highly efficient usage context. Since the transformation is implemented digitally within the system they are called Digital Transformatives.

The thesis initially provides a defining schema for the identification of Digital Transformatives. The schema is complemented by a model of efficiency in human communication, which is developed based on evidence based cognitive research, and validated on practical examples. Based on those findings a concept is deduced, describing the Digital Transformatives working principle. Hereby the importance of cognitive prototypes is highlighted and further investigated.

The work follows an iterative research methodology, gradually evolving functional characteristics and design guidelines for the development of cognitive prototype oriented systems; also applicable for human machine interaction in general. Moreover, certain cognitive findings are described, providing a selective perspective on psychological aspects, especially involved in communication of enhanced efficiency. Hereby it should be noted that the structure and relations among the presented processes have been deduced by the author from cognition literature, and may vary from typical presentations in this field. This thesis also provides explanations on the efficiency advantages of further implementations, such as Tangible User Interfaces, User Interface Metaphors, Transitional Objects, Persuasive Technologies, and comparative assessment in user evaluations. Finally, this text highlights the importance and chances of social network analyses for the identification of cognitive shared prototypes in various application fields apart from interactive system design, including innovation management, marketing strategies, communication or product development.

Kurzfassung

Aufgrund immer besser werdender Entwicklungsumgebungen und Vertriebsstrukturen, steht Endanwendern ein immer größer werdendes Angebot an digitalen Systemen zur Auswahl. Es bieten sich meist mehrere Alternativen gleichen Funktionsumfangs. Somit wächst in zunehmendem Maße die Bedeutung einer intuitiven, natürlichen Handhabung. Aber was genau bedeutet natürlich oder intuitiv? Für die meisten Anwender und Entwickler stellen sich diese Faktoren als kaum messbare Empfindungen dar, wodurch es schwer wird entsprechende Kritikpunkte zu adressieren. Dieser Text gibt, basierend auf kognitionspsychologischen Studien und praktischen Beispielen, ein Verständnis für Intuition und deren Messbarkeit.

In erster Linie wird in dieser Arbeit jedoch eine neue Klasse höchst effizienter interaktiver digitaler Systeme definiert und charakterisiert, so dass solche gezielt entwickelt werden können. Nach dem Vorbild kognitiver Techniken, wie etwa der Mnemontechnik, aktivieren diese sogenannten Digitalen Transformative versteckte Nutzerpotentiale indem sie den ursprünglichen funktionalen Kontext in einen effizienten Benutzungskontext überführen.

Zu Beginn der Arbeit wird ein definierendes Schema herausgearbeitet. Dieses wird nachfolgend auf Basis evidenzbasierter kognitionswissenschaftlicher Erkenntnisse an praktischen Beispielen validiert und um ein Modell effizienzsteigernder kognitiver Mechanismen in menschlicher Kommunikation erweitert. In einem weiteren Schritt wird aus diesem Modell nachfolgend ein Konzept abgeleitet, welches die Wirkungsprinzipien Digitaler Transformative darstellt. Wesentlicher Bestandteil ist die Verwendung Kognitiver Prototypen (Kognitive Schemata), welche in diesem Zusammenhang genauer untersucht werden.

Weiterhin werden funktionale Charakteristika sowie Design Richtlinien erarbeitet, welche allgemein auf Mensch Maschine Interaktion übertragbar sind. Ferner werden Evidenz basierte kognitive Methoden, speziell auf die Thematik der Effizienzsteigerung dargestellt. Die Form und Zusammenstellung der Darstellung ist speziell auf den Anwendungsbereich dieser Arbeit ausgelegt. Auch angrenzende Gebiete wie Tangible User Interfaces, metaphorische Benutzungsschnittstellen, Transitionale Objekte, Persuasive Technologies und vergleichende relative Bewertungsverfahren für Benutzerevaluationen werden in diesem neuen Kontext beleuchtet. Nach einer Abschließenden Bewertung, werden die Bedeutung und Chancen sozialer Netzwerkanalysen zur Bestimmung geteilter kognitiver Schemata in den verschiedensten Anwendungsfeldern wie Innovation, Marketing, Kommunikation oder Produktherstellung kurz herausgestellt.

Foreword

How to read this text

The following text combines insights from the research fields of computer science, cognitive psychology, developmental psychology, and learning. Therefore, explanations are comparably comprehensive, trying to address also non-expert readers of complementary domains. For example, psychologists might be familiar with most experiments and concepts described in the cognitive sections of this text, and may skip the detailed executions, while the same information could be very helpful for readers from the field of computer science.

Consequently, all sections, chapters, and the text as a whole provide conclusions, referring to each other, hierarchically giving a top-down view on this work. Based on this structure one may read the text on demand, starting with the concluding chapter 6 on page 253. The conclusions refer to passages with more detailed descriptions on chapter level, which further refer to more detailed sections. This way known information can be skipped easily.

In order to understand the evolutionary path of this achievement, it is better to read the text sequentially from beginning to end. This way the methodological steps to the final results become more apparent.

1 INTRODUCTION 1

1.1 INITIAL IDEA AND RESEARCH QUESTIONS 3

1.2 RESEARCH QUESTIONS 4

1.3 RESEARCH METHODOLOGY 4
1.3.1 Hypothesis-Feature Graph 6

1.4 OUTLINE 7

1.5 BASIC SCHEMA OF DIGITAL TRANSFORMATIVES 7
1.5.1 Performance and context 7
1.5.2 Basic Schema of Digital Transformatives 8

1 Introduction

Imagine you were given a marble sphere with a random pattern of thousands of dots on it, similar to the one illustrated in Figure 1-1. You were asked to map this dot pattern onto a white sheet of paper, only by using a ruler and a pencil. How would you approach this problem if you were not allowed to use any further physical aids?

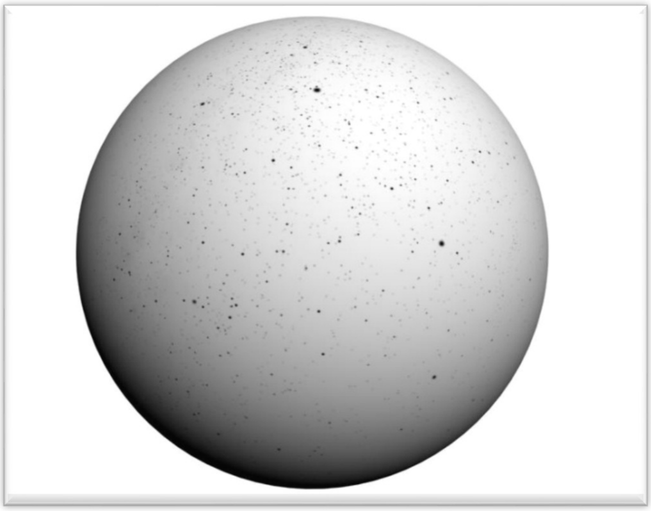


Figure 1-1. Random dots on a sphere.

Brute force, one could start with an arbitrary dot, measure the distance to the neighboring dots, estimate their relative direction, map them onto the sheet of paper, and iteratively proceed with all adjacent dots. However, one might easily loose track due to the infinite nature of the sphere. Even if we take a more sophisticated approach by estimating the surface of the sphere, first, and then determine the scale ratio to the size of the paper, we run into similar problems. Although we have the size of the surface, we would still be missing orientation. Compared to the sheet of paper the sphere does not have a well-defined outline and thus has no clear direction. There is no beginning or end, no left or right border. Every dot could be North Pole, South Pole or the center. Statements such as “dot A is left of dot B” would always be ambiguous.

However, the stated problem has probably been subconsciously solved by anyone of us multiple times before, just in a different subject domain. Let us imagine, Figure 1-1 would not show a marble sphere, but a hollow spherical aperture mask, where every dot is a hole. We were sitting inside the sphere seeing reams of dots with light falling in, just like a star sky’s firmament (such a sphere is illustrated in Figure A-1 on page 292). The solution becomes apparent now. In the star sky, by using constellations, we make use of a simple but powerful trick. Instead of working on an abstract set of dots, we concretize their arrangement, and give meaning to certain patterns by associating familiar shapes with them. These patterns are easily recognized, remembered, and serve as a reference system for orientation.

The given example showcases the power of our mind's imaginary and associative abilities. The effect demonstrated above, and the use of constellations, can be described as a mnemonic device. As defined in (Wikipedia.org, 2012f) "*Mnemonics rely on associations between easy-to-remember constructs which can be related back to the data that are to be remembered. This is based on the observation that the human mind much more easily remembers spatial, personal, surprising, physical, sexual, humorous, or otherwise meaningful information, as compared to retrieving arbitrary sequences.*" (Paivio, Rogers, & Smythe, 1968)

Since mnemonics work on very basic cognitive processes, we are not always aware of the commonness of using such aids in everyday life. However, their positive effect seems to be clear since several hundred years. Multiple studies have been conducted, showing performance gains through the use of mnemonic devices (Atkinson & Raugh, 1975; Atkinson, 1975; G. H. Bower, 1970, 1972; J. H. Douglas, 1987; Garcia & Diener, 1993; J. R. Levin, Levin, Glasman, & Nordwall, 1992; Raugh & Atkinson, 1974, 1975; Solso & Biersdorff, 1975; Solso, 2005).

Numerous books have been published on learning techniques for the enhancement of mental capabilities (Bolzoni, 2004; M. J. Carruthers, 1992; M. Carruthers, 2000; Spence, 1984; Voigt, 2001; Yates, 1966). Even in ancient times orators were using memory techniques, often helping in memorizing a thought-out composition of a speech (Cicero & Caplan, 1954; Crowley & Hawhee, 2004; Enos, 2005; Quintilian, 2006). Today, mnemonic techniques reach broader publicity through impressive performances achieved by memorization specialists at public venues. For example, records "*The World Memory Championships*" competition show surpassing performances, such as, memorizing a sequence of 1456 cards in one hour, 4140 binary numbers in 30 minutes, or 164 faces in 15 minutes (World Memory Sports Council, 2011; www.sueddeutsche.de, 2005).

Improving performance means optimizing efficiency. Pursuit of efficiency is an intrinsic driver of human behavior and can be a key selling point of digital systems. Mnemonics trigger performance improvement within their applicants. Consequently, mnemonic methodologies hold high potential for improving the design of human machine interfaces, if their working principle is also applicable for digital systems.

This observation immediately raises the question whether such transformation processes can be aided digitally? Moreover it would be interesting to analyze the basic principles behind mnemonic devices, and utilize this knowledge for the creation or enhancement of existing digital systems. This way, one might extract features, or even tools, which help with the creation of new kind of user interfaces for activating superior user efficiency. Analogous to mnemonic devices, efficiency would be released intrinsically, within the user, by building on cognitive transformations.

1.1 Basic Idea

The following work will investigate the basic idea of designing interactive systems, which release user potentials by transforming the usage context. The usage context is dependent on the system user interface. The interface provides access to system functions. Oftentimes user interface and system functions are closely related. One might think of changing heat with a radiator valve, using pedals to accelerate a bicycle, or locking a door by turning a key. The close relationship between interface and function of the given examples might have mechanical reasons (compare Figure 1-2 (left)).

Today, many systems offer a digital connection between interface and function. A good example is given by *Fly-By-Wire* technology used in airplanes. Traditionally, the control stick is mechanically connected to the wings. With *Fly-By-Wire* the mechanical movements of the control stick are digitized and communicated electronically to actuators (schematized in Figure 1-2 (middle)).

A digital connection uncouples the interface from functions. This gives more flexibility for connecting system functions to high performance usage contexts, in order to release user intrinsic potentials in a way mnemonics do. Such transformations are conducted logically, at comparably low cost, through digital computing units (depict in Figure 1-2 (right)).

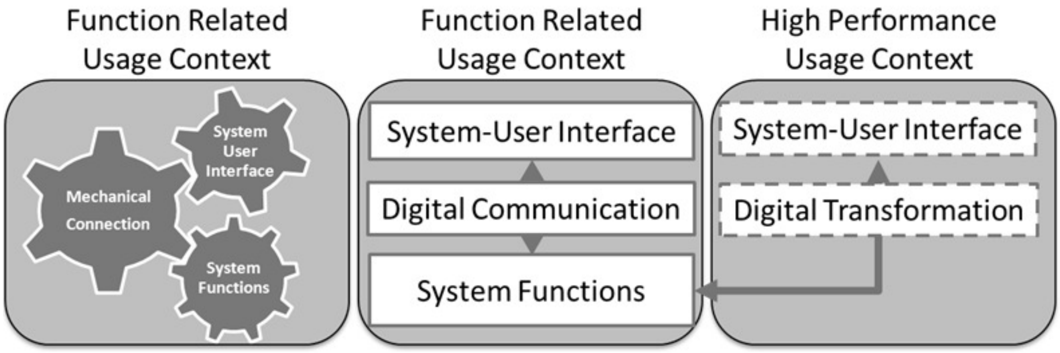


Figure 1-2. Connection between system functions and system user interface of a classical mechanical system (left). More system design flexibility through digital connection (middle) allows for digital transformations (right).

Those systems will in the following be referred to as *Digital Transformatives*. It is the goal of this work to investigate Digital Transformatives in order to design them systematically. In a long term those investigations and related new insights might establishing this new class of interactive digital systems.

1.2 Research Questions

In short terms, the investigations seek for several basic questions to be answered:

RQ1. Is it possible to learn from mnemonics in order to improve human machine interfaces?
RQ2. Are the key working principles of mnemonic devices applicable through human machine interfaces (HMI) of interactive systems?
RQ3. Are there methodologies for systematically applying the working principles, in order to foster the creation of such enhanced systems?

1.3 Research Methodology

The basic idea of Digital Transformatives (DT) seems relatively simple; however, the subject domain is highly complex. Many cognitive processes in our brain are still not completely understood. Most physical or chemical processes in our brain can be recognized, but do not provide access to actual information. Thus, Digital Transformatives are explored from an evidence based perspective on user behavior.

Since any subject related to human behavior is of rather empirical matter, a pragmatic approach is taken, trying to narrow down the key problem area through a consecutive series of iteration cycles. The investigations of this work follow the procedure of a typical iterative process (Dix, 2004; Preim, 1999; Rogers, Sharp, & Preece, 2011; Sommerville, 2001), outlined in Figure 1-3.

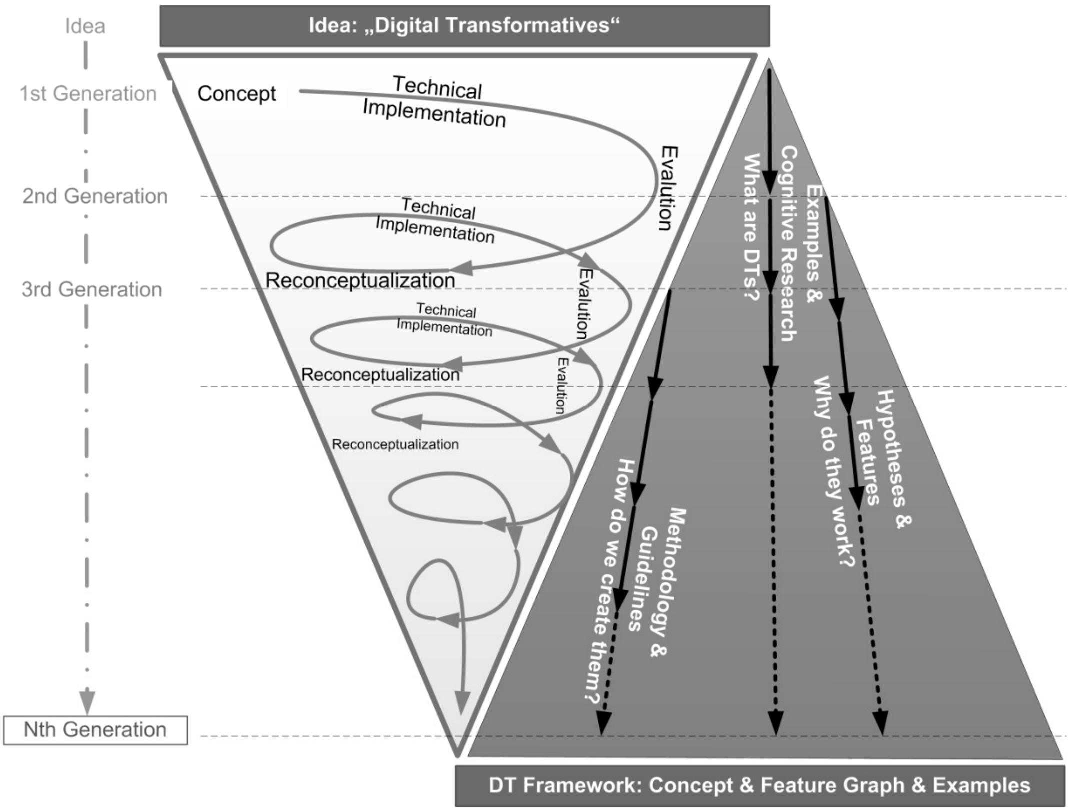


Figure 1-3. Schematic view of the iterative research methodology of this work. A reduction of the fuzziness of the initial idea correlates to increasing elaboration of the conceptual framework.

The basic idea provides a vague conceptual ground, in the beginning. Throughout various iterations of conceptualization, implementation and evaluation, it will be refined towards a sound conceptual and practical framework. While the fuzziness of the basic idea is reduced the defining framework is refined and broadened.

The first iteration is based on existing examples and evidential findings in cognitive research. In this phase an initial concept is developed and validated. The initial concept should allow for identifying Digital Transformatives ("What are Digital Transformatives?"). The second phase mainly aims at determining the working principle and functional characteristics of Digital Transformatives ("Why do they work?"). In the third phase those functional attributes will be further validated, extended and complemented by design methodological aspects and guidelines ("How do we create DTs?").

The framework given at the end of this work should provide a strong conceptual basis, including a feature set, guidelines, and several use case examples.

1.3.1 Hypothesis-Feature Graph

Each step of iteration raises new hypotheses, features, and guidelines. Validated hypotheses may be transformed into features. Hence, during this thesis a catalogue of most relevant features, for determining and using DTs, and guidelines, for designing them, is being developed. Iterations will go on, until most relevant hypotheses are satisfyingly validated. The process is started with the expression of the fundamental feature for Digital Transformatives:

Feature 1. Digital Transformatives aim for superior user performance by shifting the usage context.

This feature is assumed to be true if the following hypothesis is true.

Hypothesis 1: The system usage context shift of Digital Transformatives releases user intrinsic potentials.

This fundamental hypothesis is being tested comprehensively throughout the work.

Schematically the relation between the initial feature and hypothesis can be expressed in a hypotheses-feature graph.

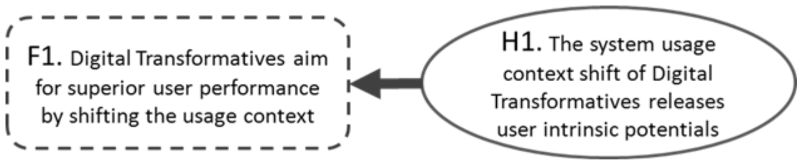


Figure 1-4. Initial hypotheses-feature graph.

Feature 1 is not validated at this moment, indicated by the dashed outline. The arrow shows a validation dependency. It is being assumed that DTs are able to aim for superior user performance (F1) if a system usage context shift releases user intrinsic potentials (H1). The graph is refined and further extended in the following. A final overview of all features and hypothesis is given in chapter *Appendix B - Hypotheses,* Features, and Guidelines (page 294).

1.4 Outline

The investigations on the above hypothesis will start with a basic schema of Digital Transformatives. This definition allows for exploring the background and related work of digital Transformatives, both, from a cognitive and from a practical side, as elaborated in chapter 2. In chapter 3 those insights will be used to refine the basic concept, and for presenting a more distinct concept of the working principle. After a first concept validation, systematic design methodologies for Digital Transformatives are investigated and developed in chapter 4. Chapter 5 describes major use case prototypes, which were designed during the iterative research process throughout this thesis. Chapter 6 provides a condensed summary of this work, and all results. Finally, the potential of the described findings beyond digital system design is detailed, a critical revision is given, and possible future work is described.

1.5 Basic Schema of Digital Transformatives

In this chapter the initial idea of Digital Transformatives will be further refined towards a distinct definition. First, the terminology of context and performance will be detailed. Afterwards a fundamental schema will be offered, which helps identifying Digital Transformatives and differentiating them from other systems.

1.5.1 Performance and context

Initially, the understanding of some basic terms needs to be clarified. The term *context* is used differently in various backgrounds. In this work, it will be used based on a definition given by (Ungerer & Schmid, 2006), who distinguish between context and *situation*. Situation is defined as interaction between real world objects, while context relates to the cognitive conceptual representation of such situations.

In this work, if not explicitly specified differently,

> *context refers to the usage situation created by a digital system. It includes the user's real world situation and the cognitive context, or perception induced by this situation.*

Moreover, if not explicitly specified differently,

> *performance refers to actions conducted by users. These actions may not explicitly be of physical nature, but also comprise cognitive activities.*

1.5.2 Basic Schema of Digital Transformatives

Mnemonic devices follow a long tradition. Some of the currently used techniques date far back, to a time, were ancient Greek orators utilized them for holding free speeches (Cicero & Caplan, 1954; Enos, 2005; Quintilian, 2006). (Grey, 1756) describes memo techniques as “[t]he Design of which is not to make the Memory better, but Things more easy to be remember’d”. Other great minds of that time see mnemonics as a method to remember certain thoughts by associating them with other, already known thoughts (da Signa, 1892; Kant, 1800; Kästner, 1805; Voigt, 2001). The basic understanding of how mnemonics work has not changed much since then. Nowadays, researchers describe mnemonic devices as methods, which help structuring information during the encoding phase, to enhance the storage and recall of information in memory. Mnemonics are seen as specific methods for information encoding and decoding, improving the suitability of certain problems (Becker-Carus & Herbring, 2004; Solso, 2005).

They may be compared to mathematical operations such as the *Fourier Transformation*, which provide a ground for solving complex cases with simplified arithmetic. Like mathematical transformations, mnemonic devices are described through a definite set of operations, such as encoding numbers into objects, or embedding incoherent information into a narrative context. Overall performances often improve, despite the fact that applicants of such techniques take the extra transitional effort for encoding and decoding information, as visualized in the following schema (Figure 1-5).

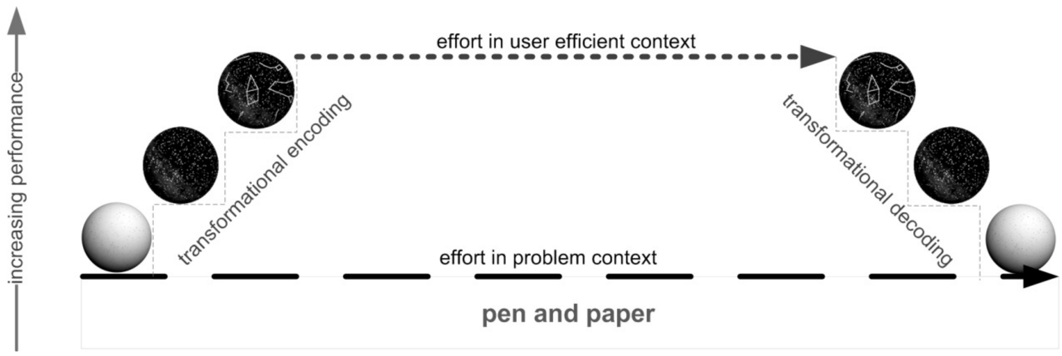


Figure 1-5. Reducing effort through information encoding and decoding as applied in mnemonic devices. The length of the arrows indicates the effort. The transformation steps of the marble sphere give a concrete example.

The base line arrow indicates the applicants’ effort in the original problem context. The top line symbolized the reduced effort in the transformed context. The applicant may either work on the original problem with bigger effort, or takes extra steps for encoding the problem into a more suitable context. Referring to the marble sphere example, given in the introductory part of this text, these steps comprise, encoding into a firmament, mapping of recognizable constellations, and decoding back into the marble sphere context.

Looking at this schema from a perspective of digital system design, it becomes apparent that most digital systems are created problem-based, according to their eventual function. For example, a function driven implementation for supporting the marble sphere problem, could provide a digital pen and paper, garnished with additional functionality for zooming, panning, rotating, and measuring. Another example we are all well aware of are number code input element, as we find it on security doors, cash machines, or in online banking forms.

Implementing number passphrases as digital keys is straightforward for a system developer, however, human users might have problems with memorizing such digital keys. Unfortunately, code cracking tools progressively advance, demanding more efficient security interfaces. Either one could implement technically more advanced mechanisms, such as eye scanners, or simply leave efficiency gains with the users by increasing the minimum amount of demanded digits. In such cases, users often take the extra transformational load to change the usage context, by transforming number codes into something more memorable, such as shapes, or dates (Herley, van Oorschot, & Patrick, 2009), as visualized in Figure 1-6.

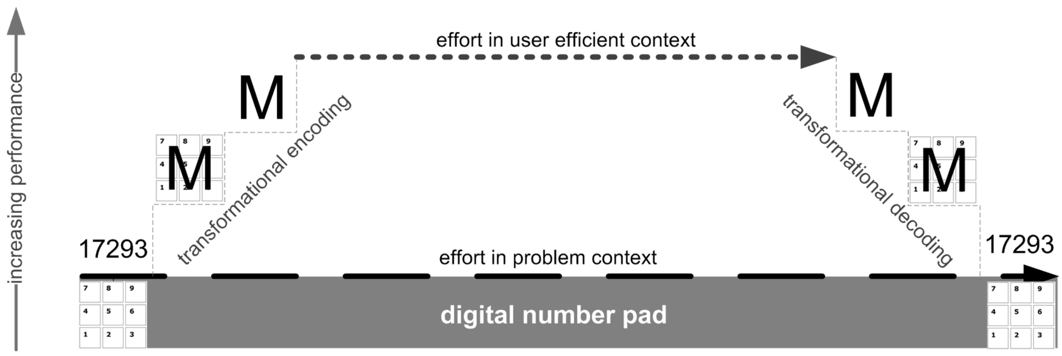


Figure 1-6. Problem based system user interface providing access to support functionality. Illustrating the task of memorizing a numeric access code.

In contrast to common problem based systems, Digital Transformatives offer a user interface in a context of improved user efficiency. This enables users to act in a context of high performance, while their actions are digitally mapped onto the original problem context. Users might not even get to know the original context. Instead of letting users apply a context change every time they use the system, an improved performance context is considered once, during the design process, and applied by the system. Therefore, Digital Transformatives need to close the gap between a performance driven user interface and the problem based action context. This is done by implementing some form of transformational encoding and decoding as shown in Figure 1-7.

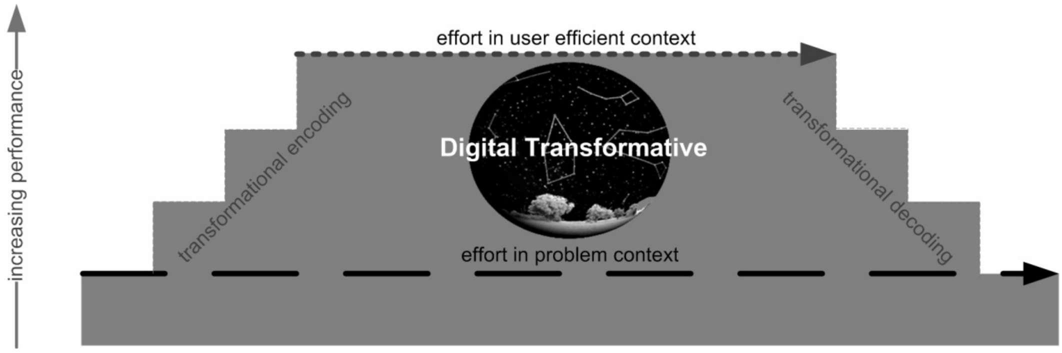


Figure 1-7. The schema of a Digital Transformative. Opposed to common digital systems, shown in Figure 1-6.

Compared to standard digital systems, DTs require an additional implementation of the definite sequence of transformational encoding and decoding operations. This way, new performance potentials can be released within the user. Such gains may even exceed those achieved through the use of mnemonics, since users perform directly in a suitable context, and do not have to burden the extra load for encoding and decoding information.

Reconsidering the examples given above, a Digital Transformative addressing the marble sphere task would provide users with a star sky map instead of a digital pen and paper interface. In case of the access code, a graphical, *drawmetric*, or *cognometic* system could improve efficiency for recalling passwords. In this example, system security may additionally be enhanced since users are able to use longer passwords (Biddle, Chiasson, & Van Oorschot, 2011; J. M. Clark & Paivio, 1991; Paivio et al., 1968). The number pad example shows that many systems already exist, meeting the above definition of Digital Transformatives. Some systems, such as the *Android Pattern Lock*, or the *Desktop Metaphor*, demonstrate high potential through commercial success (Meacham, 2013; D. C. Smith, Irby, Kimball, & Harslem, 1982). The definition given here helps with rudimentary identifying Digital Transformatives in order to analyze them comparatively, and determine critical features.

A major challenge in the design of Digital Transformatives lies in finding an optimized user action context, which allows for bilateral mapping to the problem context. Existing systems and mnemonics will be investigated in the following, in order to extract a common working principle.

2 RELEVANT COGNITIVE AND PRACTICAL BACKGROUND 13

2.1 RELEVANT COGNITIVE MECHANISMS 14
2.1.1 Mnemonics Devices 14
2.1.2 Human Communication as Transformative 19
2.1.3 Conceptual Metaphors 21
2.1.4 Concepts, Prototypes, and Categories 23
2.1.5 Performance, proceduralization, and category prototypes 34
2.1.6 Pervasiveness of Similarity Comparisons 44
2.1.7 Cognitive Efficiency Catalysts in Communication 46
2.1.8 Efficiency in Semiotics 50
2.1.9 Summary of Human Cognition in Digital Transformatives 52

2.2 RETROSPECTIVE PROPERTY EXTRACTION OF EXISTING DIGITAL TRANSFORMATIVES 55
2.2.1 General Property Extraction Procedure 55
2.2.2 Determine Relevant Digital Transformatives 58
2.2.3 Cognitive Domain 59
2.2.4 User Interface Metaphors 69
2.2.5 Affective Domain 76
2.2.6 Psychomotor Domain 89
2.2.7 Affordances from UI Concepts 99
2.2.8 Summary of the Retrospective Property Extraction 101

2 Relevant Cognitive and Practical Background

In this chapter, the key working principle will be investigated, by combining evidence based research on relevant cognitive mechanisms with an epidemiological property extraction method of existing Digital Transformative (DT) systems. Figure 2-1 gives an overview of the procedure, as well as involved mechanisms and systems.

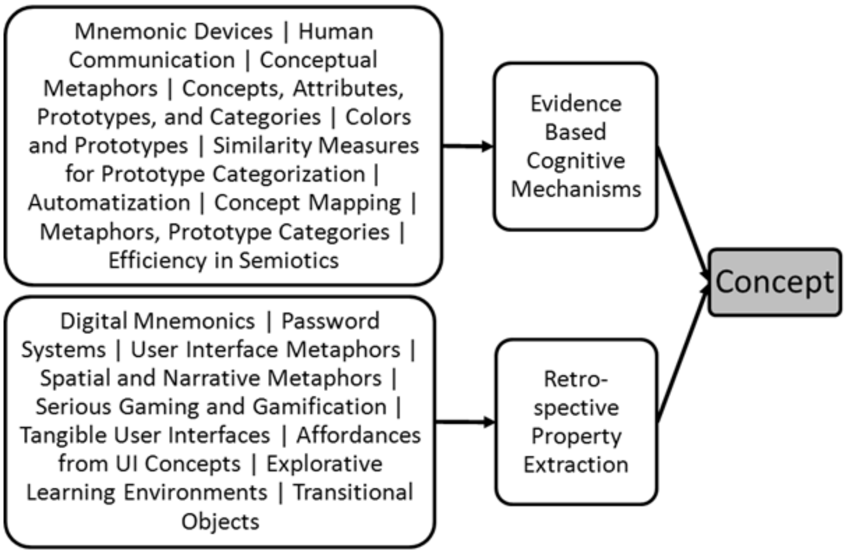


Figure 2-1. Fundamental cognitive (top-left) and practical investigations (bottom-left) for developing a Digital Transformative concept.

It is started with an analysis of evidence based findings, important for understanding the cognitive mechanisms addressed by Digital Transformatives, in section 2.1. Therefore, the cognitive background of Mnemonic Devices is detailed first. Afterwards analogies between human communication and the schema of Digital Transformatives are outlined, looking for transferable insights from language and cognition research. Such cognitive fields include conceptual metaphors, concepts, attributes, cognitive prototypes, categories, process automatization, and semiotics.

In Chapter 2.2, the cognitive findings are evaluated against practical examples in an epidemiological property extraction procedure. Therefore existing systems following the Digital Transformative schema are examined. Those include User Interface Metaphors, Serious Games, explorative playful environments, Transitional Objects, Tangible User interfaces, and Affordances of User Interfaces. The findings of the cognitive and practical investigations are combined into an elaborate concept, describing the working principle of Digital Transformatives, in chapter 3.

2.1 Relevant Cognitive Mechanisms

In this chapter, related evidence based research on mnemonic devices and associative cognitive tools will be revised. The revision provides a selective view on cognitive processes. Hereby it should be noted that the structure and relations among the selected processes have been deduced by the author from cognition literature, and hence may vary from typical presentations in this field.

Starting off with a historic view on mnemonics, it will be further elaborated on conceptual metaphors, which are based on categorization, cognitive prototypes, and automatization, as depict in Figure 2-2.

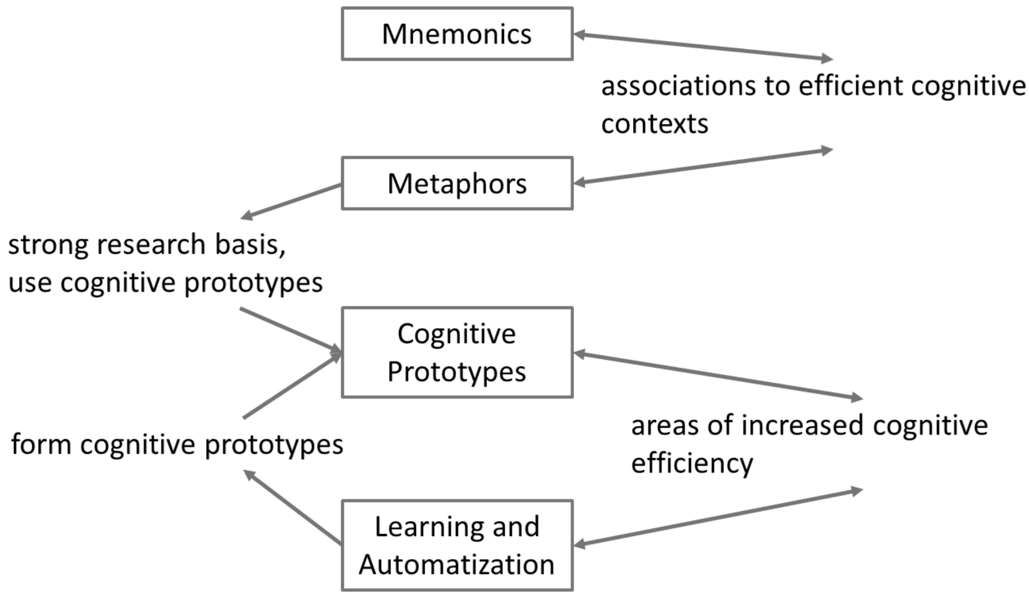


Figure 2-2. Cognitive mechanisms related to Digital Transformatives.

Mnemonic devices work by offering associations to efficient cognitive contexts. Conceptual metaphors use the same mechanism, additionally, providing a broad linguistic and cognitive research basis, for determining and investigating the fundamental working principle. It will be emphasized how efficiency gains of metaphors are mainly induced through cognitive prototypes, which are also related to automatization and learning processes. Finally, further efficiency improvements through cognitive prototypes, similarity comparisons, and categorization will be examined on semiotics.

2.1.1 Mnemonics Devices

Mnemonic devices follow a long history of phases alternating between academia and entertainment. Latest waves of academic interest were in particular based on three antique references from, Cicero (55 B.C.)[1], Quintilian (95)[2], and Rhetorica ad Herennium (28 - 40 B.C.)[3] (Cicero & Caplan, 1954; Enos, 2005; Quintilian, 2006). Around thirteen hundred it was Albertus Magnus, trying to build a scientific fundament for mnemonics. Around seventeen hundred Döbel (1707)[4] and Schenckel elaborated on mnemonics. Kästner, Aretin, Paris, Reventlow, Kothe, and Feinaigle had a big impact for another wave of interest on mnemonics, in the 19th century. However, mnemonics never ought to be recognized as a scientific domain, which was regularly taught at universities. Nowadays, we are speaking of the "art" of mnemonics rather than its methodology. Mnemonics are interesting for entertainment more than for science (Voigt, 2001, p. 25 ff).

The heterogeneous history might be a reason for the absence of a well-defined concept description of mnemonic devices. Most definitions are rather abstract. For example, Alsted (1610) defined them as any device fostering memorization. (Helvetius, 1758) believes that mnemonic devices improve recall abilities by bringing objects in order, while (Kästner, 1804)[5] sees their major advantage in connecting new ideas with associated known ideas. (da Signa, 1892) uses the terms *artificial memory aids* (subsidium artificiale) and *signs for memory* (signum memoriale). He sees such signs everywhere around us, where something is manifest, concrete, or just remarkable to serve as a sign for memorizing. In his view paintings, statues, memorials, bell towers, pillories, concisions, or even the act of anointing someone with oil are examples to define signs for memory. Voigt (2001) critically remarks da Signa's definition, stating that mnemonic devices always relate to the applicants themselves. However, Voigt (2001) builds on the classical views given by (Dommerich, 1765)[6], (Kant, 1800)[7], (da Signa, 1892), and (Kästner, 1804), defining mnemonics as follows: mnemonics "[...] give access to new or hardly memorized information by associating them to something known or easily memorized. The known information is utilized as

[1] Cicero (55 B.C.) – de Oratore

[2] Quintilian – (A.D. 95) Institutio Oratoria

[3] written by an unknown author

[4] Döbel – (A.D. 1707) Collegium Mnemonicum

[5] „Alle Regeln, die sie [die Mnemonik] vorträgt, laufen in der einigen zusammen: Verknüpfe eine Idee mit einer anderen, an der du jene stets hebeyzuziehen im Stande bist" (Kästner, 1804, p. 9)

[6] „Man verbinde die Sache, die man behalten will, mit anderen Vorstellungen, so geben diese insgesamt mnemonische Mittel ab" (Dommerich, 1765, p. 57,58)

[7] "eine Methode gewisse Vorstellungen durch Assoziation mit Nebenvorstellungen dem Gedächtnis einzuprägen" (Kant, 1800, p. 94)

a tool to memorize the desired information.“ (Voigt, 2001, p. 36)[8]. He substantiates this definition with a model which follows the idea of building bridges in form of memory anchors. A schematic illustration is given in Figure 2-3.

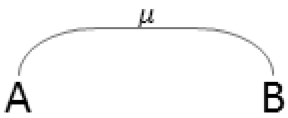


Figure 2-3. Classical Schema of Mnemonic Devices after (Voigt, 2001).

Hereby, he defines three variables:

- A - the content to be remembered (“Erinnerungsinhalt”),
- B - the aid to remember the content (“Erinnerungsstütze”),
- µ - the Association (“Verknüpfung”) between A and B

This schematic concept describes the typical classical interpretation of the rudimentary working principle of mnemonic devices. It provides a common design schema of mnemonic devices, which can be found in most applications, as illustrated in an example given by (Beniowski, 1842). In his botanical studies at university Beniowski had to memorize Latin names for several plants, however, he had big problems with learning and assigning such names. On the other hand, he was wondering, why he could easily memorize nicknames of friends. Beniowski also observed an association between nicknames and a certain appearance or behavior. For example, some of his friends were called Long Cloak, Old Boot, or Big Nose. Hence, he decided to transfer this knowledge onto the problem domain for learning Latin plant names. Therefore, he was looking for a connection between the appearance of the plant (µ) and its’ Latin name (A) with a nickname (B).

If such plants were human like friends of him, they would get meaningful nicknames according to their appearance, or some remarkable behavior. For example, the plant officially called Achillea Millefolium had the look of a roof covered with snow. The first part of its Latin name sounds like Aquila – the Latin word for eagle – mille stands for thousand, and folium is the Latin word for leaf. From here he draws the following connection between those two associations: in his imagination an eagle ap-

[8] Translated from German: „*Die Idee ist, das Neue oder Schwierige durch eine Bindung an Bekanntes oder leicht Faßbares beherrschbar zu machen, das, was wir im Gedächtnis behalten wollen, mit etwas anderem zu verbinden, um dieses andere sodann als Mittel zu benutzen, das Gewünschte zu erinnern“* (Voigt, 2001, p. 36)

pears in a setting of high snow covered mountains, which is also true for houses with snow covered roofs (visualized in Figure 2-4).

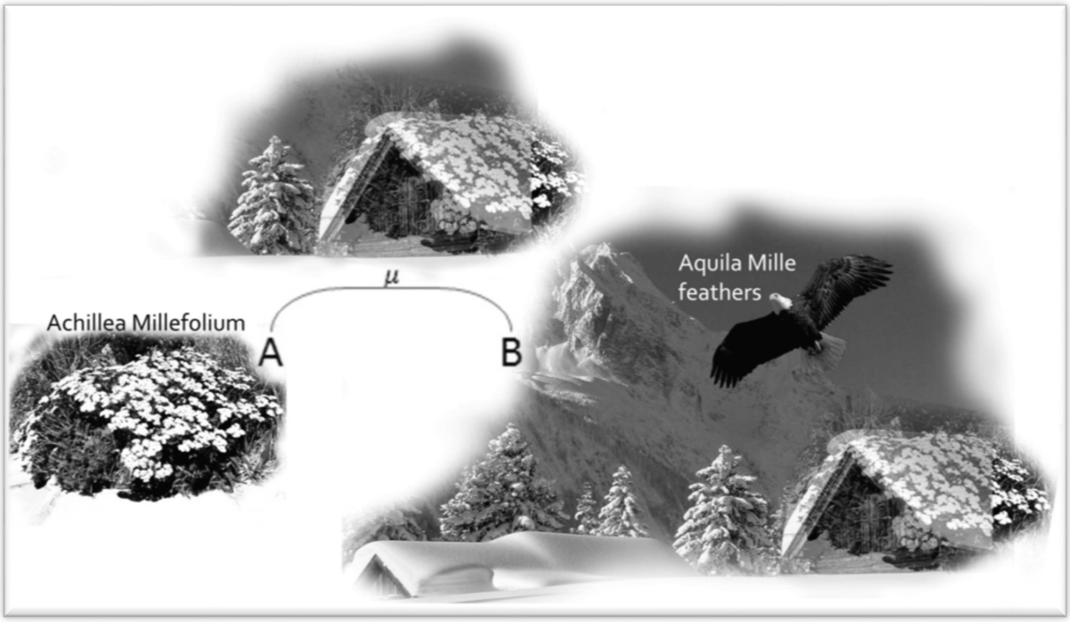


Figure 2-4. Achillea Millefolium – for Beniowski looked like a house covered with snow. The Eagle (Aquila) with thousand (mille) feathers / leafs (folium).

When Beniowski sees the plant he remembers its nickname, snowy roof, from its look. The snowy roof raises images of mountains, and opens the scenery of an eagle with thousand feathers gliding in between the mountain tops. From the eagle with thousand feathers he deduces the Latin words "aquila mille folium", which brings him to the name of the plant "Achillea Millefolium", because he already knows the Latin words for eagle, thousand, and leafs.

This way, Beniowski created a chain of associations, which he was able to follow in both directions, from the name to the appearance of a plant, and back. He used the same technique to remember all other plants, and gained comparably great success in his class (Beniowski, 1842).

COGNITIVE EVIDENCE BASED RESEARCH ON MNEMONICS DEVICES

The classical view on mnemonic devices is well in line with findings from the perspective of cognitive research. (Solso, 2005) defines mnemonics as a technique that uses familiar associations to enhance the storage and recall of information in memory (Becker-Carus & Herbring, 2004; Solso, 2005). Nowadays, the key mechanism of mnemonics is often seen in proper organization of information, which is structured during the encoding phase, to enhance storage and recall in memory. While the conceptual basis on effects of mnemonic devices seem to be rather vague, several studies provide cognitive evidence on their advantages.

In a small scale test (Ericcson, Chase, & Faloon, 1980) analysed a subject, who used mnemonic systems to memorize digits. In most tests the subject rapidly achieved performances similar to those of experts with lifelong training. For example, his memory span was increased from 7 to 79 digits within 230 hours of practice. In their studies they found evidence that training advantages did not lead to improved short-term memory capabilities. From their view, mnemonic devices rather relieve the workload of short-term memory through a single association to already stored information, which allows easy retrieval of complex target information. In other words, complex unknown material is associated with something familiar.

(Atkinson & Raugh, 1975; Atkinson, 1975; Raugh & Atkinson, 1975) analyzed the keyword method for learning foreign vocabulary. The technique is similar to the method described by (Beniowski, 1842) for learning plants. The learner reads a foreign word such as “zronok”, which is the Russian word for bell. The last syllable of “zronok” sounds like the English word “oak”. Hence “oak” can be used as a keyword, providing an imaginary connector to the English meaning. Applicants of this method demand on creating a remarkable image with sufficient source and target elements. In this case, it could be an oak hanging full of bells instead of acorns.

In one of the experiments, test persons had to learn 40 words a day over 3 days. One group was provided with a visual representation of the keyword method, while the control group only had the English translation. The keyword group achieved significantly better. Already after 2 days of learning they memorized more words then the control group after 3 days. In a second test, which was conducted 6 weeks after the learning session, the experimental group could still remember 43% of the words, while the control group only remembered 28%. (Raugh & Atkinson, 1975) also found it was better to provide keywords, instead of letting the test persons create keywords themselves.

(G. H. Bower & Clark, 1969) studied the effect of cognitive elaboration to structure a list of words through narratives. Cognitive elaboration builds on increasing the amount of associations (retrieval cues) to facts that have to be memorized. It is assumed that facts are easier recalled the better they are connected with other information. Narratives foster interconnectivity of otherwise independent words and provide a structural frame of reference for organizing information (Becker-Carus & Herbring, 2004; G. H. Bower & Clark, 1969). (G. H. Bower & Clark, 1969) evaluated two groups of subjects on memorizing ten specific words. One group was assigned to think of a story which included all ten words, while the other group was left without any instructions. Subjects structuring information through narratives performed more than six times better than individuals not using such a mnemonic device.

CONCLUSIONS

Mnemonics follow a long history of being used as memory enhancements. Classical anecdotic descriptions on improvements are supported by more recent evidence based studies. The findings and explanations of the effect correlate between classical and recent views. Mnemonic devices make heavy use of associative abilities of our brain. Instead of memorizing "plain" information, we increase meaning by associating new complex information with proper known information, giving new information a familiar structural frame. Cognitive elaboration also seems to be advantageous, for memorization, by increasing interconnectivity of knowledge.

Increasing interconnectivity of knowledge, in general, and associating new information with existing knowledge, seem to be major mechanisms for intentionally improving recall abilities of selected information. However, mnemonics are constructed with care, in order to work. Simply increasing cognitive elaboration or connecting new information to random knowledge is not necessarily sufficient. For utilizing and transferring this mechanism from the domain of memorization to the broader domain of user system interaction, it is necessary to get a more fundamental understanding of the working principle. Therefore, in the following, it will be further detailed on cognitive mechanisms which give hint on,

- why some associations are better suited than others, and
- how cognitive elaboration affects performance.

2.1.2 Human Communication as Transformative

The Digital Transformatives schema for human computer interaction, as introduced earlier, also serves as a general model for communication. Figure 2-5 shows how the schema relates to human communication.

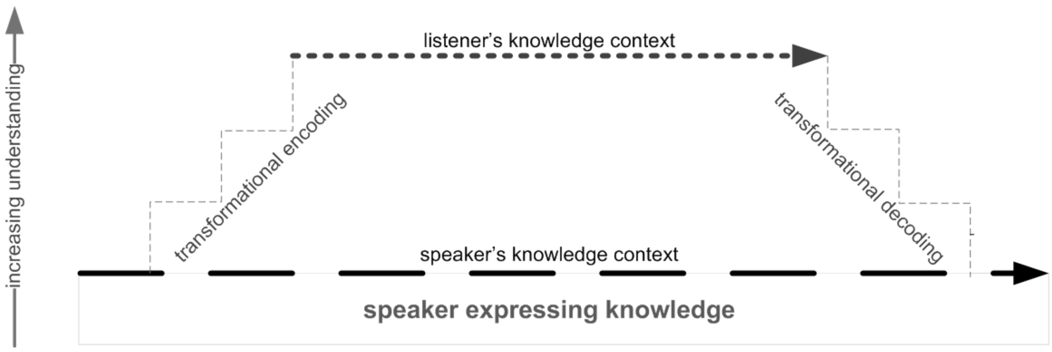


Figure 2-5. The Digital Transformatives schema transferred to human communication.

In human communication, information is expressed based on the speakers' knowledge, and understood based on the listeners' knowledge. The interpretation of information depends on the knowledge context. In the age of five, we understand the same information different to when we are fifteen, or twenty-five years old. Hence, understandability is increased if information is transformed into a knowledge context common to most listeners. The same is true for human computer communication, where humans need to exchange information with computers, for example via graphical user interfaces.

The basic similarity between human-human and human-computer communication indicates possible analogies. Thus, the research field of cognitive linguistics offers a great empirical and conceptual ground for further investigations on the idea of Digital Transformatives. The transformational schema, shown in Figure 2-5, can also be used to describe linguistic constructs, such as metaphors or similes. In order to communicate information more efficiently, and increase understandability, a context shift is induced by the speaker, transforming the original context into a similar one, which is thought to be better understood by the listener.

Starting from conceptual metaphors, such linguistic constructs will be elaborated in the following. Additionally, knowledge about underlying cognitive processes provides an evidence based frame and helps with validating later findings regarding Digital Transformatives.

2.1.3 Conceptual Metaphors

"*Having lots of ideas doesn't mean you're clever, any more than having lots of soldiers means you're a good general*"[9] (quote of Sebastien-Roch Nicolas De Chamfort (van Bever & others, 1923, no. 446; Wikiquote.org, 2012)). Everyone of us knows, and uses metaphors in various context in daily life. As the above example shows, they open comparative perspectives, which might be hard to find but are surprisingly valuable. (Lakoff & Johnson, 1980) show the pervasiveness of metaphors through numerous examples and metaphor schemata. Accordingly, metaphors are of fundamental nature, not just in language, but also in thought and action (D. Gentner, Bowdle, Wolff, & Boronat, 2001). "*The essence of metaphor is understanding and experiencing one kind of thing in terms of another*" (Lakoff & Johnson, 1980, p. 5).

Metaphors typically are considered to be a rhetorical instrument. For example, we talk about ideas as if they were objects. We "have" ideas, "lose" or "find" them, or just can't "get" ideas out of our mind. As another example, emotion is mapped into a form of motion: one is "moved by a poem", or "went into transports of joy"(Lakoff & Johnson, 1980; Lakoff, 2012).

In linguistics, metaphors work on a level of comparison. They are closely related to *similes* or *analogies* by working on items that share primary attributes (D. Gentner et al., 2001). For example, if we say a person is like an elephant, in our mind the person inherits primary attributes, such as having thick skin, being robust, and stable. An elephant also never forgets, or we consider elephants to be slow and inflexible.

(Leech, 1969) describes the key items involved in a metaphor as the *tenor*, the *vehicle*, and the *ground*. In the former example, the person is considered the tenor, inheriting attributes of the elephant, which is the vehicle. The ground provides the base of comparison, given by salient similar attributes, such as having a thick skin or a good memory. The tenor is the target of our comparison, or the explained element, the vehicle is the source or explaining element (Ungerer & Schmid, 2006). Additionally, it should be noted that each metaphor also holds attributes which are dissimilar, usually not intended to be used for comparison. In the given example, this part of a metaphor, also referred to as *tension*, might comprise attributes such as being hunt for teeth, entering the period of musth, or sleeping only for two hours a day.

[9] On n'est point un homme d'esprit pour avoir beaucoup d'idées, comme on n'est pas un bon général pour avoir beaucoup de soldats (van Bever & others, 1923)

Important for this work is the effectiveness of metaphors. Hereby it can be distinguished between *inventive* unexpected metaphors, and established *conventional metaphors*, where the later supposedly ought to be more powerful and effortless. An example for a conventional metaphor is the term “head-of-department”(Ungerer & Schmid, 2006). (Lakoff & Turner, 1989) describe such metaphors to be deeply entrenched, efficient, powerful, automatic, unconscious and effortless.

According to (Strube et al., 1996), a rhetorical metaphor conducts a transfer of an expression from one subject domain to another, based on analogies or parallels between the both. A transfer only seems to be useful, when the source subject domain is known.

(Lakoff & Johnson, 1980) further elaborate on the schema of linguistic metaphors, and state that it is not just a stylistic device. Users of metaphors actually map cognitive concepts from one domain to another. Hence, metaphors can be seen as cognitive instruments, where the cognitive *source concept* of the vehicle is mapped on the target concept of the tenor. The *mapping* refers to the ground.

Not all features are eligible to be mapped. What features are eligible, for a certain *mapping scope*, is influenced by various factors. This is, some conceptual mappings eligible in one culture might not work in a different culture. Especially effortless conventional mappings are affected by such differences. This is easily understood if we think of metaphorical phrases in one language, and translate them to other languages, such as “it is raining cats and dogs”.

Eligible mappings, and the mapping scope, build on the similarity of the source and target concept. In context of this work, it is also important to find links between similarity and effectiveness. Also conventional automatized metaphors might give further hints on how metaphors increase efficiency, there is no evident study from psychology research on this matter, yet. However, a patchwork of theories and studies from different domains of cognitive psychology is available, which might provide some guidance. Such relevant theories and studies will be briefly described in the following.

Metaphors may also be seen as a species of categorization (Glucksberg & Keysar, 1990; Glucksberg, McGlone, & Manfredi, 1997; Honeck, Kibler, & Firment, 1987; Kennedy, 1990). As detailed above, in cognitive psychology the schema of metaphors can be described as a conceptual link between a source and a target concept.

CONCLUSIONS

Evidential findings on conceptual metaphors give a deeper understanding of cognitive mechanisms analogous to those involved in mnemonics. Corresponding to associative anchors connecting known information to new content, as detailed earlier in the Mnemonic Devices schema, conceptual metaphors map salient attributes from a known source concept to an unknown target concept; this relation is visualized in Figure 2-6.

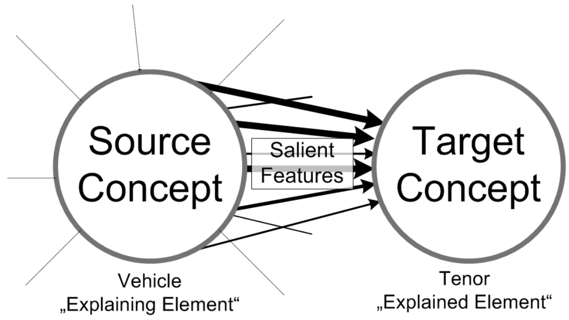


Figure 2-6. Working principle of conceptual metaphors. The source concept (vehicle) inherits most salient features to the target concept (tenor). Some features are more salient than others, indicated through different line-thickness.

Conceptual Metaphors work best if both concepts are sufficiently similar. Since the target concept inherits from the source concept, similarity is dependent on the dominant features of the source concept.

Metaphors are also often seen as a specialization of categorization. The mapping of salient features from a known source, to a new target concept increases efficiency in communication. In this sense, metaphors have different levels of efficiency. Efficiency depends on cultural acceptance and on conventionalization of a metaphor in its language. Conventionalization is reached through cognitive processes of automatization.

The next section will detail the term concept and the relations to attributes, prototypes, and categories. Later it will also be looked at efficiency gains through cognitive automatization and habituation.

2.1.4 Concepts, Prototypes, and Categories

One of the big questions still unrevealed is how in detail our brain processes and represents real world knowledge (Sternberg, 2008). Cognitive scientists, and also computer scientists in the field of *Artificial Intelligence*, investigated various concepts concerning this fundamental aspect.

In cognitive psychology, the basic theoretical unit representing symbolic knowledge is often called a *concept*. A concept could simply be the representation of the word "tree". Concepts relate to each other, thus the concept for "tree" might have a relation to "leaf", "trunk", or "plant" (JS Bruner, Goodnow, & Austin, 1956; J. Fodor, 1994; Hampton, 1997b; Kruschke, 2003; Love, 2003; Sternberg, 2008). A cognitive structure of interconnected concepts may also be referred to as *mental model* or *conceptual model* (K. J. W. Craik, 1943; Dedre Gentner & Gentner, 1982; Johnson-Laird, 1983, 2005)[10].

Furthermore, we tend to categorize knowledge. *Categories* can be seen as structures, where concepts are organized based on common *features*[11], or through similarity to a *prototype* (Coley, Medin, & Atran, 1997; Hampton, 1995; Medin, 1998; Sternberg, 2008; Wattenmaker, 1995; Wisniewski & Medin, 1994).

The feature based approach, also called *classical view* or *Aristotelian view*, for determining categories, developed out of linguistic research. At first glance, this approach provides distinct definitions of categories through a set of *defining features*. Defining features are mandatory for a concept to be part of a category (H. H. Clark & Clark, 1977; J. A. Fodor & Katz, 1963; J. J. Katz, 1972). An example is given by endothermy for mammals. By knowing that an animal is cold-blooded, we can exclude it from being a mammal. In order to be part of a category, all defining features need to be fulfilled. To give another example, a bachelor is defined by being male, unmarried, and adult.

In practice, human perceived categories often do not have such clear boundaries (Keil, 1992; Malt & Smith, 1984; Mervis, Catlin, & Rosch, 1976; E. Rosch & Mervis, 1975; Wittgenstein, 1953). (Malt & Smith, 1984) conducted a study, where subjects were asked to rate the typicality of members of various categories on a seven point scale. An oak was considered to be the most typical, while bamboo was rated be the most untypical tree. All other queried members of the category tree were rated somewhere in between, as visualized in Figure 2-7.

[10] The terms mental model or conceptual model are understood quite differently depending on the readers' background. Hence, in order to prevent misunderstandings, such terms will be used with care in this text. In general they are describing mental representations of real world occurrences.

[11] also often referred to as *attributes*, *properties* or *characteristics*

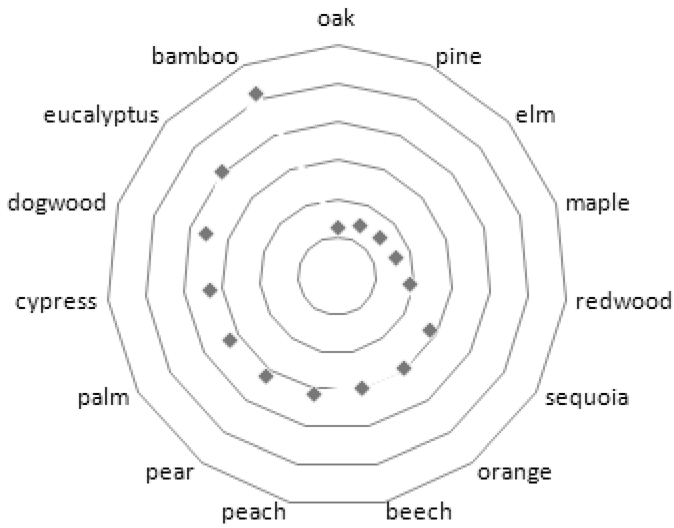


Figure 2-7. Ratings of members of the category tree, the more typical the closer to the center; adopted from (Malt & Smith, 1984).

Many more studies demonstrate the fuzziness of categories and their nature of having good and bad examples (Heider, 1971a; Labov, 1973, 1978; E. H. Rosch, 1973b). Defining features are often not sufficient for describing a category. Different approaches have been developed giving an insight into the fuzziness of our cognition.

According to *linguistic relativity hypothesis,* also referred to as *Whorfian Hypothesis,* or *Sapir-Whorf hypothesis*, we construct our understanding of the world through language (Davidoff, 2001; D. Gentner & Goldin-Meadow, 2003; Saunders & Van Brakel, 1997; Whorf & Carroll, 1956). The most famous example of this hypothesis refers to the relatively big amount of terms Eskimos use for snow (Derose, 2005; Woodbury, 1991). The hypothesis claims that the sophisticated vocabulary of snow, in their language, was a reason why Eskimos had a more differentiated understanding and perception of snow. For example, simply the existence of the Inuit word “piqsirpoq”, which means drifting snow, makes Eskimos reflect on certain characteristics of snow. Speakers of languages without a word for drifting snow may never think of such characteristics – hence not develop equivalent cognitive concepts.

While snow certainly is a matter experienced by various humans differently, perceived colors offer a much more general domain for researching linguistic effects on knowledge structuring. Based on the understanding of a uniformly distributed color space surrounding us, it has been concluded that colors should be categorized arbitrarily, as reflected in random linguistic color terms developed by different cultures. The typical green for an Eskimo should be different from the typical green for an

Egyptian (R. W. Brown & Lenneberg, 1954; Lenneberg & Roberts, 1991; Lenneberg, 1953).

In contrast, (Berlin & Kay, 1969) found that the color continuum in human perception is structured by a universal reference system. Their test consisted of two major steps. In a first test they analyzed so called *basic color terms*, which comprise most salient colors commonly known in a culture. Those colors were determined by asking for the smallest set of terms for describing every color of our spectrum (Berlin & Kay, 1969). For example, in English the rainbow colors red, orange, yellow, green, blue, and violet might almost be sufficient (Saunders & Van Brakel, 1997). In their first survey Berlin and Kay (1969) tested 98 genetically diverse languages, of which 20 were analyzed through interviews, and the rest was based on grammatical and written materials analysis[12]. They found that some ethnic groups, such as the Dani people from Papua New Guinea, only differentiated between warm and cold colors (Heider, 1971a). Other ethnos require up to eleven terms to describe the whole color spectrum. Berlin and Kay (1969) further investigated the different color hierarchies among the investigated languages, as comprised in Figure 2-8.

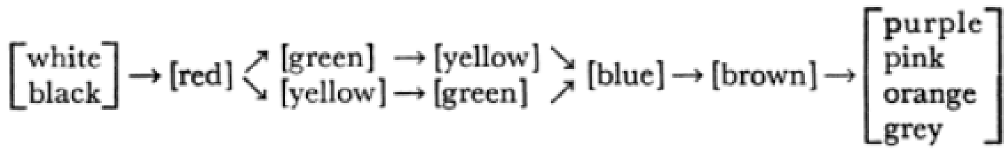


Figure 2-8. Focal color hierarchy determined by (Berlin & Kay, 1969). This hierarchy was later revised, modified, extended and transformed into a set of categories (P. Kay, Berlin, Maffi, Merrifield, & others, 1997).

If a language had three basic color terms, then those three terms were "black", "white", and "red". Languages with four color terms additionally included "green" or "yellow", language with five basic color terms usually have both terms. An optional sixth differentiation would be given by "blue", and a seventh by "brown". Languages with eight or more terms also contain "purple", "pink", "orange", "grey", or some combination of it.

In a second step, after the color terms were elicit, the instructor presented a board with various colors of the human perceptive visual spectrum (Wyszecki & Stiles, 1967). The board consisted of 329 standardized Munsell color reference chips, equally

[12] The twenty languages analyzed through interviews: Arabic (Lebanon), Bulgarian, Catalan, Cantonese, Mandarin, English, Hebrew, Hungarian, Ibibio, Indonesian, Japanese, Korean, Pomo, Spanish (Mexico), Swahili, Tagalog, Thai, Tzeltal, Urdu, and Vietnamese (Berlin & Kay, 1969)

distributed over 40 hue and 8 brightness levels, plus additional 9 gray scale chips, similar to the board shown in Figure 2-9 (Cook, Kay, & Regier, 2005; Richard Cook, Paul Kay, & Terry Regier, 2012). Participants had to go through all formerly determined basic color terms, and mark every chip meeting the term under any condition. This procedure helped finding the boundaries of each color. It turned out that the boundaries were highly unreliable, and varied not just among languages or informants, but also among different trials of the same individual.

Additionally, Berlin and Kay asked for the best example for each color, thus, for a chip showing a typical red or yellow. These results were much more distinct. In repeated trails, selected best example chips rarely exceeded an offset bigger than two. Trails were repeated three times with a full week in-between. Moreover, such so called color foci did not vary more between speakers of different languages than between speakers of the same language – indeed the tests even showed a slightly higher deviation among speakers of the same language. Figure 2-9 shows the focal color areas of all tested languages.

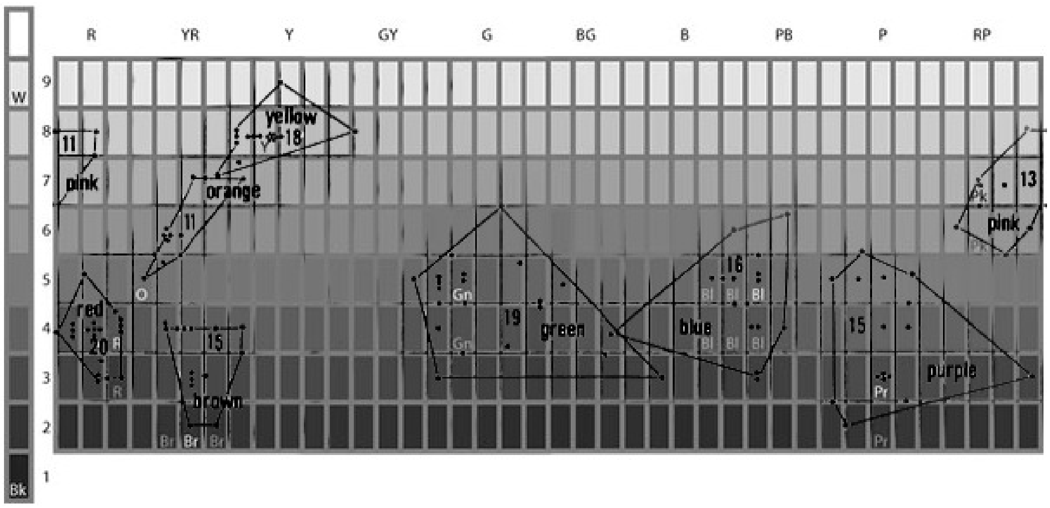


Figure 2-9. Focal colors distribution, visualized on a Munsell Grid, determined through the first experiments conducted by (Berlin & Kay, 1969). Visualization adapted from (Fred Hatt, 2011); also compare color chart provided by (Richard Cook et al., 2012).

Every dot represents the weighted average of all informants of one language. Numbers inside the foci boundaries indicate how many languages encode the corresponding color category.

Berlin and Kay's findings on the universality of human basic color terms were highly influential[13], and largely accepted by psychologists and vision researchers. Their findings led to a series of still ongoing experiments, theories about categorization, and implications on perception, recognition, knowledge organization, and knowledge representation (Ungerer & Schmid, 2006). The importance of this work, for understanding mechanisms of categorization, evolved from the concept of seeing *focal colors* as cognitive reference points (E. Rosch, 1975a). This view is supported by earlier investigations of (Wertheimer, 1938), who proposed the existence of "*ideal types*" among real world stimuli, serving as anchors for perceptional reference. (E. Rosch, 1975a) further found similar reference points in other domains, such as numbers, lines, shapes, organisms, and objects (E. H. Rosch, 1973b; Ungerer & Schmid, 2006). While in the first phase of focal color research those reference points were called *foci,* Rosch's further research led her to call them *prototypes* (Ungerer & Schmid, 2006).

The idea of prototypes is eminent for Digital Transformatives, because they may mark the sweet spots of human performance. Due to the broad basis of research on cognitive prototypes in human color perception, further investigations will be detailed by the means of color perception.

The claim of universal basic color terms seems to be controversial to findings of feature based prototype composition, since it implies that the perceived color distribution would be the same for everyone, at any place in the world. Discussions on the universality have been intensified during the past decade (Cook et al., 2005; Davidoff, Davies, & Roberson, 1999; Regier, Kay, & Cook, 2005; Ungerer & Schmid, 2006).

One of the major critics regarding the universality of focal colors, relates to the informants used in the experiment. As described in (Berlin & Kay, 1969), for multiple languages the focal color value is based on the data of only a single surrogate. Although the informants were all native speakers, except of one ethnic group, they all resided in San Francisco Bay area; all of them also spoke English and were from industrialized countries. This might explain why the focal colors of informants from different ethnical groups in average even had less dissimilarity than the ones determined by informants of the same language. By living in a similar environment, they were exposed to a similar color distribution (compare Berlin & Kay, 1969). Further experimental flaws of this study are depict in (G. A. Collier, 1973; Conklin, 1973; Hickerson, 1971). A comprehensive multi-disciplinary overview refusing the universality from various perspectives is given by (Saunders & Van Brakel, 1997).

---

[13] According to the academic search Google Scholar the work of (Berlin & Kay, 1969) has been cited approximately 3600 times as to July 2012 (http://scholar.google.de/scholar?-q=Basic+color+terms%3A+Their+universality+and+evolution)

The doubts in the universality of basic color terms are even underlined by the succeeding World Color Survey (WCS). The much more elaborate WCS aimed at overcoming the weaknesses of the initial study (P. Kay, Berlin, L. Maffi, Merrifield, & Cook, 2009; Paul Kay & Cook, 2011; Richard Cook et al., 2012). Beginning in 1976, fieldworkers all over the world collected color naming data from speakers of 110 unwritten languages, of 45 different families. Analogous to the first study, basic color terms were investigated, color boundaries determined, and focal color categories evaluated. In 1980 all data was collected, in 1991 raw data was cleaned up. By the year of 2003 the preparation of all data reached a stable state, and the first portion of it was made public (Cook et al., 2005; Paul Kay & Cook, 2011). The results of the WCS in general support the results of the first study. However, instead of distinct basic color term boundaries, the WCS revealed a more divergent distribution of focal colors. Figure 2-10 shows the accumulated color term distribution, based on colors evaluated by each test person. The centers of mass of each subjects evaluation, the so called *term centroids*, were transformed from the Munsell color system into CIE L*a*b coordinates according to (Wyszecki & Stiles, 1967).

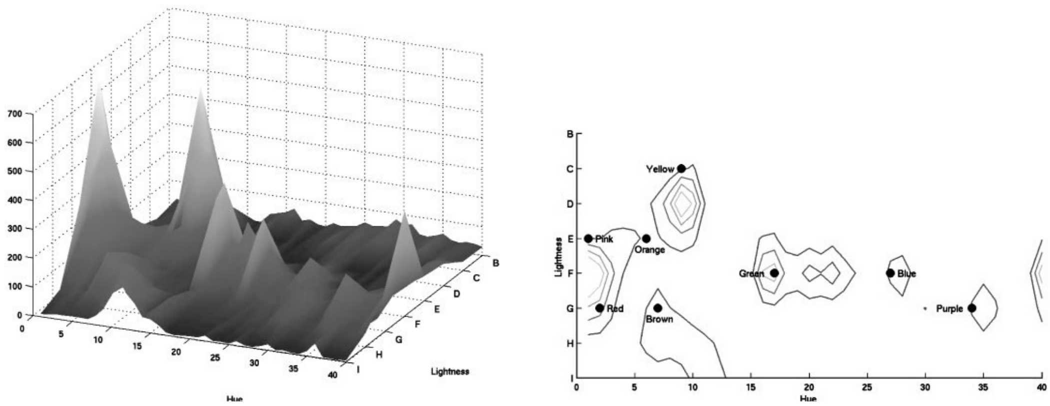


Figure 2-10. The color term centroids of the tested 110 languages in CIE L*a*b* coordinates. A 3D visualization (left). A top view relief map, with 100 occurrences per line (right). The dots mark the English color terms. From (P. Kay & Regier, 2003).

The 3D view of the relief shows how the centroids cluster at certain peaks, which are often close to the original English color terms (compare Figure 2-9). However, the valleys between the peaks are not completely flat, as it may have been induced by the first study. The color hierarchy of the originating basic color test, as shown in Figure 2-8, had to be revised after the WCS, and turned out to be much more complex (Berlin & Kay, 1969; P. Kay et al., 2009). It can be concluded that the WCS gave evidence for a predominant universality of color term clustering, but its clusters are more indistinct and diverse than expected in the first study.

Another open issue, are the influences of language on color categories. It is not clear whether language is fully responsible for categorization, as argued by (Davidoff, 2001), whether it distorts the mapping, or whether it has no influence. The major chicken-and-egg debate on the *Whorfian Hypothesis*, whether language creates categories, or whether categories are formed, and language is only used to express and communicate them, is still open (D. Gentner & Goldin-Meadow, 2003). Maybe the influence of language is depending on the type of category. It might be less influential on *natural categorys*, which are based on natural occurrences in the world, than on *artifact categorys*, which are designed by humans (Medin & Heit, 1999; Medin, Lynch, & Solomon, 2000). Categories describing natural occurrences, such as color or botany, are typically fuzzier than artifact categories, such as employee, bachelor, or researcher. Artifact categories are well defined through, usually exclusive defining features, which are common to all members. It seems that the *feature-based theory* refers to artifact categories, while *prototype theory* gives better explanation for natural categories (Sternberg, 2008). However even some artifact categories seem to have prototypes. For example, some persons consider 7 and 13 to be better examples for odd numbers than others, similar findings have been made for squares or for the category mother (Armstrong, Gleitman, & Gleitman, 1983; Fehr & Russell, 1984; Lakoff, 1987). Hence, the right way for combining both, the feature based theory and the prototype theory, is being investigated following the goal for a full theory of categorization (Hampton, 1997a; E. E. Smith, Osherson, Rips, & Keane, 1988; E. E. Smith, Shoben, & Rips, 1974; Wisniewski, 1997). If language plays a major role for the development and communication of artifact categories, fuzziness of a category might be an indicator for the influences of language.

In the field of cognitive perception, prototypes are considered to be mental representations, which serve as references for pattern perception and recognition. It is assumed that we perceive and recognize patterns, such as certain objects, shapes, or sounds, by comparing them to our previous knowledge.

In perceptional cognitive psychology it is controversially discussed whether recognition or perception is conducted top down, based on high-level cognitive processes, or whether such comparisons could also already be encoded in the stimulus, bottom up. This discussion lies beyond the scope of this work, however, it is more important to get an idea of the mental structures we use for comparison. *Template theories* suggest that we recognize patterns by looking for a perfect match among an immense set of mental models, memorized in high detail. Computer implementations of this concept, for information retrieval, tend to be highly inefficient and unreliable; unlikely characteristics for such an important process in human cognition (Sternberg, 2008).

Experiments conducted by (Chambers & Reisberg, 1985, 1992; Peterson, Kihlstrom, Rose, & Glisky, 1992) clearly suggest that our mind does not construct exact representations. In one experiment (Chambers & Reisberg, 1985) presented certain drawings to participants for a time span of 5 seconds, similar to the one shown in Figure 2-11.

Figure 2-11. Ambiguous image for testing cognitive representations. From (Chambers & Reisberg, 1985).

After exposure the visualized object had to be named. Although the image was ambiguous, most subjects recognized only one possible view. Moreover, they were not able to recall the alternative interpretation from memory. Finally, the test persons were asked to draw the image as they remembered it. After looking at their drawing all of them were able to determine the second interpretation. The results empirically show that we do not seem to store exact representation from real world stimuli, as postulated by template theories.

Unlike templates, according to *prototypes-matching theory*, we compare perceived patterns to prototypes. Prototypes are considered to form around some kind of averages of a class of objects, integrating the most typical or most frequently observed features of a class. Multiple studies support this prototypes concept (Franks & Bransford, 1971; Neumann, 1977; Posner, Goldsmith, & Welton Jr, 1967; Posner & Keele, 1967; Reed, 1972; Solso & McCarthy, 1981). A very illustrative experiment, supporting this theory, is given by (Solso & McCarthy, 1981). They created a set of different basic prototypical identikit pictures, and used those to generate 10 further gradually modified variances. The variances ranged from 75% to 25% similarity, compared to the original prototype face, as shown in Figure 2-12 (left). The modified faces were presented to 36 test subjects for memorization. The source prototype faces were not shown. After the memorization phase, the test persons were confronted with a large set of random faces, including some of the modified faces, as well as the prototypes.

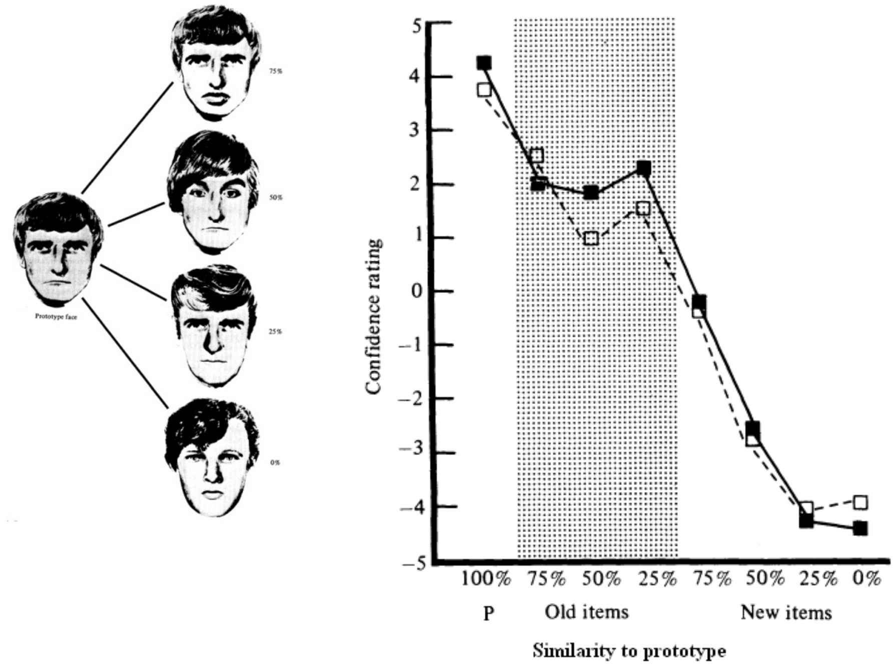


Figure 2-12. Prototype faces used in the experiments of (Solso & McCarthy, 1981) (left). Confidence rating of the faces directly after (filled markers) and after six weeks (unfilled markers) (right). From (Solso & McCarthy, 1981).

Interestingly, although identification of known and unknown faces was very good, 35 of 36 subjects also identified the prototypes, although they were never exposed to those faces. The confidence ratings for the prototypes were even exceeding those of the actually shown items, as depict in Figure 2-12 (right). In a second experiment, 24 of 25 subjects showed the same behavior also six weeks after the memorization task.

This experiment gives hint that prototypes are formed on common features of a class of objects. Multiple studies indicate that the frequency of features is relevant for prototype generation (Neumann, 1977; Posner & Keele, 1967; Reed, 1972; E. Rosch & Mervis, 1975). Although *feature theories* mainly focus on feature perception and recognition, it is still not clear how features are determined cognitively, and how exactly they compose prototypical models for comparison (Neumann, 1977; Sternberg, 2008).

SIMILARITY MEASURES FOR PROTOTYPE CATEGORIZATION

According to prototype theory, an object is part of a category, if it is sufficiently similar to a representative prototype. Different theories on how similarity is measured have been developed, the exact cognitive processes are still not clear (E. E. Smith & Medin, 1981). Because similarity measures and cognitive prototype formation likely underlie probabilistic mechanisms, the prototype theory is also often referred to as *probabilistic view.* A common simplified approximation of the probabilistic processes for estimating similarity is given by comparing the number of features shared between an object and a prototype. Additionally, features may also be weighted by importance (Sternberg, 2008). Hampton expressed this concept mathematically through the following formula.

$(0 \leq w_i \geq 1)$

$(-1 \leq v_{(i,j)} \geq 1)$

$$S_j = \sum_i (w_i \cdot v_{(i,j)})$$

Formula 1. Similarity measure by Hampton (Hampton, 1995).

According to Hampton's linear similarity measure, similarity between two items, i and j, is determined by the sum of the product of the similarity (v), and the weight (w) of all features. The weight corresponds to the salience of a feature, and v is the degree of which the instance j possesses the feature of i. A similar function for determining similarity was also modeled by (E. Rosch & Mervis, 1975; Tversky, 1977). Hampton showed that this function achieves reasonable results in discriminating members from non-members in various categories, such as birds or sports, if it used in conjunction with a threshold (Hampton, 1979, 1995). Although the measure proved certain accuracy, it still relies on many factors of uncertainty. It remains unclear how exactly salience and similarity of single features is determined. In a broader context, it is also not clear whether we use single prototypes as references for categorization, or if cognitive categories are based on multiple exemplars. Exemplars might be typical representations of an object. A category might also include exceptional exemplars, which do not follow rules seen in typical exemplars (Murphy, 1993; Ross & Spalding, 1994; Ross, 2000). However exceptional exemplars seem to be unefficient, due to their sheer amount of occurrences (M. Collier, 2005; J. D. Smith, 2005)

These challenges are also faced in more complex categorization scenarios, based on semantic categories. The classical view here is that objects of a category all share certain common features. However, as (Wittgenstein, 1953) argues, some categories do not seem to follow this premise; instead of common characteristics such categories are rather defined by a network of overlapping features. Exemplarily, Wittgenstein argues that one can think of various games, and never find a feature that is common to all of them. For instance, while some games are competitive, others, like playing

Frisbee, are completely uncompetitive[14]. He calls such categories *family resemblances*. In this case, a category can exist without an existing member including all optional features, defining features are rather spread over several members. Two members of the same category might even have no features in common at all. (E. Rosch & Mervis, 1975) did further investigations on the internal structure of family resemblance, and modeled the family resemblance Score for determining category membership, analogous to Hampton's similarity measure, as given earlier in Formula 1.

Multiple evident concepts on the generation of categories through prototypes exist, however none of them provides a complete accurate model. Hence, a combined model describing natural and artifact category formation seems to be the most promising (Hampton, 1997a; E. E. Smith et al., 1988; Wisniewski, 1997).

CONCLUSIONS

The basic unit for representing symbolic knowledge is called a concept. Concepts are characterized through relations to other concepts. We tend to categorize knowledge based on common characteristics, or features. Category affiliation is either based on explicit defining features, or through similarity to a cognitive prototype. According to popular measures the sum of all shared features helps determining similarity of two items. In this sum each feature is weighted differently by salience and based on a similarity factor specific for both compared items. While defining features are characteristics common to all members of a category, other categories may exist around cognitive prototypes, which only hold optional features, so called family resemblances.

Cognitive prototypes are ideal types also serving as reference concepts for new stimuli. Prototypes are not exact representations of real word occurrences in our brain, as proposed by the template theory, they rather form around some sort of accumulated averages of classes of objects, integrating the most typical features. Their relation to performance will be detailed in the following.

2.1.5 Performance, proceduralization, and category prototypes

Multiple studies give empirical evidence on increased performance at cognitive prototypes. Especially experiments on focal colors show improved efficiency in those areas. Focal colors are better memorized, and earlier learned, than non-focal colors (Heider, 1971a, 1972; E. H. Rosch, 1973b; E. Rosch, 1975a). For example, in an experiment

[14] Oftentimes it is differentiated between games, which are competitive, and play for uncompetitive joyful activities (Deterding, Dixon, Khaled, & Nacke, 2011). Additionally, a common feature of all games might be their relation to fun.

conducted by (Heider, 1971a), children had to pick a proper color chip from a given set, matching a formerly shown target color. Hereby they achieved significantly better on focal colors than on non-focal colors. In another experiment (Heider, 1971b) found that adults remember focal colors more accurately than non-focal colors (cf. Heider, 1972).

(E. Rosch, 1975b) investigated the effects of *priming* on the performance of recognition tasks. Priming describes influences on the recognition of certain stimuli through prior recognitions of the same, or similar stimuli (Neely, 2003). In one of the experiments subjects were presented with either two words or two pictures, such as the vehicles car and airplane. Subjects should immediately determine whether prompted items were identical, or belonged to the same category, by pressing a 'same' key; they should hit the 'different' key otherwise. The degree of category affiliation was determined through an assessment task, prior to the experiment, were the test items have been rated for their goodness of example on a 7 point scale. Additionally, for half of the items, the experimenter induced priming by speaking out loud the category of the succeeding stimuli. For example, the instructor first primes the upcoming stimuli by providing the category name "Fruit". Two seconds after the priming, the words apple and banana were presented, both good examples of the same category, contrarily nut and lemon would have been bad examples of this category. Each pair was shown twice, one time with category priming and another time without raising any expectation. In case of no priming the experimenter simply said blank. The results of the experiment are illustrated in Figure 2-13.

The experiments showed that influences on task performance, through priming for physically identical stimuli, were dependent on the goodness of example. Priming affected performance positively for good examples, had almost no effect for items of medium typicality, and was negative for bad examples. Not surprisingly, reaction times were longer for determining category membership than for marking identical matches. Also word processing was less efficient than pictures. Moreover, matching tasks are processed at a more concrete level when it is asked for identity, and on a more abstract level, when it is asked for category membership, as other experiments show. For colors, however, both performance times are very similar, giving hint that color category names are associated with more concrete object stimuli. (E. Rosch, 1975c). It could also be argued that plain color perception happens on a fundamental perception level, whereas object matching demands more complex cognitive processes.

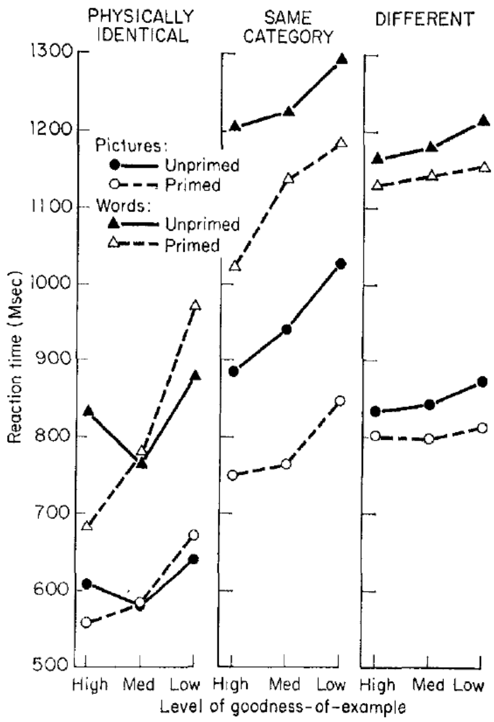


Figure 2-13. Varying reaction times for assessment category affiliation after priming with different levels of category typicality from (E. Rosch, 1975b).

The tests described above may suggest a correlation between abstraction and performance. The more concrete the cognitive level the more efficient our performance. However, cognitive concepts appear to have a *basic level*, which is preferred over more abstract and more concrete representations. If we see a red roundish edible object with a little stem on it, we are most likely thinking of an apple, although one could also refer to it as a fruit or Red Delicious (Medin, Proffitt, & Schwartz, 2000; E. Rosch, 1978; Sternberg, 2008). (R. Brown, 1958, 1965; P. Kay, 1971) suggest that the basic level corresponds to areas were most obvious differences between single concepts are perceived. Consequently, basic levels "*partition the domain of individuals in the way that correspond to the most obvious discontinuities in nature*"(P. Kay, 1971, p. 878). Similarities in basic level and prototype tests give hint that basic level and prototype categories correlate in many things. On the one hand, prototype categories are most developed on basic levels. Moreover, the structure of basic levels is very similar to the structure of prototype categories (Ungerer & Schmid, 2006). It is claimed that prototypes maximize efficiency of basic level categories by maximizing distinctiveness, since they comprise the largest numbers of attributes shared in a category, and, at the same time, the largest number of attributes not shared with other categories (E. Rosch & others, 1977; E. Rosch, 1978; Ungerer & Schmid, 2006). Basic levels seem to be changing depending on context or expertise (Tanaka & Taylor, 1991). (E. Rosch, Mervis, Gray, Johnson, & Boyes-Braem, 1976) found that people identify objects at a basic level more quickly than they identify objects at higher or lower levels. Similarly children learn typical instances of categories earlier than they learn atypical ones (E. Rosch, 1978).

The mechanisms of category prototyping may also be active in domains such as learning and automatization. Prototyping processes structure recognized patterns, and thus generate cognitive reference points (P. Kay et al., 2009; Neumann, 1977; E. Rosch & Mervis, 1975; E. Rosch, 1975a; Solso & McCarthy, 1981; Sternberg, 2008). Such reference points are probabilistic cognitive representations of occurring input stimuli, mainly based on attribute frequencies (Hampton, 1979, 1995; E. Rosch & Mervis, 1975; E. E. Smith & Medin, 1981; Tversky, 1977). The studies on focal colors in combination with perceived color frequencies, as elaborated later in this work, underline this relationship of occurrence clusters and prototypical colors (P. Kay et al., 2009; P. Kay & Regier, 2003; Richard Cook et al., 2012). As detailed above, prototypes also correlate with areas of improved performance (Heider, 1971a, 1971b, 1972; E. H. Rosch, 1973b; E. Rosch, 1975b, 1978). All those aspects also account for automatization and learning processes.

A popular evidential activity for understanding automatization and pattern learning is playing chess. (Chase & Simon, 1973) tested the ability of remembering chess figure positions with beginners and masters. They tested completely random constellations, and "real" constellations from chess games. Results are shown in Figure 2-14.

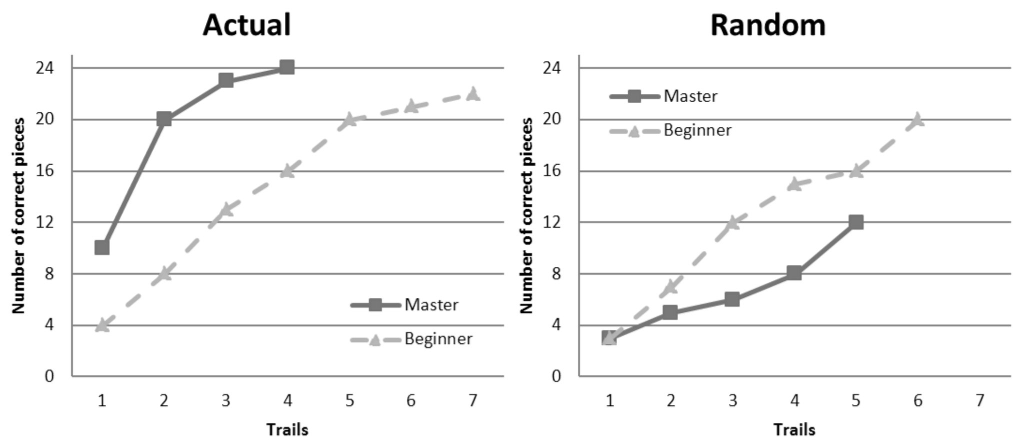


Figure 2-14. Performances on remembering actual and random chess positions distinguished between experts and novice performances (Chase & Simon, 1973).

As expected, experts performed much better than novices, when they had to remember actual chess positions. However, this difference was not observed for random distributions (also compare De Groot, 1978; Vicente & De Groot, 1990).

These results are in line with other tests conducted on performances of cognitive prototypes. Since prototypes refinement may be based on mechanisms, very similar to those also active in automatization processes, it is worth taking a closer look to cognitive processes involved in this field.

<u>FROM CONTROLLED TO AUTOMATED PROCESSES TO HABITUATION</u>

From the perspective of cognitive psychology, cognitive processes vary between highly effective *automatic processes*, and comparatively in-effective *controlled processes* (Sternberg, 2008). *Controlled processes* are new intentional procedures, typically performed step by step. Practice may lead to *automatization* of such processes (LaBerge, 1975, 1976). According to (Posner & Snyder, 2004), automatic processes are performed unconsciously, causing minimal attention interference. Transition from controlled to automatic actions seems to be continuous (J. R. Anderson, 1983; Bryan & Harter, 1899; LaBerge & Samuels, 1974; G.D. Logan, 1988). The studies of (LaBerge & Samuels, 1974; Samuels, LaBerge, & Bremer, 1978), for example, provide insights on how *proceduralization* behaves during the process of learning to read.

(LaBerge & Samuels, 1974) investigated 16 college students who had to read unfamiliar Greek, and familiar Romanic letters. While processing unfamiliar letters became more efficient over time, the reading speed of familiar letters stayed the same, as shown in the following diagram.

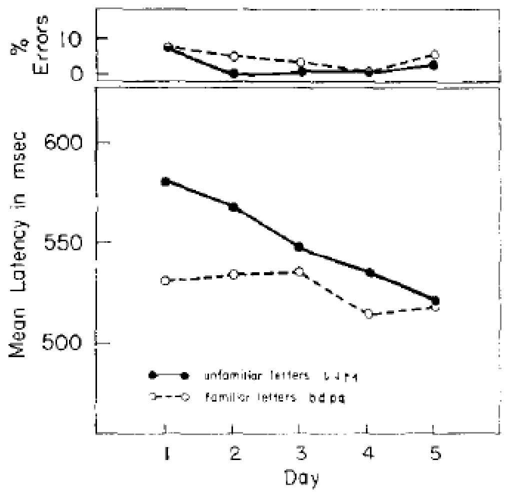


Figure 2-15. Reading speed and error rate change over time during training. From (LaBerge & Samuels, 1974).

Although the cognitive proceduralization of patterns is not fully understood in detail, there is substantial evidence that novel processes are consciously controlled, and may be gradually transformed into automated high performance processes. The nature of performance gains through automatization has been mathematically expressed in the popular, but controversially discussed formula of the *power law of practice,*

$$RT = a + b \cdot N^{-c}$$

Formula 2. Mathematical description of the power law of practice, from (Gordon D. Logan, 2002).

Hereby, a describes the asymptotic RT (reaction time), b is the maximum amount by which RT can be reduced through practice, N is the number of practice trials, and c is the learning rate. Consequently, a high learning rate results in quicker acquaintance, and also earlier stagnation (Gordon D. Logan, 2002). (Sternberg, 2008) visualized a curve similar to the *power law of practice.*

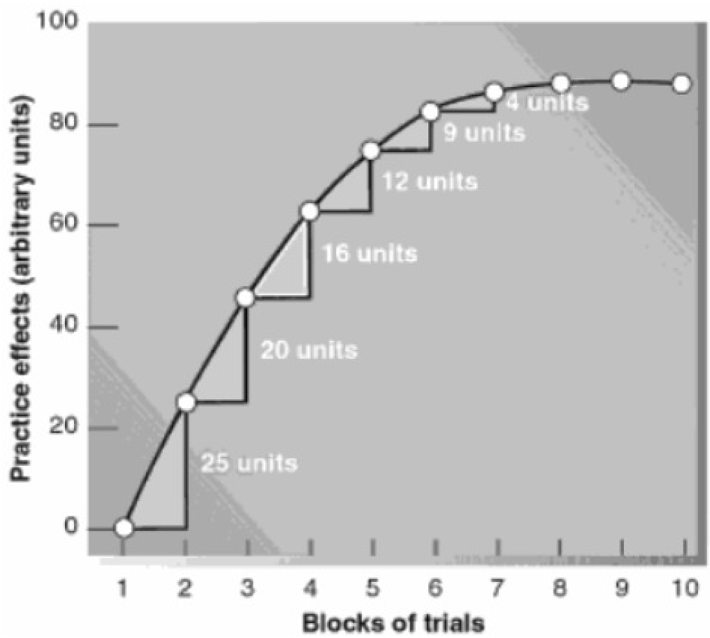


Figure 2-16. Visualization of the power law of practice (Sternberg, 2008).

The improvements of practice become smaller with every trial until they are not measurable anymore. A zero slope of performance improvements is a classic indicator for automatized tasks, compare (Palmeri, 1999; Schneider & Shiffrin, 1977). The highest level of automatization is also referred to as *habituation.* Habituation further needs to be distinguished from *sensory adaptation.* While our sensory system might adapt to certain conditions, smells, brightness, temperature, etcetera, habituation is being learned. Contrarily, adaptation is unrelated to frequency of prior exposure.

Adaptions also describe reactive processes beyond conscious control, which are induced by external stimuli. In contrast, habits can be controlled intentionally.

(Samuels et al., 1978) tested the reading speed and errorness of second, fourths, and sixth graders, as well as of college students. They tested 20 random subjects of each group for their ability to read words of three to six letters length. Results are visualized in Figure 2-17.

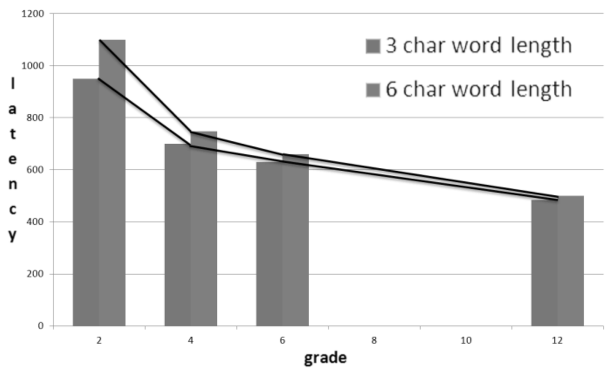


Figure 2-17. Reading speed latency of 3 graders to college students for words of 3 to 6 characters length from (Samuels et al., 1978).

Not surprisingly, reading speed increases with experience. The curve in Figure 2-17 shows a non-linear alleviative degression of latency with greater experience. For second to fourth graders, reading speed increased in average by approximately 42%. This gain drops to almost 9% for children between fourth and sixth grade, and is further saturated with increasing experience. The results also show a significant decrease of reading speed with increasing word length for second graders, while fourth and sixth graders are less influenced by word length, and college students are almost not affected at all. According to these results, second graders seem to read letter by letter, as indicated by the gradual performance increase for longer words. The flat curve of the college students suggests a holistic recognition of words. The unsteady performance curves of fourth to sixth graders suggest a component based reading style.

In summary, speed increases with practice, correlating to the processed size of reading patterns, or chunks. The results of the college students holistic recognition speed shows that *chunking* also has a positive effect on performance.

Chunks can be basically described as cognitive representations of patterns. Chunks are nested: each chunk usually consists of smaller chunks. Hence, chunks can be recognized at different levels of complexity (G. D. Bower, 2008). The term "open-minded" may be recognized on a letter basis – as it is suggested for second graders reading skills. It may also be recognized component based "open-mind-ed", which resembles recognition processes closer to those of fourth and sixth graders. More experience readers may as well recognize it as a whole (compare Samuels et al., 1978).

The term of chunking evolved from work driven by (De Groot, 1946, 1978; Miller, 1956). (Miller, 1956) found that our abilities to memorize items in short term is dramatically limited, compared to our long term knowledge. In his early experiments, participants were able to memorize around seven items of information. This number stands in contrast to our abilities for recalling huge amounts of information from long term memory. Consequently, new incoming information should be structured in chunks of information, and recognized according to existing knowledge in the Long Term Memory; otherwise the short term memory would be an obstructive bottleneck. Bower and Springston (1970) conducted the following experiment to study the connection between the long term memory and chunking. A reader had to read out loud a certain sequence of letters. A group of test persons was asked to remember those letters. The letters were read in different ways. In one way the reader said "FB … IPH … DTW … AIB … M". In the second way the reader read "FBI … PHD … TWA … IBM". The letters of the second variant were much more easily remembered, since they referred to commonly known abbreviations.

Recognizing patterns is a key performance driver, and improves efficiency. Our performance highly depends on learned patterns. Bower further found in his experiments that students learning abilities of the same input were highly minimized, when such inputs were presented in randomly changing chunks (G. D. Bower, 2008). "*An implication of this constant-chunking result is that people will readily recognize and reproduce any symbol sequence that conforms to chunks they already know, that are "familiar"*" (G. D. Bower, 2008, p. 15).

Those findings correlate with insights from other fields, such as the *elaboration of knowledge,* which refer to differences between experts and novices in the area of problem solving. Multiple experiments, from various fields, support an observation, which is well illustrated by the chess position experiment of (Chase & Simon, 1973), as described earlier (Gobet & Simon, 1996a, 1996b, 1996c; Larkin, McDermott, Simon, & Simon, 1980; Lesgold, 1988; Reitman, 1976). The chess experiment shows plastically how the different knowledge structures of novices and experts affect performance. Expert knowledge seems to build on large information units, which are well interconnected based on structural similarities. In contrast, novice knowledge seems to be based on small chunks of information, loosely connected through superficial similarities (Bryson, Bereiter, Scardamalia, & Joram, 1991; Chi, Glaser, & Rees, 1982; Larkin et al., 1980). The large base of evidence for chunking mechanisms, involved in human perception and goal oriented problem solving, also led to computational models and implementations, such as the *Elementary Perceiver and Memorizer* (EPAM), or the *Chunk Hierarchy and Retrieval Structures* (CHREST) (Gobet et al., 2001).

Chunking is a key mechanism for mastering actions and knowledge. Experts greatly reduce the amount of new learning by recognizing and building on previously known chunks (G. D. Bower, 2008). "*[T]he ability to chunk information into meaningful units allows for superior memory and capacity*" (Sternberg, 2008, p. 462). The main challenge of chunking remains in recognizing chunks. Where and why do we refine cognitive chunks? Bower determined two major principles in this process. The recognition of chunks depends on how we group information. We often tend to group information by proximity in time and space, or by similarity. For example, objects which follow in close sequence, look, or sound alike, are grouped together, in the same way we quantize the color spectrum in major colors.

The recognition and grouping of information chunks is a mechanism prevalent in many cognitive processes. Researchers try to find out where cognitive prototypes are formed, why basic levels are hierarchically neither abstract nor concrete (as described earlier), or why we perceive objects as a whole, and not as a sum of its parts. The field of *Gestalt* psychology is concerned with the question how elements or groups of objects form an integrative whole - how the whole differs from the sum of its parts (S. E. Palmer, 1999, 2000; S. Palmer, 1999). According to the Gestalt principle of *Prägnanz*, we organize objects into a stable form. One might simply think of a tree. Instead of perceiving myriad of leafs and branches we recognize such patterns as a whole. Moreover, if the tree stands in a forest, we could either still focus on the tree, or the whole forest. In this case, following this idea of *figure and ground*, the tree would be our focal figure, and the rest of the forest would be the ground. We could as well focus on the forest as a whole. Other principles that have been determined as part of the law of Prägnanz are *proximity*, *similarity*, *continuity*, *closure*, and *symmetry* (Sternberg, 2008; Ungerer & Schmid, 2006). Prototypes also play a role in the law of Prägnanz, if we consider that focal objects are likely to be prototypes (Ungerer & Schmid, 2006).

As stated above experts knowledge heavily builds on well interconnected information chunks. There is empirical evidence that increased interconnectivity, so called *semantic elaboration*, correlates with improved efficiency. In this context, Craik and Lockhart found that subjects could memorize words better, when they were part of a more elaborate task (F. I. M. Craik & Lockhart, 1972). Craik and Lockhart determined different levels of information processing and argued that more elaborate information increases efficiency in recalling this information. (Hyde & Jenkins, 1973) re-firmed the results of (F. I. M. Craik & Lockhart, 1972). They found that the memorization performance of semantic word related tasks, such as determining the type of word, was better than the performance of non-semantic word tasks, such as looking for the appearance of a certain letter. The results also showed that the different levels of processing proposed by (F. I. M. Craik & Lockhart, 1972) were very likely to be insufficient (C. D. Morris, Bransford, & Franks, 1977).

In later studies test persons were given a cloze test (F. I. M. Craik & Tulving, 1975). Probands were asked to complete a sentence with a given noun. The sentences had different levels of semantic and contextual complexity. One sentence could be a simple statement, while the other describes a whole action chain. After completing multiple sentences, test persons were asked for previous fill words. It turned out that memorization was better when the semantic context was more complex. It is assumed that the improved memorization performance is based on the higher amount of activated associations. This concept is called semantic elaboration, and it seems to be based on knowledge from the long term memory (Oberauer, Mayr, & Kluwe, 2005)

CONCLUSIONS

There is strong empirical evidence that cognitive prototypes correlate with areas of superior cognitive efficiency. Hereby, there is tendency for concrete or abstract concepts being more efficient; the most efficient areas usually lie on a level in-between. Those, so called basic levels, depend on typicality, and they correlate with cognitive prototypes. They are also differing based on context and expertise. It is claimed that prototypes maximize efficiency of basic level categories by maximizing distinctiveness. They hold the largest numbers of attributes shared in a category, and, at the same time, comprise the largest number of attributes not shared with other categories.

Additionally, priming has a positive effect on performance especially for ideal types, while it is negative for bad examples. Cognitive prototypes correlate with good examples. Hence, if cognitive prototypes are utilized to increase efficiency of user interfaces, it is more important to meet expectations. If elements of low typicality are used, it is not so important to meet expectations. Consequently, completely new functions should probably better be implemented through bad examples, while typical interface elements only release all their efficiency potential if they behave exactly as expected.

Cognitive prototypes are probabilistically formed based on stimuli frequencies, a mechanism which also is prevalent in cognitive proceduralization. Performance advantages of automatization are achieved through practice. Repetitive exposure to actions, or information, reinforces representative cognitive patterns, so called information chunks. Chunks are nested in various levels of complexity. Compared to novices, experts gain most of their performance advantages from an elaborate knowledge structure, consisting of more complex highly interconnected chunks. This structure allows them to greatly reduce the effort of new learning by building on previously known chunks. Hence, reusing and interconnecting existing chunks is fundamental for increased efficiency in several cognitive processes. Reusing and interconnecting cognitive patterns are the result of more basic similarity matching processes.

2.1.6 Pervasiveness of Similarity Comparisons

Similarity comparisons in perception and recognition are fundamental for most of our remarkable cognitive abilities, including abstraction, categorization, inference, comparison (D. Gentner & Christie, 2008; D. Gentner, 2003; Penn, Holyoak, & Povinelli, 2008). Similarity comparisons are inborn, and essential for learning in various ways; virtually every cognitive process, such as categorization or transfer, involves explicit or implicit comparisons (D. Gentner, 2003). Hence, Gentner developed a functional model based on human cognitive processes which describes the representation and mapping of conceptional structures, called *structural-mapping* theory (D. Gentner, 1983). Analogous to the role of similarity comparisons in human cognition, processes, such as perception, abstraction, categorization, or inference are essential for the structure-mapping theory. Similarity comparisons are necessary for the alignment and mapping between structured conceptual representations. The existence of internal structural representations of objects and their properties is assumed. According to the model, metaphors are considered as comparisons that share primarily relational information. This way one can consider them as analogies (D. Gentner et al., 2001). On the one hand, the model has been developed based on empirical evidences in cognitive psychology, on the other hand it has been shown that structure-mapping theory proofed accurate modeling of cognitive principles of operation of metaphors and analogies (Gentener, Bowdle, & Ortony, 2008; D. Gentner & Markman, 1997; D. Gentner, 1983, 1988).

Comparisons are based on similarity, but similarity comparisons are asymmetrical. Each comparison usually underlies an asymmetric immanent order: the more salient object serves as comparison base. We say “your new sports car looks like a Ferrari” rather than “a Ferrari looks like your new sports car” or “an elipse is like a circle” rather than “a circle is like an elipse” (Tversky, 1977). Hereby, salience seems to be related to prototypability. The object more similar to a prototype is predominantly considered as more salient. As studies show, the asymmetry of metaphors is even more fundamental. Reversing metaphors affects their interpretability since salient base features are preferably mapped first (D. Gentner & Clement, 1988; Glucksberg

& Keysar, 1990; Glucksberg et al., 1997; Ortony, Vondruska, Foss, & Jones, 1985; Ortony, 1979). Gentner proposes three possibilities of asymmetry, initial projection, initial abstraction, and initial alignment (D. Gentner et al., 2001). The strongest is possibly the temporal asymmetry, as shown in Figure 2-18 .

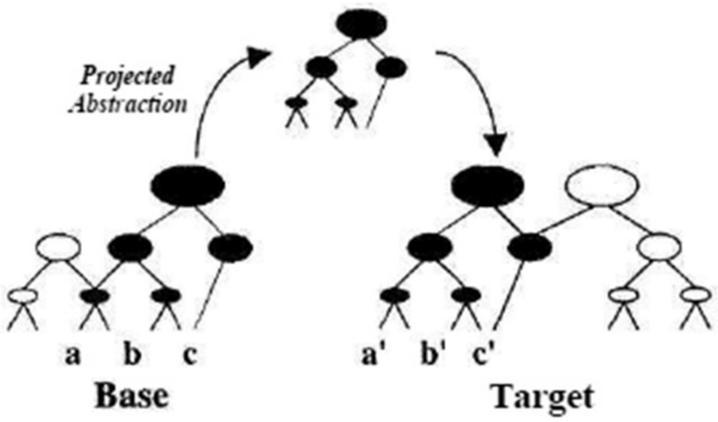


Figure 2-18. Initial Projection: temporal asymmetry of metaphors from (D. Gentner et al., 2001).

As depict, the initial projection comprises three steps. First salient features of the base are accessed. In a next step these features are abstracted, and finally projected from base to the target.

Theories similar to the structural-mapping theory are frame-based theories or the *conceptual blending* theory. Beyond similarity comparison they offer theoretical ground for other cognitive processes, such as creativity, abstraction, and many more. Such theories give explanation on how knowledge may be shaped through conceptual blends (more in (Coulson, 2001; Fauconnier & Turner, 2003)

CONCLUSIONS

The structure mapping theory provides an empirical evidence based model of human cognition, and underlines the pervasiveness of similarity comparisons. In combination with other cognitive models, such as prototyping, automatization, and chunking, it allows for reviewing and getting an understanding of efficiency advantages achieved through high level mechanisms, like metaphors or categories. The theory highlights the importance of similarity comparisons in perception and recognition. This inborn mechanism is fundamental for most cognitive processes, such as abstraction, categorization, or inference. Similarity comparisons are the basis for mapping cognitive concepts, which highly increases cognitive efficiency. Comparisons are based on similarities, typically of salient features, and seem to be asymmetric from base to target concepts.

2.1.7 Cognitive Efficiency Catalysts in Communication

As empirically demonstrated throughout various prototype experiments we prototypically categorize as we recognize objects. We do not store exact representations, but more efficiently, build on existing knowledge, and gradually refine its structure according to new stimuli (Chambers & Reisberg, 1985; Chase & Simon, 1973; P. Kay et al., 2009; Neumann, 1977; Posner & Keele, 1967; Reed, 1972; Solso & McCarthy, 1981; Vicente & De Groot, 1990 and many more). Several studies underline the improved efficiency at prototype categories (Heider, 1971a, 1971b, 1972; E. H. Rosch, 1973b; E. Rosch & Mervis, 1975; E. Rosch, 1975a, 1975c, 1978).

As it is more efficient to memorize and recognize based on prototype categories, we also improve efficiency in communication by using categories. For example, we could describe a new car as two-box design with hatchback, taller than a sedan, for five passengers, with five doors, two of them are sliding doors, however, more efficiently we rather use the category term minivan to describe such a car (Wikipedia.org, 2012e). Obviously terms like sedan or hatchback describe categories themselves.

Prototypes are available at various levels of complexity, according to our hierarchical understanding of categories from concrete to abstract. As depict earlier, we prefer cognitive operations on the so called basic levels, which usually lie on a level between the most concrete and most abstract known cognitive concept. Basic levels also comprise areas of maximal operational efficiency since they correspond to cognitive prototypes. They are thought to maximize distinctiveness, since their concepts hold the largest numbers of attributes shared in a category, and, at the same time, own the largest number of attributes not shared with other categories (E. Rosch & others, 1977; E. Rosch, 1978; Ungerer & Schmid, 2006). As forming of prototype categories is dependent on expertise, the basic level seems to be different depending on context and expertise, as well (Tanaka & Taylor, 1991).

Similarly, something like a shared basic level also exists in communication between multiple partners (compare previous section on Similarity Measures for Prototype Categorization). Finding the right level of complexity can be challenging, as it is exemplified and visualized in Figure 2-19

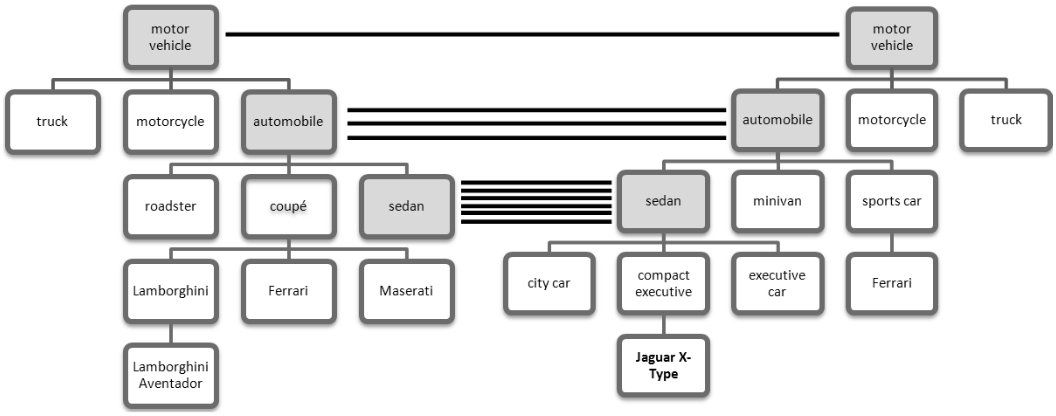


Figure 2-19. Two, partly different conceptual models, of categories shown in hierarchical organization. Sub-ordinates inherit salient features from super ordinates. Black connecting lines illustrate possible informative richness at different matching levels in the concept hierarchy.

The figure illustrates the differently developed cognitive models of two communicating persons. While they have some concepts in common, others are structured in a different way, or even completely unknown. For example, the person with the left cognitive model does not know the concept of a "Jaguar X-Type". Most efficiently this concept could be communicated by referencing the overlapping concept "sedan", since it is the next superordinate known by both sides. By being the next superordinate it has most salient features in common with the target concept to be communicated. Consequently, after finding this level, in theory "sedan" should be the most efficient level, for both persons to start talking about the more concrete object.

Practically, cognitive prototype categories, especially natural categories, are vaguely defined and may be changing dynamically over time. Encyclopedic categorization definitions are precise expert descriptions, usually exceeding the general understanding of categories. The World Color Survey and the distribution of color in our environment illustrates that perception and recognition of objects is not uniformly distributed. Natural prototype categories are formed around salient features of a class of objects, based on occurrences in our perceived environment, also compare (P. Kay & Regier, 2003; Paul Kay & Cook, 2011; Neumann, 1977; Posner & Keele, 1967; Reed, 1972; E. Rosch & Mervis, 1975; Sternberg, 2008). Prototype categories are cognitive refinements at probabilistic feature maxima allowing for high performance information processing (detailed in the previous section Color Perception and the World Color Survey).

Concluding the above, if we talk about categories we usually have some sort of prototypical probabilistic average in mind, especially when it comes to natural categories. This average is on a basic level of high efficiency, which usually is not the most concrete or abstract concept we know. For example, what comes to our mind, if we think of a sports car? Many might immediately think of a Ferrari, and if we are asked to think of a Ferrari, we might have a red wedge-shaped car in front of our inner eye, although Ferrari cars also comprise various other models.

Considering a person represented by the right cognitive model, shown in Figure 2-19; if this person does not know a Lamborghini, and we say that a Lamborghini Aventador is a sports car, the cognitive Ferrari prototype is intrinsically used as reference for comparison (see Similarity Measures for Prototype Categorization). In this case, saying that a Lamborghini Aventador is a sports car is as efficient as comparing it to a Ferrari.

This aspect of the nature of cognitive prototype categories demonstrates their fuzziness, and their close relationship to metaphors. Some researchers came to the conclusion that metaphors may also be seen as a specialization of categorization (Glucksberg & Keysar, 1990; Glucksberg et al., 1997; Honeck et al., 1987; Kennedy, 1990).

A test conducted by Rosch exemplifies this view (E. H. Rosch, 1973b). In her experiment Rosch found further evidence for the existence and importance of prototype categories in the recognition of shapes. Besides confirmation of her assumptions, she also revealed interesting findings through pre-tests. In such pre-tests she wanted to verify that the Dani did not have category names for shown items. Therefore, one Dani had to explain drawings to another, who had not seen the drawings before. The drawings showed variations of shapes similar to the ones shown in Figure 2-20.

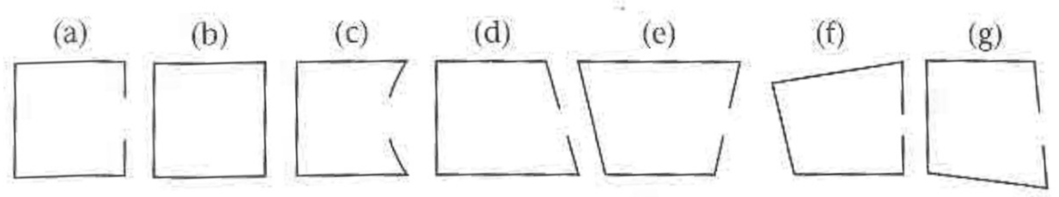


Figure 2-20. Drawing used by Rosch (E. H. Rosch, 1973b) for evaluating prototype categories in the recognition of shapes. From (Ungerer & Schmid, 2006).

How would you describe those shapes to a third person who does not see them? According to prototype theory, we should make use of conventional prototypical shapes as reference. Thus, the first looks like a square with a gap on the right hand side; the second is a square, and so on. The Dani, however, do not have abstract concepts, such as square or trapezes. Hence, they immediately utilized prototypes from their environment by comparing those shapes to “a pic” or a “a broken fence” (E. H. Rosch, 1973b; Ungerer & Schmid, 2006). While we could use superordinate categories, the Dani spontaneous create categories by using salient similar prototypes for explanation. This aspect shows how categorization is a natural mechanism much more unstructured than our conventional language based categorizations.

As elaborated above, categories and metaphors build on the same mechanisms for increasing efficiency. Cognitive prototypes are used as reference points for new information, interpreted through similarity comparison. Gentner and associates even suggests that metaphors can take an evolutionary path from comparison to categorization during conventionalization, while at the same time becoming more abstract (D. Gentner et al., 2001).

Other constructs in linguistics, closely related to metaphors, are *metonymies*. For example, we say “the press wrote” instead of “reporters wrote”. More conventionalized metonymies, such as “head of institute” or “program chair”, are often not even recognized in daily language. While metaphors build on the principle of similarity, metonymies build on the principle of *contigutity* – a salient associative relation between two concepts. For instance, the press is associated with reporters although they share no salient features (Gibbs, 2008; Ungerer & Schmid, 2006). Examples, such as “having a heated argument”, illustrate the fuzziness between metaphors and metonymies. Not surprisingly metonymies are often seen as specializations of metaphors (Ungerer & Schmid, 2006). This might be explained through their associative nature: concepts which occur in close temporal or spatial distance are mentally linked to each other, and tend to have further similarities (also compare Gestalt laws). Consequently, metonymies might unintentionally inherit characteristics of metaphors.

CONCLUSIONS

Conceptual metaphors provide vehicles for efficiently communicating knowledge. By giving a comparative conceptual context, the speaker is able to communicate many attributes through a single comparison. It is also easier to remember new information if the hearer gets them presented in a context of known cognitive concepts. The principle of operation is similar to the one of categorization. Natural prototype categories are vaguely defined in various dimensions, changing dynamically over time, with changing context, and expertise. According to prototype or instance theory, this structural mechanism might be a basic functionality of efficient knowledge representation in our brain.

Although Metonymies are sometimes defined as specialized Metaphors (Ungerer & Schmid, 2006), they do not come with any direct efficiency advantage, because they need to be learned before they transport knowledge. Metonymies describe learned associations which work on the basis of inference. During the process of inference, fundamental similarity comparison processes are active, analogous to those used in metaphors (compare Structure Theory and Conceptual blending theory briefly described previously). The function and advantage of those processes have been investigated on metaphorical basis, and will not be further investigated here.

2.1.8 Efficiency in Semiotics

Mechanisms, similar to linguistic metaphors and metonymies, can also be found in non-verbal semiotic communication. According to Peirce' model of semiotics for object-relations it can be differentiated between *icons*, *indices*, and *symbols*. Icons build on similarity comparisons, indices are based on contiguity, and symbols rely on conventions (Atkin, 2010). In the context of efficiency, it is of interest how well semiotics support communication, or how good they are understood.

The use of icons is based on the same principles as metaphors or categories. Salient features turn them into a powerful communication mechanism, which needs no further learning. By building on similarity comparison their quality depends on the commonness of the mimicked object, its prototypicality, and the selection of most salient features. Prototypical representations attract attention quicker than other stimuli, and are more easily remembered than less salient stimuli (Heider, 1971a, 1972)

As research on basic levels suggests, such objects should neither be too abstract nor too concrete (compare (E. Rosch et al., 1976) and previous section on *Concepts, Prototypes, and Categories*).

Indices build on contiguity, which demands at least some kind of inference. Analogous to metonymies, indices are based on a common understanding of associations between used concepts, among all communicating parties. The sound of paper being scrunched up, for example, is often used as an audible computer feedback of something being deleted, or metaphorically, being thrown away into a bin.

Another example for indices are stereotypical judgment (status symbols) based on consumption objects. (Belk, Bahn, & Mayer, 1982) investigated consumption symbolism among ages. It seems that people see the possessions as extensions of themselves (Prelinger, 1959; Secord, 1968). Hence, consumption goods allow them to express themselves, in the same way they also decode consumption cues from others. (Belk et al., 1982) tested 956 subjects, consisting of pre-school pupils to eight's graders, college students, and adults older than 28 years. They found that consumption stereotyping intensified with age. Pre-school students showed least ability to read symbols, while college students showed greatest degree of interpreting materialistic status symbols (Belk et al., 1982). These results support concepts of developmental psychology where children begin to infer about others, based on their actions from the age of four or five. The results are also in line with Piagets' theory that children improve their ability to judge others over the years (Piaget, 1954, 2007). Investigating the development of language in children, Ginsburg & Opper and Watts (Ginsburg & Opper, 1988; Watts, 1944) also found that children refine their judgments of others.

(Gerstadt, Hong, & Diamond, 1994) conducted a modified Stroop test to find out how children perform in the test. In the Stroop test color words, such as red green or blue, printed in a different color are presented to subjects (Stroop, 1935, 1992). In the modified test, children, between the age of 3 ½ and 7, were randomly prompted with either a white card, showing a sun, or a black card, showing a moon. The subjects were asked to respond with saying "day ", when they see the black moon card, and "night" when they see the white sun. A control group had to respond in the opposite way. They found that the younger children had clear performance deficits over the older children, due to the difficulty of the task (Gerstadt et al., 1994). The control group which performed on iconic level was much more efficient than the children working on conventional level.

Such findings are in line with those from research on structural-mapping theory. Similarity comparisons are inborn and fundamental for various higher level cognitive processes including abstraction and inference. Hence icons work on a more fundamental and advanced level of cognitive processing. Indices often demand more learning and require more complex processes, therefore they tend to be more effortful than icons. On the other hand, findings on chunking, prototyping, and automatization advocate processing advantages for habituated entities (as detailed in section *Concepts, Prototypes, and Categories* and section *From controlled to automated processes to habituation*). Cognitive refinement through repetition may increase efficiency of every semiotic instrument, whether these are icons, indices, or symbols, up to a highly efficient level of recognition.

CONCLUSIONS

According to Peirce' model of semiotics for object-relations, semiotics can be distinguished between icons, indices, and symbols. Icons build on comparison similarity, indices are based on contiguity, and symbols rely on conventions. Empirical findings on prototype categories, automatization, and metaphors in cognitive psychology suggest that recognition of new icons requires least complex cognitive processes and may be most efficient.

Indices build on associative processes. Recognizing new indices may therefore be more effortful than recognizing icons. Research on cognitive basic level gives hint that both instruments should utilize salient concepts at prototype categories, which are likely not to be based on a very concrete or highly abstract level. In general, icons are comparatively least dependent on prior learning. While indices may require more learning effort, symbols usually cannot be recognized without prior learning.

2.1.9 Summary of Human Cognition in Digital Transformatives

In the previous sections related evidence based research on mnemonic devices and associative cognitive tools has been revised. Major aspects will be briefly summarized here. After the epidemiological property extraction of existing implementations the identified aspects will be combined in a concept of Digital Transformatives in chapter 3.

In linguistics, metaphors work on a level of comparison. They are closely related to *similes* or *analogies*, as they are working on items that share primary attributes (D. Gentner et al., 2001). Metaphors can be seen as cognitive instruments, where a *source concept* is mapped onto a *target concept* (Lakoff & Johnson, 1980; Lakoff & Turner, 1989; Strube et al., 1996; Ungerer & Schmid, 2006).

Metaphors may also be seen as a species of categorization (Glucksberg & Keysar, 1990; Glucksberg et al., 1997; Honeck et al., 1987; Kennedy, 1990). *Categories* are structures, organizing concepts based on common *features*, or through similarity to a *prototype* (Coley et al., 1997; Hampton, 1995; Medin, 1998; Sternberg, 2008; Wattenmaker, 1995; Wisniewski & Medin, 1994). It can be distinguished between natural categories and artifact categories. Natural categories are based on natural occurrences in the world. Compared to artifact categories they are fuzzier and formed around cognitive prototypes. Artifact categories are designed by humans and commonly described through defining features. (Medin & Heit, 1999; Medin, Lynch, et al., 2000). While language seems to have minor influences on natural categories, it plays a major role in the development and communication of artifact categories – as demonstrated by the basic color terms and world color survey (Berlin & Kay, 1969; P. Kay et al., 2009). Hence, fuzziness of a category might be a reciprocal indicator for the influences of language.

During perception we seem to compare patterns of natural occurrences to cognitive prototypes. Prototypes are probabilistic clusters, which form around some kind of averages of a class of objects (Franks & Bransford, 1971; Neumann, 1977; Posner et al., 1967; Posner & Keele, 1967; Reed, 1972; Solso & McCarthy, 1981). They integrate most typical observed features for an object class. The frequency of feature occurrence is of fundamental relevance in prototype formation (Neumann, 1977; Posner & Keele, 1967; Reed, 1972; E. Rosch & Mervis, 1975). Natural category membership may be well approximated with measuring similarity to centers of occurrence clusters, but are hardly predicted in boundary areas. For determining appropriate similarity measures it is necessary to evaluate features in a weighted way (Hampton, 1979, 1995; E. Rosch & Mervis, 1975; Tversky, 1977). In combination with human perceived color distribution, the World Color Survey gives a sound basis for investigating natural prototype formation in detail, on a basic stimulus level. The perceived color distribution and determined categories give an understanding of the different weights of features and on family resemblance. This is, several objects may be necessary to fully cover all relevant features of a category (E. Rosch & Mervis, 1975; Wittgenstein, 1953). Colors are most exemplary at prototype categories, and salient forms are the good forms of gestalt psychology (E. H. Rosch, 1973a). Such colors and forms are more efficiently processed, better Remembered, and attract attention predominantly over other stimuli (Heider, 1971a, 1972). "*When category names are learned, they tend to become attached first to the salient stimuli (only later generalizing to other instances), and by this means "natural prototypes*" become foci of organization of categories."(E. H. Rosch, 1973a, p. 330)

There is empirical evidence, from several studies, showing that cognitive prototypes describe areas of improved efficiency and correlate with increased performance (Heider, 1971a, 1971b, 1972; E. H. Rosch, 1973b; E. Rosch & Mervis, 1975; E. Rosch, 1975a, 1975c, 1978). Priming has a positive effect on good examples and a bad effect on bad examples (E. Rosch, 1975c). Studies on basic level theory show that performance does not correlate to hierarchical abstractness, but are highest somewhere in between the most concrete or most abstract cognitive concept (E. Rosch et al., 1976; E. Rosch, 1978). Additionally, prototype categories are most developed on basic levels (Ungerer & Schmid, 2006). It is claimed that prototypes maximize the efficiency of basic level categories, by maximizing distinctiveness, since they comprise the largest numbers of attributes shared in a category, and at the same time the largest number of attributes not shared with other categories (E. Rosch & others, 1977; E. Rosch, 1978; Ungerer & Schmid, 2006). The correlation between basic levels, prototypes and efficiency is also supported by studies showing that basic levels are depending on expertise (Tanaka & Taylor, 1991).

The boundaries from prototype formation and cognitive automatization (or proceduralization) are fluent, as depict by experiments on recognizing and memorizing chess position patterns (Chase & Simon, 1973; De Groot, 1978; Vicente & De Groot, 1990). Multiple analogies exist, indicating similar or identical underlying cognitive processes in prototype formation and automatization. The frequency of an exposure to a stimulus is relevant for automatizing processes, as shown in many experiments, and expressed in the formula for the power law of practice (LaBerge, 1975, 1976; Gordon D. Logan, 2002; Posner & Snyder, 2004; Samuels et al., 1978; Sternberg, 2008). Like prototype categories, proceduralized processes also determine areas of increased performance. Such processes require only minor resources of attention, and happen unconsciously or automatically (Posner & Snyder, 2004).

Chunking plays a key role for increasing performance in automatization processes. Chunking is also a major element of prototype categorization, and its effects have been researched in the area of Gestalt psychology. Prototypes are nested depending on the area of attention. Being able to recognize objects as a whole, or to change the focus of attention to sub-parts, is important for automatization processes, as well as for prototype categorization. Recognizing chunks at different levels of complexity is a key performance driver, as empirically proofed in multiple experiments (G. D. Bower, 2008; G. H. Bower, 1970, 1972; Chase & Simon, 1973; De Groot, 1978; Gobet & Simon, 1996a; Larkin et al., 1980; Lesgold, 1988; Reitman, 1976; Samuels et al., 1978; Vicente & De Groot, 1990). Increased chunk size and interconnectivity among information is a key to superior cognitive efficiency (Bryson et al., 1991; Chi et al., 1982; Larkin et al., 1980; Sternberg, 2008). Analogous to prototype categorization, dominant basic processes for recognizing and grouping information into chunks, are proximity and similarity (G. D. Bower, 2008; Sternberg, 2008; Ungerer & Schmid, 2006). In Gestalt psychology, other principles are continuity, closure, and symmetry (Sternberg, 2008; Ungerer & Schmid, 2006).

Many remarkable cognitive abilities, such as abstraction, categorization, or inference, are based on implicit or explicit similarity comparisons. Hereby, the structure mapping theory provides an empirical, evidence based model of human cognition, and underlines the pervasiveness of similarity comparisons. (Gentener et al., 2008; D. Gentner & Christie, 2008; D. Gentner & Markman, 1997; D. Gentner, 2003; Penn et al., 2008). Comparisons are based on similarity and fundamental for many pervasive cognitive processes, such as abstraction, categorization, or inference.

Metaphors build on mapping cognitive concepts from a base to a target domain, and therefore improve efficiency. Herewith, metaphors work similar to categories, which inherit salient features to category similar objects. Cognitive prototypes are used as reference points for new information, which is interpreted through similarity comparison. Some researchers see metaphors as a species of categorization (Glucksberg & Keysar, 1990; Glucksberg et al., 1997; Honeck et al., 1987; Kennedy, 1990). Indeed the boundary is vague, especially between metaphors and natural categories (E. H. Rosch, 1973b; Ungerer & Schmid, 2006).

Similarity comparisons are also fundamental for the recognition of semiotics. While icons purely build on similarity, indices work associatively via contiguity. Empirical evidence in cognitive research suggests a higher effort for learning indices, and typically a better understanding of new icons, rather than new indices.

2.2 Retrospective Property Extraction of Existing Digital Transformatives

As previously detailed, research on human cognition of mnemonic devices, metaphors, and semiotics provides some clear indicators for possible working principles of Digital Transformatives. In the following, the basic principle will be approached from a practical side by analyzing existing implementations meeting the determined schema of Digital Transformatives.

2.2.1 General Property Extraction Procedure

Methodologically, the procedure for finding the main efficiency features will be similar to a retrospective cohort study, often applied in medicine or social science. The extraction process follows the steps illustrated in Figure 2-21.

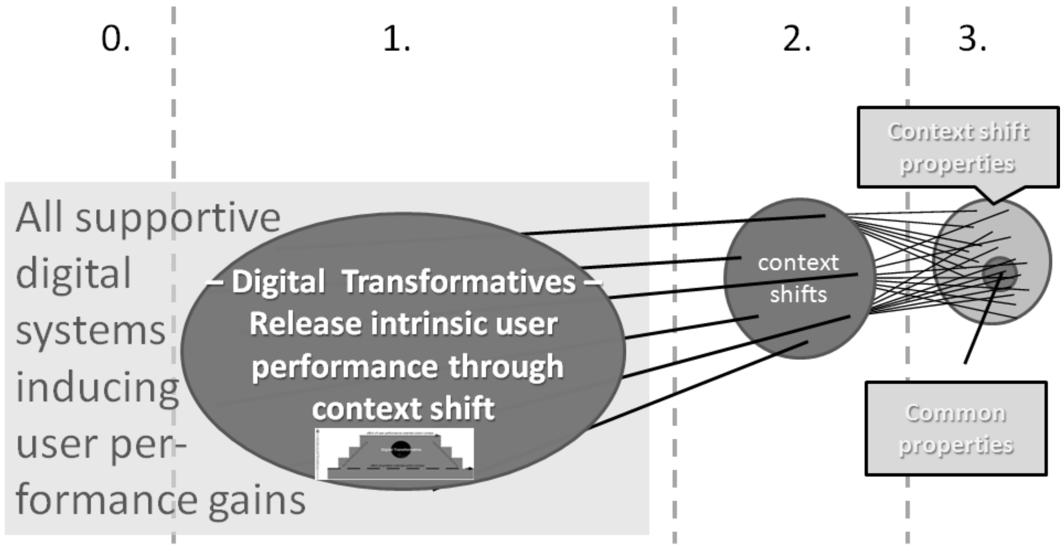


Figure 2-21. Property extraction methodology.

Starting from the vast pool of user performance enhancing digital systems:

1. Determine a set of relevant systems, shifting the user context.
2. Extract the context shifts implemented by those systems.
3. Collect characteristics of all context shifts, and find characteristics common to all of them.

The three steps of the procedure will be detailed in the following.

DETERMINE RELEVANT SYSTEMS SHIFTING THE USER CONTEXT

The analysis is based on a representative group of existing systems. According to the Digital Transformative schema, introduced earlier, two criteria are prevalent:

a. The systems should increase user performance.
b. Performance gains should be induced intrinsically by the user, and not programmatically by the system.

Car navigation software, for example, may reduce our navigational effort for reaching a destination by car; however, the performance gain is reached fully programmatically, since the system automatically determines the perfect route and generates instructions. Another example may be a password manager, which stores passwords and provides automated access to several systems. Users only need to memorize a single password, providing them access to their key chain. This way, the key chain offers an additional extrinsic tool to reduce user effort. Users might become highly dependent on such services. In case of the navigation system, drivers are often fully relying on the commands given by the navigation system. In case of the password manager, one does not use the separate passwords of the key chain, and tends to forget them in favor of the master password.

Hereby, it is often easy to exploit intrinsic potentials. For example, the memorization abilities of users may be much higher in contexts different from numbers (Abdullah, Abdullah, Ithnin, & Mammi, 2008; Biddle et al., 2011). Instead of prompting users with a number pad, such a system could provide a digital canvas, which maps simple drawing patterns to numbers (compare Jermyn, Mayer, Monrose, Reiter, & Rubin, 1999).

Hence, in this initial step of the procedure, it is the main task to distinguish, whether efficiency gains are caused extrinsically or intrinsically. This differentiation demands a first look at involved context changes.

EXTRACT THE CONTEXT SHIFTS

In a next step, the identified Digital Transformatives are investigated in detail. Therefore, possible shifts of the user context are emphasized and characterized. The transformation from a number to a pattern passphrase interface, as described previously, changes the usage context from memorizing symbols to memorizing spatial structures. Thus, this transformation may be characterized through a change in visualization, including features of spatial structuring.

Additionally, the investigated systems should also be characterized on a more fundamental cognitive level. In order to do so, it is necessary to derive relevant cognitive assessment metrics from previously determined cognitive mechanisms.

Cognitive Assessment Metrics for Digital Context Shifts

As described earlier, several major conceptual mapping instruments, such as categories, metaphors, or semiotics, seem to be based on cognitive efficiency structures. Hereby, prototype categories provide a reference system, determining areas of increased cognitive efficiency. In order to evaluate efficiency changes induced through context transformations, it is important to identify the prototypes addressed by each usage contexts.

As psychological research shows, prototypes are best determined implicitly. A major implicit measure is typicality. Thus, prototypes can be found by asking subjects for typical examples. Additionally, frequency of stimuli occurrence provides another implicit indicator for prototypes. Cognitive automatization is promoted through practice and repetition, which correlates with procedural familiarity. Consequently, first fundamental metrics are *typicality or familiarity* of a usage context.

Although basic levels are neither the most concrete nor the most abstract concept, it is not fully clear if this is also true for shared basic levels, at the area of maximum conceptual similarity. Consequently, it has to be investigated whether context shifts follow the tendency of *abstraction* or *concretization*. Hereby, it will mostly be assessed, whether feature richness increases or decreases.

The base concept might also always be associatively better interconnected with other concepts. Such kind of *knowledge elaboration* of information chunks has been determined to be a major difference between expert and novice knowledge. Expertise related performance gains is based on a more elaborate knowledge structure, characterized through concepts of increased chunk size, which are associatively well interconnected (Chase & Simon, 1973; Gobet & Simon, 1996a, 1996b, 1996c; Larkin et al., 1980; Lesgold, 1988; Reitman, 1976). Additionally, semantic elaboration seems to correlate with performance (F. I. M. Craik & Lockhart, 1972; F. I. M. Craik & Tulving, 1975; Hyde & Jenkins, 1973; Oberauer et al., 2005). It is hard to determine elaboration of knowledge without specific setups. Knowledge elaboration goes along with information being related to many other concepts. Thus, knowledge becomes more *meaningful* since it is useful in different contexts. Although, meaningfulness might also be related to prototype categories, it will here be seen as a basic indicator for knowledge elaboration.

Summarizing the above, the following types may be essential for the conceptual mapping from target to base:

- familiarity or typicality,
- abstraction and concretization,
- and meaningfulness through knowledge elaboration.

Interestingly, even historic sources vaguely accentuate those characteristics, as elicit earlier in this chapter on the pages 14 following. (da Signa, 1892) describes the signs for memory as manifest, concrete and remarkable (abstraction and concretization, meaningfulness). According to the classic view they are used to connect known information with new information (familiarity or typicality). Connecting associative thoughts improves memorization (meaningfulness through knowledge elaboration) (compare da Signa, 1892; Dommerich, 1765; Kant, 1792, 1800; Kästner, 1804; Voigt, 2001).

COLLECT CHARACTERISTICS AND FIND SHARED CHARACTERISTICS
Finally, the properties of all context shifts will be enumerated, giving an overview of all salient working principle characteristics. Ideally, there should be a set of properties common to all Digital Transformatives. In this case the principle of operation should be among those common properties, or may be characterized by all of them. This analysis, in combination with the previously conducted elaboration of cognitive processes, will broaden the conceptual basis of Digital Transformatives.

2.2.2 Determine Relevant Digital Transformatives

The starting point for the property extraction procedure is a sufficient fundament of sample systems. The challenge lies in selecting systems, which enhance performance user intrinsically, and cover the full spectrum of performance relevant characteristics. Therefore, it is assumed that performance categories correlate with major learning categories. One of the most elaborate taxonomies in developmental psychology, offering a sound structure of learning, is given by Bloom's taxonomy (L. W. Anderson, Sosniak, & Bloom, 1994; Shane, 1981), which offers a taxonomy of educational objectives driven by the classification of educational goals. The taxonomy is considered to be a foundational and essential element within the education community. According to Bloom's Taxonomy, the main human learning domains are cognitive, affective, and psychomotor. In the following, several systems will be analyzed for each of such categories, as shown in Figure 2-22.

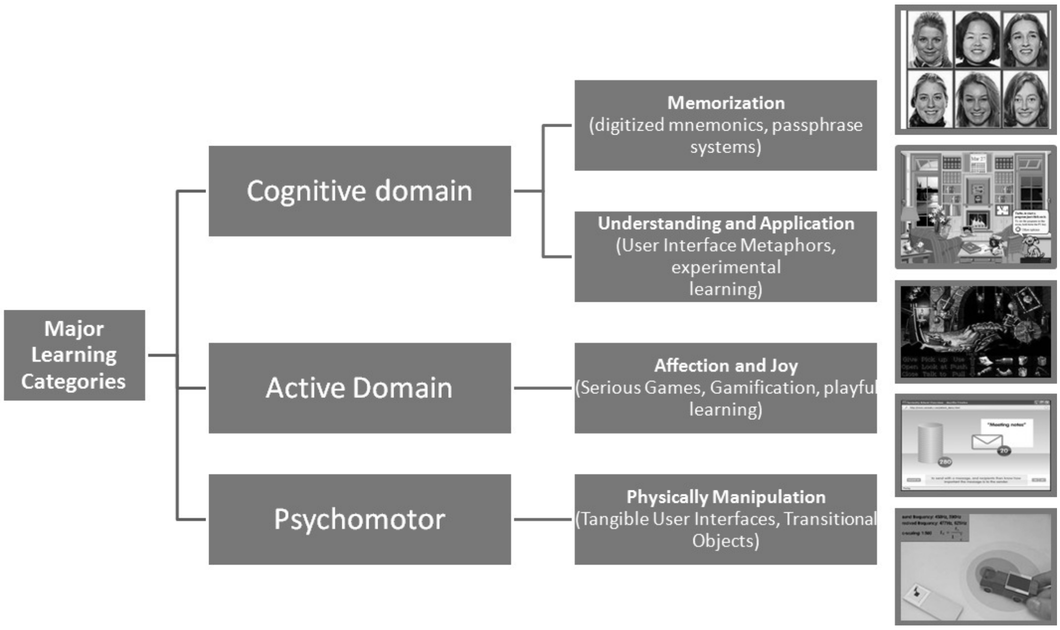


Figure 2-22. Digital system types addressing major learning categories of Blooms Taxonomy.

The cognitive domain will be covered mainly through memorization systems, such as digital mnemonics and passphrase systems. Additionally systems will be investigated, which utilize user interface metaphors to cover human understanding. Serious games, Gamification approaches, and systems implementing playful learning, seem to be good examples of the affective domain. The psychomotor domain is represented through Tangible User Interfaces and Transitional Objects.

2.2.3 Cognitive Domain

The cognitive domain is covered by myriads of systems which provide context shifts for better memorization. These tools are situated in the area of learning, in the following referred to as digital mnemonics, and also in the area of data secured through passphrases. Ideally password systems are also evaluated for task efficiency, since they seek to improve memorization abilities, and hereby increase security.

DIGITAL MNEMONICS

In this text digital mnemonics are referred to as digital implementations of mnemonic devices. *Joglab* offers a word finder tool intended for supporting the creation of acrostics or backronyms. As shown in Figure 2-23, word selection lists for each letter of an input acronym are offered.

Figure 2-23. JogLab online word finder supporting the creation of acrostics and backronyms.

The word finder allows for the creation of personal word hotlists, and the webpage provides multiple chances to share acrostic. (joglab.com, 2012).

2Know is a computer program offering support for the Phonetic System, or Major System, as it is described in (Higbee, 2001). In the major system numbers are transformed into consonants and phonemes. Silent consonants and vowels are not mapped. For example 2 is mapped on the letter n, since a tilted 2 looks like an n. 3 is mapped on m. Thus, 3 may be remembered by the word “me”, “home” or “aim”. 23 could be encoded in the word “name”. 2Know supports number to word encoding, and decoding of simple words or complex phrases, by offering the input of a number. The system then encodes such numbers and suggests possible words fitting the major system technique (see Figure 2-24).

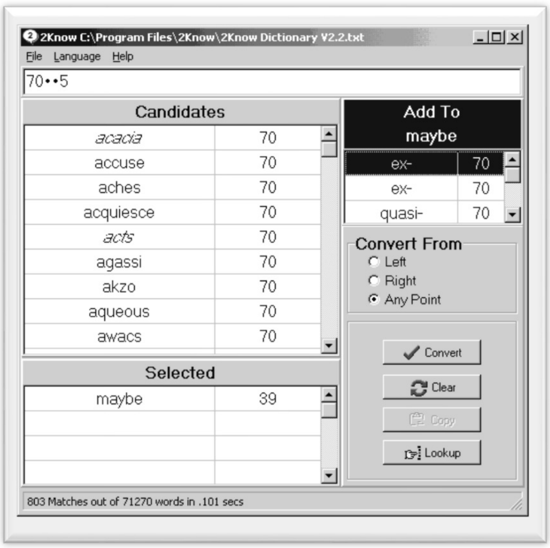


Figure 2-24. Screenshot of the 2know software for supporting the major memorization technique (right).

According to the homepage, the software is not meant to fully replace the major technique, applicants are rather supposed to be supported in finding good phrases and words (got2know.net, 2012).

Password security has always been a hot topic. Growing processing power demands increasingly better passwords. Most passwords inputs are textually, which implies a complicated dilemma: Good passwords need to be long and abstract, at the same time. While such passwords are more secure, they are also harder for us to remember. (Klein, 1990) investigated 14000 Unix passwords, and found that 25% could be cracked by using a dictionary of 3 million meaningful words. Other studies also show that the practical password entropy is far lower than the theoretically possible one, due to our preference for meaningful phrases (Feldmeier & Karn, 1990; R. Morris & Thompson, 1979; Wu, 1999). The dilemma of good passwords has been faced with several approaches. Many of them were digital implementations of mnemonic devices. Hence, passphrase systems provide a great pool of applied Digital Transformatives.

(King, 1991) suggested to use computer generated rebus passwords, based on the Keyword method as described in (Atkinson & Raugh, 1975). The Keyword method is usually applied for learning foreign languages. Translations are conducted via two links:

1) an acoustic link from the foreign language to a keyword, which is similar in pronunciation,
2) an imagery link from the keyword to the meaning.

For example, the Spanish word for horse, caballo, is pronounced “cob-eye-yo”. This gives us the keyword “eye”, which is used to create a remarkable mental image, such as a cyclopean horse with only one eye in the forehead. A good example for this method can also be found in (Beniowski, 1842, p. 36ff), where he describes his method for learning Latin names of plants. The Rebus Password Mechanism uses a variation of the keyword method to improve memorization of arbitrary computer generated passwords.

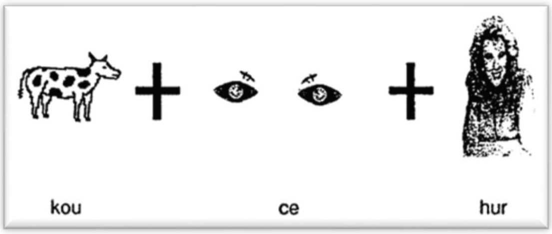


Figure 2-25. Supporting imagery link provided by the rebus system for the abstract password “kou-ce-hur” (King, 1991).

In this variation, the computer generates a sequence of random keyword syllables, and suggests associations to a sequence of similar phonemic sounds. All together is displayed to the user as a Rebus consisting of images, as shown in Figure 2-25. The Rebus Password Mechanism is meant to help users memorize random passwords more easily.

Many other mechanisms make use of our imaginary abilities, usually referred to as Graphical Passwords. Some build on the recognition of faces, some on creation and memorization of geometrical shapes. Others offer the possibility to encoding and decode information by selecting and recalling spots of interest in a given picture.

(Wiedenbeck, Waters, Birget, Brodskiy, & Memon, 2005b) suggest categorizing graphical password systems in pure recall based, recognition based, and cued recall based systems. In pure recall systems users do not get any hint for remembering their passwords. In recognition based systems passwords are set by users choosing images, which have to be identified among a bigger set of images for authentication. In contrast to recognition based systems, cued recall based systems provide the user with a reference frame of hints for authentication.

One of the first graphical passwords was presented in (Jermyn et al., 1999). The system called Draw A Secret (DAS) allows users to roughly sketch something on an input grid. Based on the grid the input is transformed into a binary number representation (see Figure 2-26).

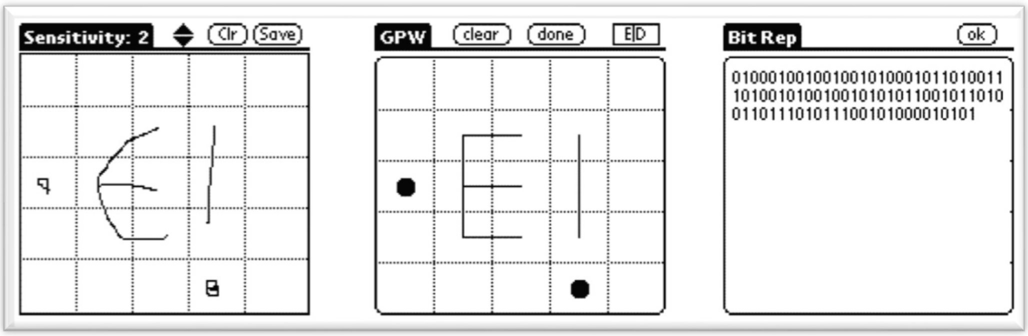


Figure 2-26. The Draw A Secret user input (left). The grid based interpretation of the system (middle). The transformed bit representation (right); from (Jermyn et al., 1999).

(Jermyn et al., 1999) argues that DAS is harder to crack then textual passwords. They state that the theoretical password space of a 5 x 5 grid exceeds the one of textual inputs, and that a grid provides a broader set of shapes for generating memorizable content, opposed to meaning full words, which are often used for textual passwords.

The commercial product PassFaces™ offers a login mechnism through memorization and recognition of certain faces, which ought to be recognized among random unknwon faces (as shown in Figure 2-27)

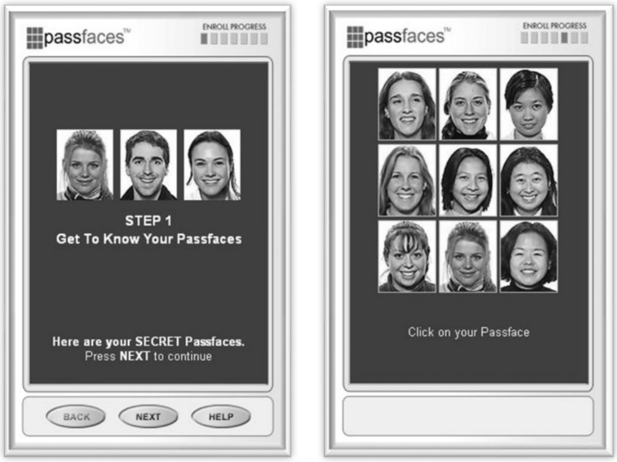


Figure 2-27. PassFaces™ learning screen (left) and selection screen (right) (passfaces.com, 2012).

A comparative study with 34 participants indicates that PassFaces™ passwords are easier to remember than textual passwords, but also take more input time (Abdullah et al., 2008; Brostoff & Sasse, 2000). (Davis, Monrose, & Reiter, 2004) tested the predictability of the PassFace™ system, and found that users' choices were biased by race, gender and attraction. Several other systems are similar to PassFaces regarding their performance context. The Cognitive Authentication Scheme demands users to memorize and recall iconic images (Figure 2-28).

Figure 2-28. Cognitive Authentication Scheme (Weinshall, 2006).

The usage is comparably uncommon. A password is encoded by traversing a grid based on game like movement rules. One needs to move down if an item is in the passphrase set, and move right if it is not. By following these rules one finally gets to the left or bottom boundary of the grid, which offers an encoded input. For a single login this traversal should be repeated multiple times, resulting in login times ranging from 1.5 to 3 minutes (Biddle et al., 2011; Weinshall, 2006).

The Deja Vu scheme also builds on the memorization and recognition of personal images (Rachna Dhamija & Adrian Perrig, 2000). The Déjà vu images are abstract random images. The login screen presents 25 images, including five images selected by the user. The login succeeds if those five images are determined. (Rachna Dhamija & Adrian Perrig, 2000) tested the system with 20 participants, and found that it was resistant to dictionary attacks, since only a few pictures were selected by more than one user. This gives indication that abstractness of images reduced predictability, however, this finding was based on a very limited test size and needs further evidence.

Cued-recall systems, also called locimetrics (De Angeli, Coventry, Johnson, & Renaud, 2005), form another set of graphical password inputs. In addition to recognition, they make use of the location of certain artifacts in an image. The basic scheme has been patented by (Blonder, 1996) (Figure 2-29 right). According to it, a password is encoded by selecting certain areas in an image, and decoded by recalling those positions in the right sequence. The click points need to be within a certain tolerance. Depending on this tolerance, and the resolution of the image, the possible password space may be comparably big. Many variations of this cued-recall graphic password scheme have been developed and investigated since then (Abdullah et al., 2008; Biddle et al., 2011; Wiedenbeck, Waters, Birget, Brodskiy, & Memon, 2005a; Wiedenbeck et al., 2005b).

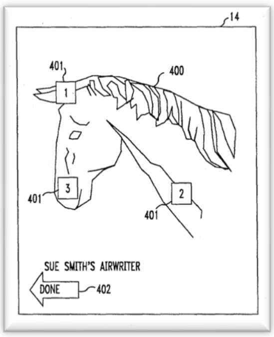


Figure 2-29. PassPoints interface, showing feedback after points have been chosen (left) (Biddle et al., 2011). Blonder patent for graphical passwords (right) (Blonder, 1996).

Although such passwords offer a relatively big possible password space, they also suffer from predictability (Dirik, Memon, & Birget, 2007). Users tend to select hotspots.

Based on the PassFace™ system (Davis et al., 2004) introduced the story scheme, and evaluated it against PassFace™, as described above. The story scheme presented a set of images showing persons, faces, and everyday items from categories such as food, automobiles, animals, children, sports, and scenic locations. In order to set a password, users required to choose 4 images in sequence. They were advised to memorize the sequence by connecting the images through a story. At the login process 9 images were presented, including the 4 user chosen images (compare Figure 2-30 left). For a successful login it was required to select those images in the right sequence.

Figure 2-30. Story system proposed by (left) (Davis et al., 2004). Password encoding through a "Repeated Sequence of Actions" (right) (Abdullah et al., 2008).

154 subjects participated in the study. The study found that the story passwords were less predictable than a chosen face phrase, while on the other side users had more difficulties in remembering the right order of the items. It turned out that many of those who could remember the items, but not the sequence, did not create their own story.

In contrast to the purely graphical approaches Oracle-Passlogix Inc. offered an action based password approach, by providing a minimal interactive environment, which lets users conduct certain action sequences (see Figure 2-30 right).

Users could choose between multiple environments such as kitchen, bedroom, or bathroom, and interact with items via mouse clicks, or drag'n drop. This allows users to enter a password, for example, by selecting ingredients for the preparation of a certain meal in the right chronological order (Abdullah et al., 2008).

(Abdullah et al., 2008) investigated usability and security features of graphical passwords. They differentiate between recognition and recall techniques. Then they rate memorability by meaningfulness, human faces, organized by theme, user assign image, icon based, abstract image, navigating image, and freedom of choice. Additionally they rate efficiency through input reliability, accuracy, simplicity, and fun to use, for both, grid based and drawing passwords.

CONCLUSIONS

The systems word finder and 2Know both map random numbers or letters onto phrases. The Rebus Password Mechanism builds on mapping random letters phonemically on words which are represented through images. All those systems map random abstract concepts onto bigger chunks of information, which are more concrete, familiar, and well connected to other concepts.

Draw A Secret is based on a mapping of abstract random letters onto shapes. The visualization adds the possibility to spatially structure information in a bigger context, and provides a frame of reference for the comparison of formerly unrelated elements. The spatial dimension also increases possibilities for associating input concepts to other concepts, hence increasing meaningfulness. The systems PassFaces™, Cognitive Authentication scheme, and Déjà Vu also add a visual dimension. They map random letters onto faces, comics, or abstract images. In contrast to Draw A Secret, those images are not self-generated by the user. Therefore, they incorporate a familiarization phase for memorizing key visuals. In so called cued-recall systems, such as PassPoints, Blonder scheme, or Cued Click Points, users are requested to mark and recall certain areas in images. While those systems also add a spatial and visual dimension, other systems, such as Story Scheme or PassLogix, also require sequential correctness; this way they additionally build on narrative elements.

The context shifts of all systems, and the properties of those shifts, are summarized in the table below.

<table>
<tr><th>Performance context</th><th>Problem context</th><th>System Names</th></tr>
<tr><td>Words/ Phrases</td><td>Memorize random letters</td><td rowspan="2">word finder, 2Know, Rebus Password Mechanism</td></tr>
<tr><td colspan="2">Increase familiarity, elaboration, and concreteness</td></tr>
<tr><td>Generation, memorization and recognition of key visuals</td><td>Memorize random letters</td><td rowspan="2">Draw A Secret, PassFaces™, Cognitive Authentication scheme, Déjà Vu, PassPoints, Blonder scheme, Cued Click</td></tr>
<tr><td colspan="2">Increase familiarity, elaboration, concreteness, visualization, comparability, objectivation, spatial structure</td></tr>
</table>

		Points
Chronological action sequence	Memorize random letters	PassLogix, Story scheme
Increase familiarity, elaboration, concreteness, visualization, comparability, objectivation, narrative structure		

Predictability is one of the reasons for weak textual passwords, reducing the theoretical password space dramatically. Graphical passwords seek to improve security through longer passphrases, however, the problem of predictability stays the same. Users chosen passwords are biased. Although passwords offer a relatively big possible password space, they also suffer from predictability (Dirik et al., 2007). Users tend to select visual hotspots, in the same way they select common sequences for numerical passwords. Those preferences likely correspond to cognitive prototypes, as it is well illustrated in (DataGenetics, 2013). They investigated 3.4 million four digit pin numbers from released, exposed or discovered password tables or security breaches.

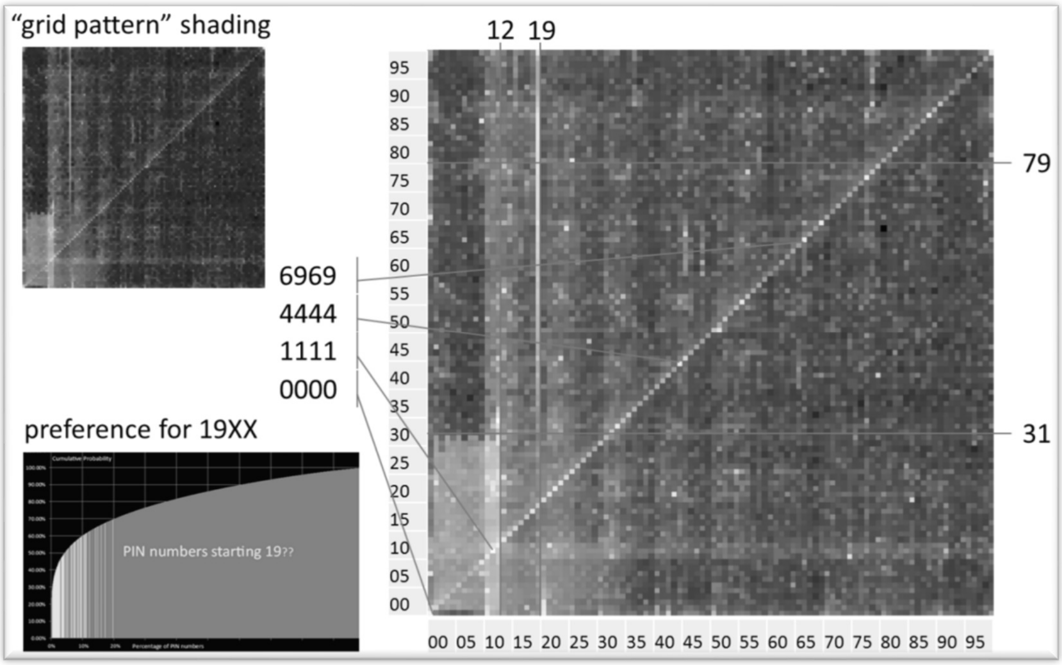


Figure 2-31. Visualizations of four digit passwords frequencies from (DataGenetics, 2013).

They found some interesting patterns, which may correspond to cognitive prototypes. In general the heat map reveals a "grid pattern" shading, although it should be uniformly distributed. The highlights in within the heat map mark areas of high frequency, which seem to correlate to number occurrences of high frequency in our external environment. For example the tendency to select passwords starting with 19, likely related to birth years. In the same way the combination of days and month is very dominant, as it is visualized in the heat map. Other repetitive occurrences can also be seen, and may not be related to prototypes, but rather to improved entropy, such as 0000 or 1212. Interestingly, the combination 6969 is dominant, within those repetitive patterns again, associated to a prototype again.

Although we seek to be highly individual, general environmental influences seem to induce general cognitive concepts, shared among a broader group of individuals. Cognitive prototype research on basic colors evidentially supports this view (Cook et al., 2005; P. Kay et al., 2009; P. Kay & Regier, 2003). As detailed in chapter 2.1, cognitive prototypes are formed probabilistically based on real world occurrences (Franks & Bransford, 1971; Neumann, 1977; Posner et al., 1967; Posner & Keele, 1967; Reed, 1972; Solso & McCarthy, 1981). Such findings may also be linked to cultural or social conditioning. Since our cognitive structure is heavily dependent on environmental influences, such influences may be important for Designing Digital Transformatives. Accordingly the findings are expressed in Feature 2:

Feature 2. Similar user environments induce similar cognitive prototypes (advantages and disadvantages of cultural and social conditioning)

The tested systems also indicate a correlation between abstractness and predictability (compare Biddle et al., 2011; Brostoff & Sasse, 2000; Weinshall, 2006; Rachna Dhamija & Adrian Perrig, 2000). (Rachna Dhamija & Adrian Perrig, 2000) tested the system with 20 participants and found that it was resistant to dictionary attacks, since only a few pictures were selected by more than one user. This gives indication that abstractness of images reduced predictability, however, the indication is loose due to the limited test size.

Investigations of the PassFace mechanism by (Davis et al., 2004) showed clear tendencies for race, gender, and attractiveness. Psychological studies also indicate a relation between the sense of attractiveness and cognitive prototypes. Several studies show that we consider an average face to be more attractive than a face which does not lie on a probabilistic maximum (DeBruine & Jones, 2013; Langlois & Roggman, 1990; Langlois et al., 2000; Rhodes, 2006). The *Averager* on faceresearch.org offers a good tool to interactively experience this correlation (DeBruine & Jones, 2013). We are probably most familiar with average faces, because we recognize a little bit of it in every face. In this case, familiarity correlates with attractiveness. Because cognitive prototypes tend to be hot spots of high user efficiency, and they are related to attractiveness and familiarity, this may be a valuable characteristic for Digital Transformative design.

Feature 3. Familiarity correlates to attractiveness, trust, or faithfulness. Both are related to cognitive prototypes.

Both, Feature 2 and Feature 3 help refining the picture of Digital Transformatives, however, they are only weak implicit indicators for Hypothesis 1 (*The system usage context shift of Digital Transformatives releases user intrinsic potentials.*).

2.2.4 User Interface Metaphors

User interface metaphors are an interesting case, since they can be well compared to linguistic metaphors, which are cognitively comprehensively analyzed. User interface metaphors are implemented in various facets in user interfaces. Hereby, it will mainly be focused on the economical successful desktop metaphor utilized in operating systems, which incorporates multiple sub metaphors. Additionally, spatial and narrative metaphors are being used with great success in learning and graphical point and click adventures.

THE DESKTOP METAPHOR

Nowadays, the *Desktop Metaphor* is probably the best established human computer interface metaphor (Dix, 2004; D. C. Smith, Irby, Kimball, Berplank, & Harslem, 1990). It was introduced by Xerox in 1981, and re-implemented multiple times by all major commercial operating systems provider, such as Apple, Microsoft, or Amiga (guidebookgallery.org, 2012). The desktop metaphor aims at improving computer usability by resembling a typical desktop through a metaphorical interface. Traditional command line interfaces were function and parameter oriented, on a noun-verb syntax (J. Nielsen, 1993). The graphically represented, familiar environment of a desktop, however, is likely to provide better anchors for guessing functions and understanding system states, without the need of learning commands (D. C. Smith et al., 1990, 1982). Thus, those graphical user interfaces are object based, and not function based. According to (J. Nielsen, 1993) this allows users to focus on their actual task rather than on operating the computer.

Desktop systems basically transform computer commands into actions, and data into virtual objects. Applications within such a desktop environment offer graphical user interfaces presented in *windows*. *Icons* are used to give files, or links, an object like look and feel, in the same way as *menus*, and *pointing* device interactions support the conceptual mapping of functions onto actions. With the desktop metaphor, still being the standard interface of current operating systems for desktop computers, this basic pattern has not changed much until today, and is often described as WIMP (*windows*, *icons*, *menus*, *pointing*) (J. Nielsen, 1993). This is also illustrated in Figure 2-32 – showing the first implementation of a desktop environment.

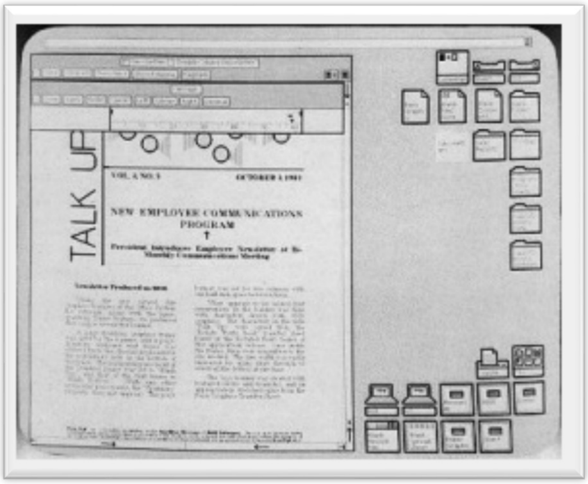


Figure 2-32. Screenshot of the first commercially sold Xerox 8010, in 1981 (guidebookgallery.org, 2012).

The wide use of the desktop is a strong indicator for its commercial potential. It offers user performance gains through improved usability, making it a primary example for Digital Transformatives.

The key to creating such a successful system has mainly been attributed to the reverse approach of starting with a conceptual model and tailoring functionality around it (D. C. Smith et al., 1990). A detailed development methodology of *Star* is being given in (Seybold, 1981). It emphasizes the importance of a task analysis preceding the design process. The design process was also lined by informal design principles, formulated before and during the Star development. Such guidelines and methodological experiences, are useful information for any further conceptual and methodological design of Digital Transformatives, and will be further investigated in according contexts later in this text. There seems to be no clear methodology, on how the idea of the desktop metaphor evolved. Only the final decision on implementing the Metaphor is being described "*We decided to create electronic counterparts to the physical objects in an office: paper, folders, file cabinets, mail boxes, and so on – an electronic metaphor for the office. We hoped this would make the electronic "world" seem more familiar, less alien, and require less training. (Our initial experiences with users have confirmed this.) We further decided to make the electronic analogues be concrete ob-*

jects. Documents would be more than file names on a disk;[...]“(D. C. Smith et al., 1990)..

Alternatively, those systems can still be operated via a command shell. From a metaphorical perspective, commands offer a textual semantic interface to the computer, historically driven by natural language communication. Since, these commands do not follow natural language semantics, and do not leave any space for interpretation, they, can be seen as an example of a deceptive metaphor. Deceptive metaphors show one of the major pitfalls of using this technique. Negative analogies evoke misunderstandings, leading to improper use of the system (Don Gentner & Nielsen, 1996; Halasz & Moran, 1982; Rogers et al., 2011). They are also often discussed in the context of user interface affordances, as detailed in *Affordances from UI Concepts* (pp. 99).

SPATIAL AND NARRATIVE METAPHORS

Other user interface metaphors build on spatial knowledge, in order to navigate through vast information spaces, for example. (Buchholz, 2005) describes how a spatial metaphor is used to organize content in a hypermedia educational game for kids, which is exemplary for many other games of that kind. Technically the game consists of separate pages with learning content, which are interconnected through hyperlinks. However, for the user they are presented in spatial conjunction, describing the environment around a site caravan Figure 2-33 (right). While one page shows the caravan, a click on the door opens the inside view. A click on the stairs, leading to the roof top of the caravan, would open the according webpage, instead. This way users do not get the feeling of requesting one web page after the other; they rather seem to explore a spatial environment.

Those spatial metaphors might have their origin in predecessors of *point-and-click adventure* video games, such as *The Monkey Island Series* or *Leisure Suit Larry* (M. L. Black, 2012; Dillon, 2004; LucasArts, 1990; Sierra Entertainment, 2006). Adventure games offer a story based artificial environment, including an avatar controlled by the user (Cavallari, Hedberg, & Harper, 1992). The avatar is able to interact with predefined objects of the game world environment. The game world consists of spatially associated locations, as described above. Locations usually consist of still images showing a scene, the avatar, and interaction objects including non-player characters (NPCs) controlled by the computer, as shown in Figure 2-33 (right).

Figure 2-33. Screenshot Monkey Island 2: LeChuck's Revenge (left) (Warren, 2003). Screenshot Löwenzahn (right).

Graphical point and click adventures evolved out of text-based adventures, such as ADVENT, or Infocom's Zork, where predefined textual commands had to be used to interact with the system (Dillon, 2004). Parallels between adventure games and operating systems are highly visible. In both cases command based syntax is transformed into actions performed on objects. On the other hand, games much more relied on spatial and narrative metaphorical structuring then desktop based operating systems.

During the mid to late nineties, some approaches for graphical user interfaces of operating systems have been conducted aiming for advanced spatial structuring, as it was known from point-and-click adventure video games. Most noticeable were the approaches made by *Microsoft Bob* (Microsoft, 1995), *Packard Bell Navigator 3.5* (Packard Bell, 1995), and *Magic cap*(General Magic, 1994). Those systems have been said to be utilizing the *home metaphor*. The reference to a home metaphor becomes apparent from the screen shots shown in Figure 2-34 and Figure 2-35.

Figure 2-34. Screen capture of Microsoft Bob (Rose, 2008).

The most elaborate, of the mentioned systems, is *MS Bob*. In general, despite the narrative component, *MS Bob* appears like an extrapolation of the desktop metaphor to *point-and-click adventure games (pnc games)* of that time. Instead of an abstract iconic desktop representation, a house environment is presented consisting of various locations, such as the “Family Room”, “Study Room”, or the “Castle Kid’s Room”. Every location is composed of a background still image, representing a scene, and vector graphic overlays, representing objects. All Objects and scenes appear in three dimensional cartoon style drawings. Objects can be positioned on the screen plane, layered, overlapped, or resized, until they perfectly fit their graphical environment. The house metaphor allows for connecting different locations spatially through location anchors. For example, a click on the left door in the family room leads to the study room. The right door of the study room, accordingly, leads back to the family room (Rose, 2008; Toastytech.com, 2012a; TopWindowsTutorials, 2009). The home metaphor provides a metaphorical ground for seamlessly mapping almost every operating system feature, such as public and private rooms, as they are incorporated in *MS Bob*. Public rooms are shared with other users of the system, while private rooms are restricted to the logged in user.

MS Bob also utilizes many aspects of *pnc games*, which are not typical for operating systems. For instance, the integration of clickable, non-functional objects of decorative nature, only, *MS Bob* also builds on a character based communication. Instead of cryptic human-computer communication, via simplified yes/no dialogs, the system offers animated characters. Those characters primarily serve as guides, meant to replace manuals. The communication with those pncs mimics expressive prosaic human to human conversation, based on multiple choice answer selection.

However, although *MS Bob* implemented many aspects relevant for making a popular *pnc games*, commercially it turned out to be unpopular and unsuccessful (Harrington, 2009; McCracken, 2010). This makes *MS Bob* an interesting case for this work, providing an example of what can go wrong when designing and marketing Digital Transformatives.

Packard Bell Navigator 3.5 and *Magic Cap* left out such typical *pnc game* features. Although they both built on the *home metaphor* they did not implement characters, transitional animations, or colorful cartoon style graphics. While *Packard Bell Navigator 3.5* provides users with feature rich graphics, including an environmental overview map (Figure 2-35 left), *Magic Cap* offers a much more iconic interface, being close to pure desktop metaphorical interfaces (Figure 2-35 right). Hereby, it may be noted that Magic Cap was originally designed for use with *Personal Digital Assistants* (*PDA*).

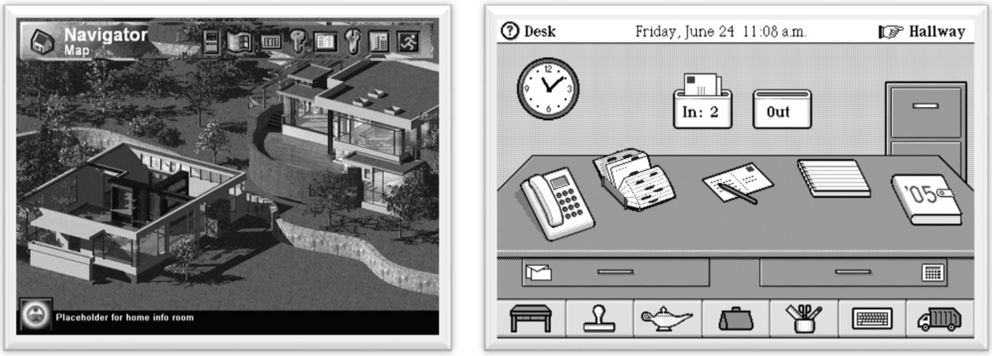


Figure 2-35. Screen capture of Packard Bell Navigator 3.5 (left) (Toasty-tech.com, 2012b), and Magic Cap (right) (Halfhill & Reinhardt, 1994).

Compared to *MS Bob* both systems were considerably more successful, being available in multiple versions for several years. The reasons why none of them became a full success are complex and also related to economic strategies. It might also simply not have been the right time. Many aspects of such early home metaphor systems were re-implemented later, more successfully. Having different rooms, and offering an overview map, can still be found in modern desktop systems in the form of multiple desktops, and the chance to set different desktop backgrounds. In the same way, icons tend to get more feature rich, including higher resolutions and more interactivity, with every evolving generation of new operating systems. Widgets, such as the clock in MS Bob, are common, nowadays; in the same way natural language interaction is getting more important, especially on mobile devices. Applications such as Siri (Apple Inc., 2011), also show their commercial potential. However, this may not be attributed to the interaction with characters in *MS Bob*, and it is much closer to what Bill Gates envisioned when Bob was presented at the CES in Vegas (Cheifet & Kildall, 1994).

CONCLUSIONS - USER INTERFACE METAPHORS

Although the previously described systems were developed independently in different contexts, there seem to be certain prevalent metaphors, utilized by most of them. One can be described as the objectification of data or files. This kind of metaphor is used in operating systems, were files are represented through iconized objects, as well as in Point and Click Adventures, were simulation model states are represented through arrangement of graphical objects. The data-object metaphor is deeply entrenched in computer science, as it is a quite common metaphor for computer programming, referred to as object-oriented programming (A. C. Kay & Ram, 2003; AlanC. Kay, 1996; Sutherland, 1964). However, it is not clear if this kind of entrenchment is responsible for the pervasiveness of this metaphor.

Another metaphor defined in the basic idea of object-oriented programming is communication with objects, by sending messages (A. C. Kay & Ram, 2003). In most programming languages this mechanism is implemented through methods, or command calls, referred to an object. Such an object-command schema can also be found, as a basis for command shells, for controlling operating systems or Point and Click adventures. Commands are used to determine or change the state of an object. The earlier described metaphorical interfaces, build on graphically represented objects, which are spatially located in a visual environment. This way the system state is visualized, and can be directly manipulated through mouse interaction. The visual representation also gives further comparative indicators of possible actions and different system states. Point and Click Adventures, and Hypermedia Learning Systems, also make use of narrative elements for structuring sequential events. By mimicking a visual interactive environment, those systems also increase predictability, since certain object behavior may be based on similar real world behavior. In the same way, such systems invite users to experimentally explore functionality instead of learning command lists. The icons and natural language interaction approach may also improve accessibility. The narrative structure and graphical art work could also improve immersion.

By adding a visual and spatial dimension, knowledge elaboration is increased. Mimicking objects, and using spatial and narrative structuring, also increases familiarity, analogous to icons demanding less initial learning then indices or symbols in semiotics (compare *Efficiency in Semiotics* pp.50). In the same way, objects and actions can be considered to be more concrete than filenames and commands, as well as narration and a spatial structure also increases concreteness.

<table>
<tr><th>Performance context</th><th>Problem context</th><th>System Name</th></tr>
<tr><td>Objects (tangible, visually represented)</td><td>Data, Files;</td><td rowspan="2">Desktop Metaphor, Point and Click Adventures, Hypermedia Systems utilizing spatial metaphor (e-learning), MS Bob, Bell Navigator 3.5, Magic Cap</td></tr>
<tr><td colspan="2">Increase elaboration, concreteness, familiarity, comparability, interactivity & adds objectification, visualization, spatial structure</td></tr>
<tr><td>Actions (mouse interaction with visual representations or menu items, drag n drop on Locations)</td><td>Commands (move, delete, copy, start, stop, …);</td><td rowspan="2">Desktop Metaphor, Point and Click Adventures, Hypermedia Systems utilizing spatial metaphor (e-learning), MS Bob, Bell Navigator 3.5, Magic Cap</td></tr>
<tr><td colspan="2">Increase elaboration, concreteness, familiarity, comparability, interactivity & adds objectification, visualization, spatial structure, curiosity, predictability</td></tr>
<tr><td>Locimetric: (Location of Objects</td><td>System State</td><td>Desktop Metaphor, Hypermedia</td></tr>
</table>

(taskbar, folder Desktop, Possible Actions)		Systems, Point and Click Adventures
Increase elaboration, concreteness, familiarity, comparability, interactivity & adds objectification, visualization, spatial structure, curiosity, predictability		
Scene, Objects, Narration	Game state	Point and Click Adventures, Hypermedia Systems utilizing spatial metaphor (e-learning), MS Bob, Bell Navigator 3.5, Magic Cap
Increase elaboration, concreteness, familiarity, comparability, interactivity, immersion & adds objectification, visualization, spatial structure, narrative structure, curiosity, predictability		
(Pseudo) Natural Language Character Conversation,	Computer Communication	Point and Click Adventures, Hypermedia Systems utilizing spatial metaphor (e-learning), MS Bob
Increase elaboration, concreteness, familiarity, comparability, interactivity, immersion & adds objectification, visualization, spatial structure, narrative structure, curiosity, predictability, accessibility		

(D. C. Smith et al., 1990) describes some basic Star™ user interface design rules, which provide further insights into the interface. Every decision should be made in favor of providing something easy over hard, concrete over abstract, visible over invisible, copying over creating, choosing over filling in, recognizing over generating, editing over programming, interactive over batch (D. C. Smith et al., 1990). All those principles could also have been expressed with a more fundamental design guidance. Besides concrete over abstract, the design should always aim for familiar over unfamiliar. We see more things than we do not see, we copy or mimic more often than we create, we choose more often than we design, we recognize more than we generate, change things rather than create from scratch, we interact more than we plan, and because we have done hard things so often that they became easy.

2.2.5 Affective Domain

The affective domain will be covered by investigating games as metaphors, including Gamification and Serious Games. While many of those systems offer a competitive environment, the second part of this chapter aims at explorative playful environments, which are not strictly based on competition.

GAMES AS A METAPHOR - GAMIFICATION AND SERIOUS GAMES

From a performance context perspective, games are highly interesting for this work, because their design is reverse to common digital systems design.

As elaborated in chapter 1.5.2, common supportive digital systems are often designed problem based and not performance based. For example, if tourists need orientation for the exploration of a new city, a problem based approach could provide tourists with a digital map. The map gives an overview and a frame of reference for orientation. Digitally it could be enhanced with GPS localization, and proper zoom and pan functionality. However, the effort of orientation is left with the users, such as creating landmarks, estimating distances, or getting a sense of space.

A performance based approach handles the orientation effort for the users, and offers an interface, which builds on user intrinsic orientation. This could for example be an Augmented Reality overlay of the users' hometown onto the new city. The hometown includes landmarks, giving a better sense of space. As emphasized earlier, Digital Transformatives are aiming for user interfaces on performance level, much like games.

Unlike typical digital systems, games are conceptualized and designed with the goal of engaging players in episodes of pleasure and fun. Computer game design usually starts with a game concept, seeking to deliver an optimal performance context of joy. The concept design is followed by a system implementation phase, which focuses on providing and realizing the desired game concept. Players are finally provided with an interface mimicking a game world (Fullerton, Swain, & Hoffman, 2008; Nacke, 2005; Ryan, 1999). The mechanics, bringing this game world to life, are based on simulation algorithms. The simulation model's only purpose lies in providing the framework for a joyful game world user interface, transporting the game concept (Figure 2-36).

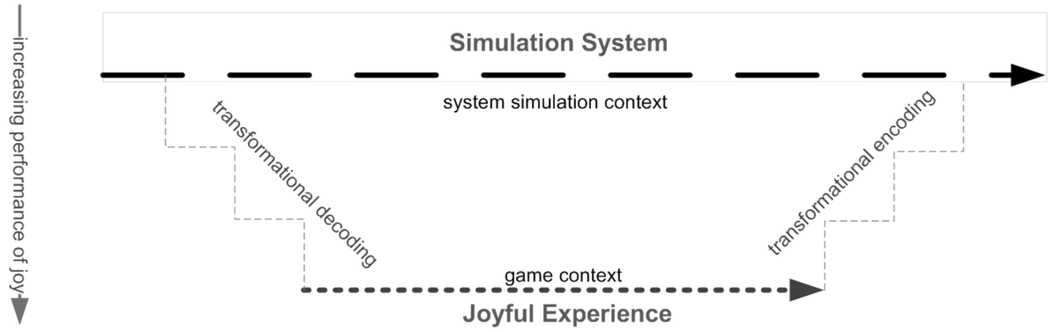


Figure 2-36. Typical computer game design pattern. The user interface is created based on a game concept. The game concept is developed for delivering a joyful experience in the game context. This experience is enabled through an underlying hidden simulation system.

Underlying digital system mechanics are implemented only for enabling a performance driven advanced user interface. They are designed reversely to application based digital systems. Games implement system mechanics only to provide a joyful user interface, while normal digital systems implement the interface to access system mechanics.

However, due to the increased engagement players show in games, it is logical to combine both approaches, and create playful experiences for real problems or valuable tasks.

Using games for conducting valuable work has been a hot topic ever since. Especially with the advent of computer mediation, it has been approached from various perspectives producing terms and trends, such as *Serious Gaming* (Abt, 1966, 1970, 1987), *Playful Learning, Edutainment, Game-Based learning, Games with a purpose* (Von Ahn & Dabbish, 2008; Von Ahn, 2006), or, most recently, *Gamification.* All those approaches imply that gaming is not meaningless per se. From research in the field of developmental psychology we know that play has the clear purpose of preparing, training, and acquiring new skills (Lerner, 1998; Mussen, Flavell, Carmichael, & Markman, 1983; Piaget, 1962). Research on creating purposeful games may easily be dated back to the 1900's (Avedon, Sutton-Smith, & Sutton, 1971; Juul, 2001).

Play in general is closely related to developmental psychology. Many models and concepts offer answers to fundamental question, such as the motivation of play. Dependent on the research perspective, concepts of play are interpreted in different forms. In the context of Gamification, social drivers of play are competition and socialization, and personal drivers are achievement, immersion, and exploration (Reeves & Read, 2009, p. 27).

The creation of joy and affection is as diverse as human individuals. Every one of us has different anchors of joy. A smell of something might remember one person of a great moment, while another is remembered of a sad incident. Someone feels joy when seeing a piece of art, while another does not even recognize that it is art. The study of joy, curiosity, or affection is a broad field, which cannot be covered within this work. Thus, play describes a very vague set of joyful experiences (compare Wittgenstein, 1953). Due to the diverse nature of joy, every system may turn into a joyful system for some users, in a way that some human beings are attracted to numbers, or fractals, while others are affected by lyrics, or paintings. While most of us consider crunching numbers as work, others associate it to be a joyful experience. Some like playing an instrument, others don't. From this perspective every system potentially induces joy and could be considered as a Digital Transformative. Therefore, for further elaboration this work will focus on the definition of games as a distinct subset of play, where compared to play in general, the boundaries of a game are set by a clear rules.

Due to the diversity and richness of the field of affection and fun, the following investigations will be limited to systems in the field of play – and here it will be focused on explorative and competitive software. The main explorative driver is curiosity. Thus, aspects such as immersion, mimicking, experimentation should be covered by such games. Competitive games utilize the urge for the comparison and improvement of skills. Additionally, competing in groups is part of many competitive games, which includes socializing.

EXPLORATIVE PLAYFUL ENVIRONMENTS

Business games are successfully utilized in education since several years. They usually offer environments for playfully experimenting with various business cases and hence train economical decision making. *MACRO* is a typical example for business games (Starbatty, 2009). *MACRO* models economic processes based on an environment consisting of two countries. Each of the countries holds various actors. Actors are enterprises, labor unions, federal banks, or governments. Players need to fulfill various tasks in order to win the game. Maximizing prosperity, assessed via consumption or savings indicators, may be one of such tasks. A set of instruments and parameters allows players to influence the game in order to reach their goals. By playing the game, users actually modify a simulation. For being successful it is essential to understand relationships among simulation elements, and understand the mechanics of the underlying model. The simulation model is a representation of real economic mechanics, making it easy for players to transfer knowledge about relationships of economic parameters and instruments into real world cases.

Simulation systems mainly build on curiosity. Players are curious to explore system mechanics in order to gain control and achieve goals. Hence, simulation systems are well established tools for learning complex real world principles of operation. The *World Water Game* provides an ecological simulation environment (Adib, 2006). Players have 45 minutes to provide optimal water and food supply for the world population. The user is presented with a graphical representation of the world, indicating the achievement status. Various actions are available for modifying the simulation, such as starting campaigns on family planning.

The *Power of Politics* is a political simulation where players are virtually able to start a political career (Powerofpolitics.com, 2012). They steer their virtual character by scheduling a political program and actions for raising popularity. The *Power of Politics* also incorporates real political information collected from newspapers. Thus, the simulation is also driven by real political situations and keeps its users informed. In the same way popularity measures for virtual political decisions are derived from the outcome of real cases. The Power of Politics was originally developed to work against political sullenness (Powerofpolitics.com, 2012; Preuster, 2010a). It was later supported by the Austrian government as an educational tool for high school students. (Parlament Republik Österreich, 2006). The creators of the *Power of Politics* also created a simulation game for playing a medical scientist, called the *Power of Research.* Players take over the role of doctors treating patients in their own hospital. The goal of the game is to become a successful doctor and virtually gain scientific reputation. Therefore, the players are able to virtually conduct research on DNA, cloning, microscopy, protein isolation, and so on. An included database provides real background information on given topics (Powerofresearch.eu, 2012; Preuster, 2010b).

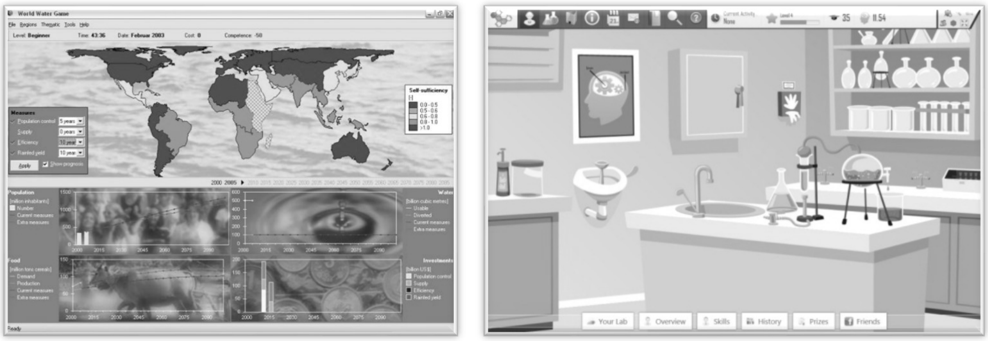


Figure 2-37. Screenshot of The World Water Game (left), and Power of Research (right) (Adib, 2006; Powerofresearch.eu, 2012).

As stated above, simulations combined with competitive tasks are common approaches for giving access to complex domains in a playful manner. Usual drivers for such simulation games are curiosity and competition. Those simulations cover many parts of natural play. Realism seems to be an important fun factor, as well as the chance to play roles. These aspects are underlined by the similarities between simulation games created for learning and simulations games created for entertainment. Consequently, many simulation games also make a good learning environment. Examples of such simulation games are *Microsoft Flight Simulator, The Sims, Sim City, Roller Coaster Tycoon* etc.

Not all simulation games are primarily developed for educating players. Numerous simulation environments are also offered to obtain user information, for the main part. A bigger group of those special educational games are online stock market games, which surely hold great potential to collect profitable user information. Online stock market platforms, such as *Wall Street Survivor, StockWatch Share Trading Game, The Stock Market Game™* (Smgww.org, 2012; Stockwatch.com.au, 2012; Wallstreetsurvivor.com, 2012), allow users to trade under real market conditions with virtual money (compare Figure 2-38). They help users to learn about stock market trading mechanics based on real data. The players input allows for evaluating economic strategies, as well as providing a *crowd-sourced* information ground for predicting the market.

Other approaches go a step further. They offer training platforms, primarily offered for analyzing strategies and determine players that may later be recruited for a real job. A publicly well rewarded system of that kind is *America's Army* (US Army, 2012). *America's Army* offers a virtual multiplayer online 3D-ego-shooter environment, allowing every player to start a virtual military career (screenshot shown in Figure 2-38). Players virtually move their *avatar* through virtual 3D environments. They are equipped with a weapon, and need to solve tasks in teams. The teams are hierarchically organized according to real *US Army* structures. The game is played for fun. The developers, on the other hand, use the game as a virtual test bed for new weapons and the analysis of combat strategies, as well as for the recruitment of talented soldiers. It also serves for propaganda purposes, at the same time educating and training its players (Galloway, 2004; Nieborg, 2004; US Army, 2012).

Figure 2-38. Online stock market platform Wall Street Survivor (left). Screenshot of the multiplayer online 3D ego shooter America's Army (right) (Wallstreetsurvivor.com, 2012) (US Army, 2012).

While education and training usually are considered to be a positive side effect, in this case, it is being criticized that glorify violence or play down danger situations may change behavior in a negative way.

All such simulation games commonly build on fantasy (immersion, narration) and curiosity (exploration). Therefore, they need to utilize immersive user interfaces as a medium for transporting complex relations of a subject domain through a simulation model, represented by game mechanics.

As previously mentioned, many of the systems are used to capture un-computational aggregated user data, such as combat strategies, political trends, economical behavior, or human judgments. Those systems are often classified as *Human-Based Computing* software or *Human Computation*. Some of such systems have been created especially for training AI and capturing common sense knowledge (Burgener, 1999; Lieberman, Smith, & Teeters, 2007; Von Ahn, Kedia, & Blum, 2006). Other games seek to improve search engines, for example, by letting users annotate images in a playful way (Russell, Torralba, Murphy, & Freeman, 2008; Von Ahn & Dabbish, 2004; Von Ahn, Ginosar, Kedia, & Blum, 2007; Von Ahn, Liu, & Blum, 2006), or more generally by letting users formulate proper questions to randomly shown websites (Ma, Chandrasekar, Quirk, & Gupta, 2009),

A well-known implementation of a Human-Based Computing software was the *ESP Game* (Von Ahn & Dabbish, 2004). It also proofed economical capabilities of this approach, since it was transformed into the commercially used product of the *Google Image Labeler* (Google Inc., 2011). The *ESP Game* is an online game for tagging images. Determining proper associations for images is a fundamental problem in image recognition and retrvieval. In the ESP Game two random strangers play together remotely, see Figure 2-39. The players are presented with the same random image. In a certain amount of time they need to input terms associated with the image. As soon as a matching term is found, both players get points and proceed to the next image. The goal is to gain as many points as possible, which might be credited with a good position in the overall high-score list (Google Inc., 2011; Von Ahn & Dabbish, 2004, 2008). Evaluating human generated input is not always trivial, and can become computational complex, as shown by *Tagatune*. Analogous to the ESP game, *Tagatune* aims at tagging music (E. L. M. Law, Von Ahn, Dannenberg, & Crawford, 2007). However, its evaluation is much more complex, since players tend to describe music more elaborately (E. Law, West, Mandel, Bay, & Downie, 2009).

Many of the recent developments in the area of Human Computation followed the pioneering work conducted by the group around Von Ahn, and it's so called design schema of *Games With A Purpose (GWAP)* (Von Ahn, 2006).

Figure 2-39. Games with a Purpose: The ESP Game (left), Tagatune (right)(Google Inc., 2011; E. L. M. Law et al., 2007; Von Ahn & Dabbish, 2004).

According to (Von Ahn & Dabbish, 2008) a game can be fully specified through a winning condition and rules. The rules should be defined in a way that players perform the right steps for solving the computational problem. The key for every game is to provide an experience of fun or enjoyment. In order to design a successful GWAP (Von Ahn & Dabbish, 2008) suggest to aim either for an *output-aggreement game*, *inversion-problem game*, and *input-aggreement game*. All three variants are based on random strangers playing together in pairs, but competing against each other at the same time. The three variants mainly differ in the way the players acquire points. The designer of a GWAP needs to choose whether players gain points when they agree on the same in- or output-data, or whether one player has to guess input data from given output data. When users have to agree on data, they try to find the most common data which could be guessed by the unknown co-player. Alternatively, in so called *inverse-problem games*, one player tries to generate proper output data, which helps the other player to guess the right input data. (Von Ahn & Dabbish, 2008).

Abstracting such design guidelines, GWAPs actually build on two major principles of operation. Enjoyment is exclusively fostered through competition. Hereby points and high score lists serve as the key drivers. Secondly, the acquiring points in ad hoc teams with unknown strangers can be seen as social surveillance, and is used to control the proper execution of the game.

Competition seems to be a simple but powerful tool for inducing engagement. As a driver it is also very common for many approaches of *Gamification* (compare Reeves & Read, 2009). Gamification basically is about adding game elements to non-game contexts and can be found in many products, nowadays (Deterding, Dixon, Khaled, & Nacke, 2011).

A typical Gamification example in the context of Digital Transformatives is *Attent* with Serios (Reeves & Read, 2009; Seriosity Inc, 2010). Attent tries to tackle the increasing information overload by introducing a virtual currency for information sent via email or other channels. The virtual money called *Serios* allows users to rate information. In turn they receive Serios from their recipients. This way one can give feedback on the value of received information.

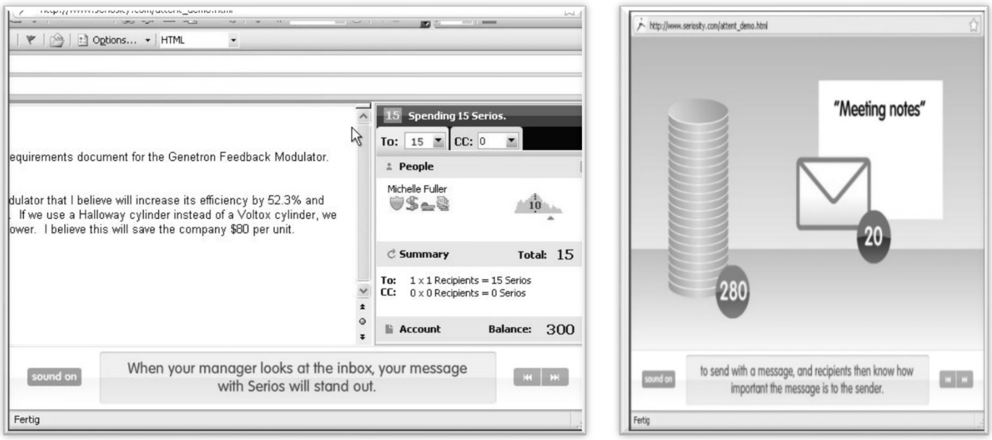


Figure 2-40. Screenshot of Attent with Serios as an Email extension (left) and in general (right) (Seriosity Inc, 2010).

Further examples for Digital Transformatives are given by the *email game* or Microsoft's *Ribbon Hero*. The *email game* adds a competitive time based point system to normal email communication (Baydin Inc, 2010). Ribbon Hero incorporates game elements into Microsoft Office products, for learning new office features in a playful manner. Office features are mapped on challenges, which improve the users' virtual Ribbon Hero skills and points. Every new use of an Office feature simultaneously advances the game state reflected through a user skill score. The game environment also gives feedback on unsolved challenges, for further advancement in the game (Microsoft Corporation, 2011, 2012).

The Nethernet, formerly known as *Passively Multiplayer Online Game (PMOG)*, consists of a browser plugin, which turns the whole internet into a multiplayer online game. The plugin integrates a toolbar as shown in Figure 2-41.

Figure 2-41. The Nethernet toolbar (thenethernet.com, 2012).

The Nethernet introduces the *Datapoints* internet game currency. Players earn Datapoints with every page visit and by time spent on website registered in the *Nethernet.* In turn, players are can invest their earned Datapoints to buy items from an arsenal of *tools*, *upgrades* and *abilities*. Participants of the game are able to interact or place those items, like *mines* or treasure *crates*, on web pages. Obviously mines are not good to interact with, while a treasure is always welcome. Players may also create *missions* which consist of a certain tasks, such as visiting a specific sequence of webpages. By fulfilling missions users improve their virtual character. Additionally, users are rewarded with *achievement badges* if they visit a certain page for a couple of days in a row, or avoid other pages, such as *google.com*, for a certain time period. The Nethernet also provides player interaction, building up a network of *Followers, Rivals,* and *Alies.* It also offers a narrative dimension explaining the history of characters. (Spiegel.de, 2008; thenethernet.com, 2012).

Game elements may not only be added to standard PC-Software. *Bottle Bank Arcade* provides a real world example, showing the power of engagement through competition Figure 2-42.

Figure 2-42. Bottle Bank Arcade systems for changing behavior with fun elements (Volkswagen & thefuntheory.com, 2009a).

The Bottle Bank Arcade project added scores and interaction to a normal bottle bank. Via a light installation, the modified machine gave users indication where to insert the next bottle. Hereby, the user gets point for quick and proper insertion. An additional high-score showed the best scores. The interactive bottle bank was used at maximum approximately 50 times more often than a nearby normal bottle bank (Volkswagen & thefuntheory.com, 2009a).

While points and competition is a valuable instrument for Gamification, curiosity may also be a good driver for changing behavior. Further curiosity driven exploratory installations of that kind are *The World Deepest Bin*, *Piano Staircase*, or *The Speed Camera Lottery* (Volkswagen & thefuntheory.com, 2009b, 2009c, 2010).

<u>CURIOSITY DRIVEN LUDIC EXAMPLES</u>

The World Deepest Bin basically consisted of a simple sensor and sound enhancement for a normal bin. Whenever pedestrians used the bin, a sound was played back from inside, giving the impression of a surprisingly deep tube. During one day twice as many pedestrians preferred the modified installation over a normal bin, standing nearby (Volkswagen & thefuntheory.com, 2009c).

For another curiosity based installation, a staircase, next to an escalator, was transformed into a big fully functioning clavier, as shown in (Figure 2-43 left). The so called Piano Staircase made 66% more people use the stairs than normally (Volkswagen & thefuntheory.com, 2009b).

Figure 2-43. Real world systems for changing behavior with fun elements. Piano Staircase (left), The Speed Camera Lottery (right) (Volkswagen & thefuntheory.com, 2009b, 2010).

The Speed Camera Lottery installation, altered the typical procedure of a speed camera. Every speeding fine, determined by the special speeding camera, as shown in (Figure 2-43 right), was collected in a lottery pot. In turn, all drivers, preserving the speed limit, automatically participated in the lottery with a chance to win parts of the pot.

Although all of the described evaluations conducted by thefuntheory.com were not representative, they give a good indicator on how curiosity can change behavior.

Conclusions

The systems described above can be separated in two groups. On the one hand, there are education based systems, offering learning and training support for users; on the other hand there are systems which aim for user generated content or user information. Both types of systems heavily rely on certain usage times to reach their primary goal. Hence, they seek to increase affection by implementing joyful elements. Resnick compares this mechanism to "*bitter medicine that needs the sugar-coating of entertainment to become palatable*" (M. Resnick, 2004, p. 1).

Reeves and Read (Reeves & Read, 2009, p. 27) determined social and personal drivers of play. They identified the social drivers of competition and socialization, and the personal drivers of immersion, exploration and achievement. The investigated systems mainly utilize the urge for exploration and competition. Exploration is driven by curiosity and competition by the need for improvement of skills and abilities. Additionally, competing in groups is part of many competitive games, which includes socializing. Based on the knowledge that groups are stronger than individuals, socializing may also be seen as improvement of abilities.

As stated above Games With a Purpose build on two major principles of operation. Enjoyment is exclusively fostered through competition. Hereby points and high score lists serve as the key drivers. Secondly, the social component of playing with an unknown stranger is used to control the proper execution of the game (Von Ahn & Dabbish, 2008). Gamification systems also fundamentally build on competition and achievements for evaluating personal growth. While game element enhancements typically consist of add-ons to non-entertaining systems, other approaches, such as serious games or simulation games, additionally offer a more immersive environment through a rich narrative context, garnished with realistic simulations, and beautiful art work.

In general, the shift to an affective layer seems to increase cognitive elaboration, since the game world adds additional anchors for players to relate information. Most of such systems also offer a more concrete usage context. Especially serious games or simulation games offer interactive use case scenarios for otherwise abstract concepts. However, systems, lacking immersion, often map to a context, which cannot definitely be considered to be more abstract or more concrete than the original one. For example, The World's Deepest Bin is not more concrete or abstract than a normal bin. In the same way it is hard to tell whether Tagatune is more concrete than simple tagging. Referring to the aim of changing user behavior by shifting to an affective context, it can be stated that all systems build on typical elements for affection, such as competition, exploration, and fantasy through immersion. The characteristics of context shifts of the investigated systems are concluded in the following table.

Performance context	Problem context	System Name
Competition or Points (if in line with other random user)	Tag images, image recognition, tag sounds and songs	ESP, Google Labeler, Tagatune
Increase of familiarity, elaboration, concreteness, competition, comparability, recognition, curiosity		
Competition (survival, ranks), Narrative Context (story to become a hero), Complex interactive Simulation (Beauty/ Realism of Simulation), Visualisation (3D Graphics), Objectivation (objects)	Recruiting Soldiers & analyzing war strategies	Americas Army
Increase of familiarity, elaboration, concreteness, competition, comparability, recognition, curiosity, objectification, interactivity, spatial structure, accessibility, immersion, beauty of simulation, narrative structure		
Competition (ranking through courses), Narrative Context	Trading Analysis	BörsenSpiel, Traders game
Increase of familiarity, elaboration, concreteness, competition, comparability, curiosity immersion		
Competition (ranking through courses, money, wealth aspects), Narrative Context (story to become a hero)	Learn about economics, or ecology	Business Games, Ecology, World Water Gate, Power of Politics, Power of Research
Increase of familiarity, elaboration, concreteness, competition, comparability, recognition, curiosity, objectification, interactivity, spatial structure, immersion, beauty of simulation, narrative structure		
Points, Achievements (personal growth), Competition	Use a bottle bank	Bottle Bank, Attent Seriosity, Ribbon Hero
Increase of familiarity, elaboration, concreteness, competition, comparability, recognition, curiosity, interactivity, immersion, narrative structure		
Curiosity, Instrument, physio skills	Use Stair about escalator	Piano Staircase, The Worlds Deepest Bin
Increase of familiarity, elaboration, concreteness, compe-		

tition, comparability, recognition, curiosity, objectification, interactivity, spatial structure, accessibility, immersion, beauty of simulation		
Play lottery, win money	Follow rules	Speed Camera Lottery
Increase of familiarity, elaboration, concreteness, competition, comparability, recognition, curiosity, objectification, interactivity, accessibility, immersion, beauty of simulation		
Play an online role game while serving the internet and interacting with others	Obtain serving behavior and social networks	The Nethernet
Increase of familiarity, elaboration, concreteness, competition, comparability, recognition, curiosity, objectification, interactivity, immersion, beauty of simulation, narrative structure		

2.2.6 Psychomotor Domain

The psychomotor domain has a comparatively long history in learning. Especially constructivistic advocates emphasize the importance of holistic hands-on learning experiences, which led Papert to develop the concept of transitional objects (Papert, 1980; Mitchel Resnick & Silverman, 2003). Transitional objects are seen as learning mediums for accessing new knowledge domains; this way they are conceptually meeting the schema of Digital Transformatives. Further digital systems will be analyzed in this chapter, which heavily build on current advances in Augmented Reality, Tangible User Interfaces, and body pose input (Ishii & Ullmer, 1997; Ullmer & Ishii, 2000).

COGNITIVE AND PRACTICAL BACKGROUND CONSTRUCTIVISTIC LEARNING WITH TRANSITIONAL OBJECTS AND DIGITAL MANIPULATIVES

In his book "Mindstorms" Papert elaborates on gears as *Transitional Objects* (Papert, 1980; Mitchel Resnick & Silverman, 2003). Papert developed an affection for cars, and everything associated with them, when he was a young child. This favor led to a distinct interest for gears, on a functional and emotional level. He projected many abstract problems onto his beloved gears, to give problems a connotation of pleasure. Piaget's work provided the epistemological basis for Papert's view on gears.

Piaget theorized that children must first construct knowledge through "concrete operations" before moving on to "formal operations" (Piaget & Mays, 1972; M. Resnick et al., 2009). During the past decade, a new wave of research has suggested that Piaget, if anything, understated the importance of concrete operations. Sherry Turkle and Seymour Papert, for example, have argued for a "revaluation of the concrete", suggesting that "abstract reasoning" should not be viewed as more advanced than (or superior to) concrete manipulations (Turkle & Papert, 1990). Piaget formulates the concept of a progression from concrete to abstract during children's stages of knowledge development, where children construct concrete operations first before they construct formal operations (Piaget, Wedgwood, & Blanchet, 1976).

Thus, based on a very strong emotional connection, gears gave Papert access to abstract mathematical ideas, while at the same time being connected to sensorimotor body knowledge. Papert was able to project himself into the place of gears to joyfully map abstract information on concrete objects. This way, they carried "powerful" mathematical concepts into his mind.

While gears gave good access to mathematical models for Papert, he was looking for a universal Transitional Object, which he found in the simulation power of computers. In this context, he worked on *LOGO Turtles,* shown in Figure 2-44.

Figure 2-44. Children playing with a LOGO Turtle (Logo Foundation, 2000).

LOGO Turtles are programmable real robotic objects, equipped with a simple pen tracing their movements. For the programmer the position, orientation, and pen are accessible. This gives possibility to implement algorithms for drawing shapes and other structures. Drawings are programmed through procedural commands by telling the turtle how to proceed from its current position. The procedural programming of geometrical shapes gives learners access to higher mathematical concepts, such as the angular sum of triangles, or the importance of the number pi. Later, when displays became less expensive, the physical turtle was more extensively used in a virtual variant, within the so called turtle graphics (Mitchel Resnick & Silverman, 2003).

Another constructivistic learning environment, building on programming computers, is Squeak. *Squeak* is a Smalltalk based authoring environment inspired by LOGO. It offers a full featured object based hypermedia environment for creating, accessing, and changing simple text, movies, sound, or even 3D virtual content. Squeak aims for a simple but powerful graphical user interface, allowing its users to adapt all parts of the system. Users may simply interact with given parts of the environment, modify existing objects, or create own simulation models and tools. The environment is meant to provide access to various levels of complexity, meeting the needs of novices, as well as experts. This way Squeak seeks to offer the “low floor” and “high ceiling”, as postulated by Papert (M. Resnick et al., 2009). Squeak is open source and its community tenders a variety of programming and authoring tools. One kind of such tools is *Etoys*, an authoring environment, which enables digital novices to create simulation models from a set of building blocks (A. Kay, 2005).

A specialization of Etoys may be seen in *Scratch*, which consists of a Squeak environment fully dedicated to programming with building blogs. Scratch is well connected with a web community platform for sharing projects (M. Resnick et al., 2009). While Scratch, Etoys, or Squeak focus on advanced graphical user interfaces, several other projects further built on the tangible idea of transitional objects described by Papert.

Resnick et al. (M. Resnick et al., 1998) introduced *Digital Manipulatives*, which put emphasis on learning with physical objects. The basic concept is the integration of computational and communications capabilities in traditional children’s toys. Information technology is implemented into toys for playful and experimental learning. The idea mainly focused on extending toys in a way that they can be programmed. Therefore, programmable bricks, so called *crickets*, where embedded into different kind of toys. These could be programmed, and even communicate with each other via infrared. For example, a common ball was equipped with a color LED, an accelerometer, and a programmable brick. The cricket could then be programmed to react on different ball movements detected by the accelerometer. This way “mood” could be mimicked by displaying a changing glow dependent on movements made with the ball.

A similar approach is undertaken by Lego with their so-called *Mindstorms* (LEGO.com, 2012). Mindstorms extend normal Lego blocks by adding motors, cameras, sensors and even a mini computer. Such computers can be programmed, enabling the building of a variety of different creations, which typically resemble simple robots (Bagnall, 2007). By using Lego Mindstorms children take first steps into programming. The usage of light or temperature sensors, on the other hand, allows them to learn about other traditional topics in physics (compare Roberta Fraunhofer IAIS, 2012).

Figure 2-45. Topobo interactive learning toy (Raffle, Parkes, & Ishii, 2004).

Another example for a Digital Manipulative is given by *Topobo* (Raffle et al., 2004). As illustrated in Figure 2-45, Topobo lets learners create real robots from a small set of simple generic building parts. The joints of these parts include servo motors, which are wired to electronics inside the housing of each part. Learners can connect multiple robot parts to bigger creatures. Such creatures are able to record and playback movements. For recording, the learner simply switches connected parts into record mode, and manually forces a movement (Raffle et al., 2004).

Technically Topobo makes great use of the ideas of Tangible User Interfaces (TUI). They enable users to interact with the computer in a natural way. Instead of using mouse and keyboard, the appearance, position, and orientation of physical objects is interpreted by the computer, providing specialized, well adapted input devices (Ullmer & Ishii, 2000).

Many other constructionistic approaches are building on tangible learning tools (O'Malley & Stanton Fraser, 2004). Recent approaches, such as the Science Center To Go, showed high potential in combining TUIs with *Augmented Reality* (AR) technology to enhance science teaching with a hands-on learning experience (Buchholz, Brosda, & Wetzel, 2010; Buchholz & Wetzel, 2009; Larsen, Buchholz, Brosda, & Bogner, 2012; Lazoudis et al., 2012).

(Buchholz & Brosda, 2012) determined the following fundamental schema behind Transitional Objects and their successors. The main goal of Transitional Objects lies in their function of helping learners to acquire new knowledge domains. This function may simply be achieved by raising interest for a new target domain. As previously detailed, Papert's love for gears gave him access to abstract mathematical models (A. Kay, 2005). Thus, an object rudimentarily works as a Transitional Object, if a learner has a strong emotional connotation to an object, which is used to interface a new target domain.

The transitional object needs to be known to the user and also be connected to the new matter. Therefore, it should be part of the learners' knowledge and the target domain as shown in Figure 2-46.

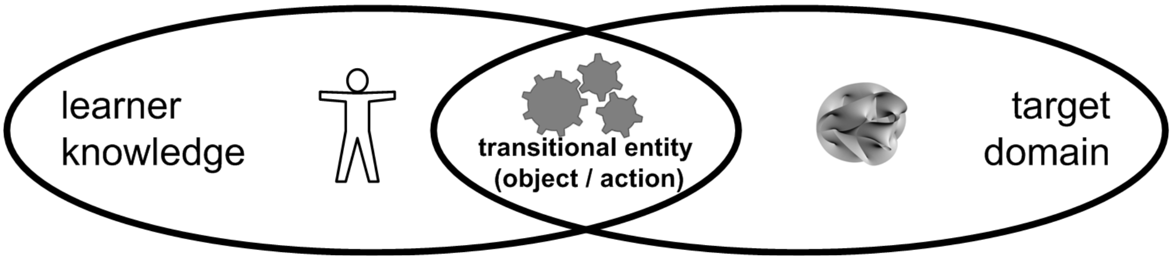


Figure 2-46. The transitional object interfacing the learners knowledge and new target domain. The area of intersection should be sufficiently big.

For Papert it is also important that such objects are tangible. The embodiment of gears, for example helped him to project himself into them. However, in the view of (Buchholz & Brosda, 2012) the transitional effect is not limited to tangible objects, and should also include actions. For example, dancing, singing, hiking, or playing an instrument, might help in acquiring otherwise uninteresting domains more easily. Oftentimes, one is not affected by an object but by its behavior. For instance, a ball would lose much of its attraction as a toy, if it loses its predictable behavior. Extending the idea of Transitional Objects by actions also extends the number of accessible target domains. Physical objects often limit the target domain to physical problems, which reduces access to certain domains, such as social interaction. (Buchholz & Brosda, 2012) refer to transitional objects and actions as *Transitional Entities*.

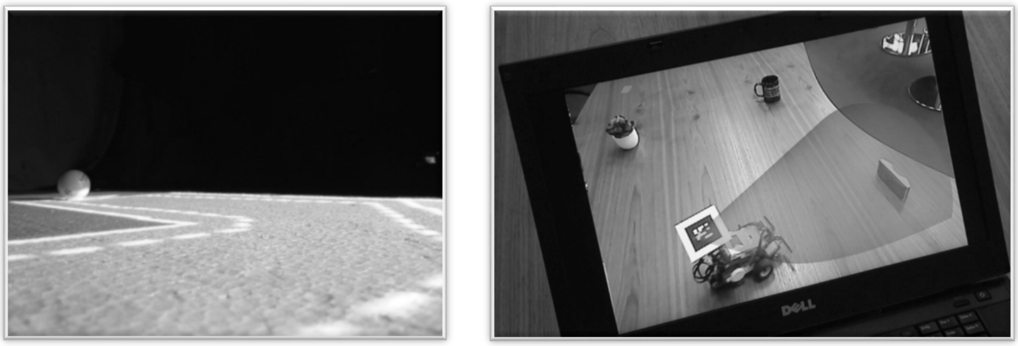


Figure 2-47. ARGolf (left), RobertAR(right) (Buchholz & Brosda, 2012).

They created two systems as experimental ground for further exemplifying and testing (Figure 2-47). With ARGolf they investigate the abilities that Augmented Reality technique offers to address multiple target domains using the simple transitional object of a mini golf setup. Via computer vision tracking the transitional object becomes the input interface for the learner. With RobertAR they created a more complex test ground, which also covers the interface to the user. Its generic appearance allows to mimic multiple transitional object.

TANGIBLE USER INTERFACES

Tangible User Interfaces (TUI) are specific human computer interfaces, building on direct sensomotoric human interaction. They seek to integrate computing power into everyday live by connecting physical objects to computers. The development of TUIs follows Mark Weiser's ubiquitous computing vision, of weaving technology into the fabric of physical objects (Ishii, Lakatos, Bonanni, & Labrune, 2012; Weiser, 1991).

Virtual digital functionality is assigned to real physical objects for a more intuitive and effective system interaction (Ishii et al., 2012). Interaction devices have a digital and a physical representation. Finding the right balance between those two representations is a major challenge in designing tangible interfaces (Ullmer & Ishii, 2000). There is no clear boundary between conventional and tangible user interfaces. (Ishii et al., 2012) compare graphical user interface controls, such as mouse or keyboard, with remote controls for virtual representations on the screen, while tangible user interfaces provide direct manipulation. However, a mouse itself is a tracked tangible object, showing the fluent transition between tangible and non-tangible interfaces.

One of the most influential early tangible interfaces has been *The Marbel Answering Machine*, designed by Durell Bishop (Ullmer & Ishii, 2000). The Marble Answering Machine resembles an answering machine, which provides graspable access to received phone calls. Every unanswered phone call releases a marble. Once such a marble is put back into the machine the recorded message of the caller is played back. Furthermore, if a marble is put on a telephone, the sender of the message will be called automatically. Hence the marbles serve as tangible keys to auditory information (Poynor, 1995).

A very active and influential group promoting the idea of tangible interfaces is the MIT Tangible Media Group (MIT Media Lab, 2012). Some of the prototypes for tangible interaction, such as *metaDESK*, *Phoxel-Space*, or *Glume*, will be described in the following (Ishii & Ullmer, 1997; Parkes, LeClerc, & Ishii, 2006; Ratti, Wang, Piper, Ishii, & Biderman, 2004; Ullmer & Ishii, 1997). The metaDESK system consists of a back-projected horizontal surface and tangible interaction objects, as shown in Figure 2-49 (Ullmer & Ishii, 1997).

Figure 2-48. The tangible user interface systems metaDESK (Ullmer & Ishii, 1997)

It implements several tangible objects as interaction tools. Those tools are physical instances of metaphors used in graphical user interfaces, such as a lens for zooming, physical icons (phicons) as placeholders, or a physical handle for panning. The interaction objects are identified and tracked by a camera system. Their usage is based on counterparts known from graphical user interfaces. For example, if one puts a model of a certain building on the desk, metaDESK automatically loads the according map, and pans it to the right position. The lens on the other hand, shows certain sections of the map in 3D, and can be used for zooming (Ishii & Ullmer, 1997; Ullmer & Ishii, 1997).

Phoxel-Space fosters the exploration of 3D voxel datasets through interaction with similar physical materials. The system consists of shapeable interaction material of different granularity, such as clay or sugar cubes. The shape of the material is tracked using a laser in combination with an infrared system. At the same time the top surface of the material is used as a projection plane, virtually changing the material texture, or superimposing other computer generated information. The system enables users to explore virtual 3D voxel datasets, such as seismic data, by digging or piling real material. Inversely they could also use the physical material for virtual modeling (Figure 2-49 left).

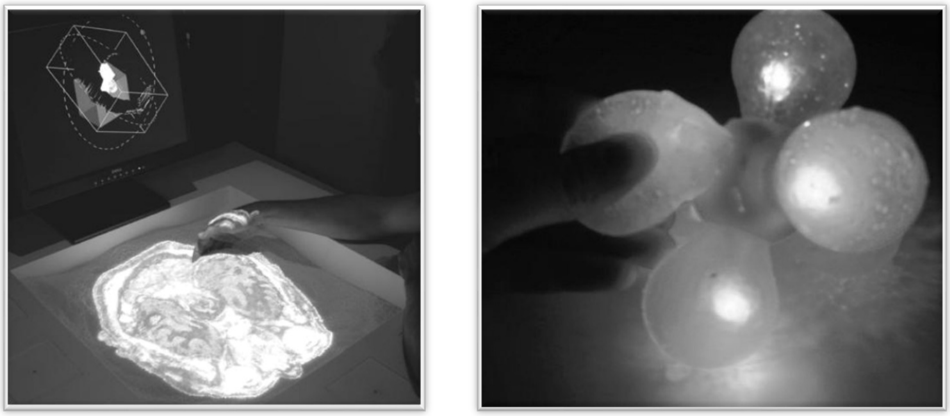


Figure 2-49. Phoxel-Space (left), and Glume (right) (Parkes et al., 2006; Ratti et al., 2004).

Another tool for physically constructing and manipulating virtual models, visualizations, and simulations of organic, three dimensional data sets, is given by Glume (Parkes et al., 2006). Glume allows users to build virtual models by interlocking multiple Glume modules, and shaping the nodes into place (Figure 2-49 right). A Glume module is a system of six bulbs, made of a transparent silicone skin, filled with ductile gel. The system automatically detects the morphology of the model. Furthermore, users can manipulate or retrieve the properties of Glume nodes through particular object modifiers or probes (Parkes et al., 2006).

While Glume and Phoxel-Space are not responsive, kinetic tangibles are equipped with force feedback mechanisms, such as Topobo described in the previous section (Raffle et al., 2004). Further examples for active tangibles are Goulthorpe's HypoSurface or Relief (M. D. Gross & Green, 2012; HypoSurface Corp, 2012; Ishii et al., 2012; D. Leithinger, Lakatos, DeVincenzi, Blackshaw, & Ishii, 2011). HypoSurface is a wall made up from panels, which can be actuated individually. This way, 2.5 dimensional shapes, such as water waves, can be displayed visually, and as an interactive relief (M. D. Gross & Green, 2012; HypoSurface Corp, 2012). Relief also offers a similar system, which is able to display 2.5 dimensional shapes, and lets users create or modify them (Ishii et al., 2012; D. Leithinger & Ishii, 2010; D. Leithinger et al., 2011). The relief interface is shown in Figure 2-50 (left).

Figure 2-50. Kinetic TUI Relief (left). Augmented Reality based TUIs Tiles (right) (Daniel Leithinger, Kumpf, & Ishii, 2009; Poupyrev et al., 2001).

The *Tiles* prototype is a mixed reality authoring interface for rapid prototyping and evaluation of aircraft instrument panels (Poupyrev et al., 2001, 2002). The prototypical implementation of Tiles consists of a metal white board, a set of paper cards, a book, whiteboard pens, and PostIts™, as shown in Figure 2-50 (right). The paper cards and book are equipped with fiducial markers and enhanced through Augmented Reality technology.

Unlike the phicons used in metaDESK, Poupyrev et al. attempt to detach physical properties from the virtual data, as much as possible. Therefore, Tiles gives an example of generic tangible interface controls. Tiles are not just placeolders for data (data tiles), but also for functionality (operator tiles & menu tiles). This allows users to dynamically work with data tiles, and modify them with basic operations, such as cut, copy, or remove. Physical objects like the whiteboard, a pen, PostIts™, or book tiles are used to add and organize content to the virtual tiles. However, the described version of tiles is only collaborative for co-located design, since the other physical objects are not tracked and digitized (Poupyrev et al., 2001, 2002).

ARTHUR is an augmented reality enhanced collaborative round table to support architectural design and planning decisions. It aims at closing the gap between CAD systems and Augmented Reality. Simple tangible placeholder objects have been used to improve interaction within a Augmented Reality enhanced 3D workbench (Broll et al., 2004). Figure 2-51 shows a use case.

Figure 2-51. Augmented Reality based TUIs in ARTHUR (Broll et al., 2004).

The users wear optical see-through glasses visualizing a virtual city on a physical table. They interact with this virtual model through a tangible wand and generic blocks, which are attached to virtual objects. A comprehensive overview of further TUIs may be found in (Shaer, 2009)

(Ishii et al., 2012) subdivide the spectrum of Tangible User Interfaces in two basic dimensions. On the one hand, a TUI may be static or kinetic. While static tangibles only serve as haptic input devices, kinetic tangibles also provide active force feedback, or are able to actively change their shape. Moreover, they differentiate between deformable tangibles and more discrete tabletop tangibles, which do not allow for continuous shape modifications. The systems described in this chapter cover all fundamental categories of this spectrum as shown in Table 2-1.

	Static/passive	Kinetic/active
2.5D continuous deformable tangibles	Phoxel-Space	HypoSurface, Relief
2D discrete tabletop tangibles	metaDESK, Tiles, ARTHUR	Topobo, Glume

Table 2-1. The systems described in this chapter, and their position in the spectrum of tangibles described by (Ishii et al., 2012).

Ishii and his fellows see affordances as one of the major challenges of tangible user interface design (Ishii et al., 2012; Ishii & Ullmer, 1997; Ishii, 2008; Ullmer & Ishii, 2000). “Tangible design expands the affordances of physical objects so they can support direct engagement with the digital world (Ishii et al., 2012, p. 38,39)”. Affordances seem to be a main advantage of tangible interfaces over graphical interfaces, and they are often considered to be an important performance driver in user interface design in general (Gaver, 1991; D. A. Norman, 1988, 1999; Preim, 1999).

2.2.7 Affordances from UI Concepts

The psychologist Gibson defined *Affordances* as all action possibilities for actors with their environment (J. J. Gibson, 1977; J. Gibson, 1979). He based the word Affordances on the verb afford. The word afford is often used in a monetary context, when someone does not possess enough money to buy certain things: "I cannot afford buying this bike". Gibson definition of affordances, however, primarily refers to persons' physical capabilities. If the person is capable of riding a bike he can "afford" to ride it. Throwing a bike could also be an affordance, if the actor is able to do so. However, eating it might most likely not be an affordance. Affordances are all actions that one can possibly perform with an object of his or her environment.

Based on Gibbson's objective view on Affordances Norman formulated the idea of *Perceived Affordances* (D. A. Norman, 1999). According to their name these are not necessarily actual Affordances but rather the actors' subjective understanding of possible actions with objects. Perceived Affordances might be quite different from actual Affordances. Let's imagine a fake, but completely real looking, cookie made of carbon. The cookie has an endless amount of objective Affordances: one could throw it, sit on it, or step on it. However, the predominant perceived Affordance might likely be eating. Affordances can also be misleading and incorrectly interpreted by the user (Gaver, 1991). The basic cognitive mechanisms of Affordances are also closely related to the concept of *the Law of the Instrument*, which is colloquially described through the popular phrase "*if all you have is a hammer, everything looks like a nail*" (Kaplan, 1964; A. H. Maslow, 1966, p. 15; A. Maslow, 1962).

Norman considers the awareness of Perceived Affordances as an important design momentum. A product designer, for example, could use knowledge about Perceived Affordances to design more intuitive devices. One of the standard examples in literature is the design of a door (Preim, 1999). An actor perceives a door as something he can open. A door might be opened in multiple ways. How would you, for example, open a door with a lever? Or what handle would you expect on a sliding door? The shape of the door or the door knob communicates its usage.

Norman argues that since "*[...] the required information was in the world: the appearance of the device could provide the critical clues required for its proper operation [...]*" (D. A. Norman, 1999, p. 39). Simply from their definition, Affordances already existed before Gibson expressed them, and designers surely already made use of Perceived Affordances before Norman formulated his view, however their texts raised the awareness for Affordances. Product designers might now iterate their work more often to see if they are able to encrypt useful operational information into their devices, not just because they know that it is possible, but because they want to create products with a better usability.

Norman argues that conceptual models, constraints, and affordances are essential for an individual's understanding of the operation of a novel device (D. A. Norman, 1999). Hereby, constraints were either of physical, logically, or cultural nature. Physical constraints are closely related to real affordances, logical constraints are based on the actors reasoning, and cultural constraints are described as social conventions, shared by a cultural group (D. A. Norman, 1999).

CONCLUSIONS

Similar to entertaining systems, described previously, transitional objects make use of affection to raise interest for new learning domains. The transitional objects described here put physical or mathematical models into a meaningful context. They hereby also raise familiarity to a new subject domain, and provide a concrete anchor for testing theoretical hypothesis.

Models are mapped onto tangible object behavior, which increases comparability and makes formulas more recognizable. It also increases the possibilities for quickly developing and testing hypothesis in an experimental setup, which allows for direct interaction. Force feedback devices, equipped with servo motors, such as Topobo, HypoSurface or Relief, increase interactivity and immersion even further.

The concept of Affordances has been identified as a major working principle of Tangible User Interfaces (TUI). The shape of an object implies its usage, which is based on prior experiences with similar objects. Consequently, Affordances and Tangible User Interfaces mainly build on familiarity with object usage, which also corresponds to the Law of Instrument coined by Kaplan and Maslow (Kaplan, 1964; A. H. Maslow, 1966). Users implicitly deduce further associations from familiar TUI object shapes and map it onto the virtual control. Hereby the TUI increases elaboration of Knowledge. TUIs also provide concrete multi modal access and are driven by physical constraints, which, on the other hand, lead to a predictable spatial and logical structure.

Performance context	Problem context	System Name
Programmable (physical affective) object	Physical or mathematical models	LOGO Turtel, Mindstorms, RobertAR, Transitional objects: (Scratch, Squeak, Etoys), Digital

Increase of familiarity, elaboration (meaningfulness), concreteness, comparability, recognition, tangibility, curiosity, interactivity, spatial structure, accessibility, immersion, beauty of simulation		Manipulatives (crickets), Topobo
Behavior recording and playback	Syntactic programming of behavior	Topobo, HypoSurface, Relief
Increase of familiarity, elaboration (meaningfulness), concreteness, comparability, recognition, tangibility, curiosity, interactivity, spatial structure, accessibility, immersion, beauty of simulation		
Affordance based tangible input control	Generic conventional input control	Phoxel-Space, metaDESK, Tiles, ARTHUR, HypoSurface, Relief, Topobo, Glume, Topobo
Increase of familiarity, elaboration (meaningfulness), concreteness, recognition, tangibility, interactivity, spatial structure, accessibility, immersion, beauty of simulation		

2.2.8 Summary of the Retrospective Property Extraction

Knowledge elaboration, typicality, and abstraction have been determined as possible fundamental characteristics, affecting user intrinsic efficiency in human cognition in the previous section. Additionally, the retrospective property extraction revealed further characteristics.

The conceptual mapping, induced by some of the investigated systems, builds on improved visualization or recognizability, for example when abstract passphrases are mapped on images segments. Increasing comparability has been determined as another important feature. Especially in the context of games, competition drives affection for otherwise uninteresting achievements. Comparisons with others, or former performances of ourselves, provide necessary feedback for competition, and for our self-estimation.

Many systems also build on mappings which increase interactivity, often accompanied by improved accessibility. Tangible User Interfaces, for example, improve interactivity due to raised usage familiarity, as expressed in the concept of Affordances. Games usually increase interactivity since game environments are more elaborate than underlying simulation models. Moreover, curiosity is a key driver for playful learning in experimental environments. Another curiosity based property, especially promoted by mappings into playful contexts, is immersion achieved by adding narrative structures. Narration is a familiar cognitive tool to weave information into context, and increase meaningfulness. Immersion is also often mediated through realistic comprehensive simulations or beautifying coherent artwork.

Graphical and Tangible User Interfaces also show the power of spatial structuring. Spatial structuring provides users with additional anchors for elaborating information cognitively. By implementing tangible properties into user interfaces, TUIs also make use of familiar features of haptic objects, this way, they also improve predictability of usage.

The context shift properties of all investigated systems have been individually determined and categorized, and categorized. The context shift properties of

- increased knowledge elaboration or meaningfulness,
- and increased familiarity or typicality,

seem to be of fundamental nature, since they could be found in all 76 investigated systems. Further distributions are shown in Figure 2-52.

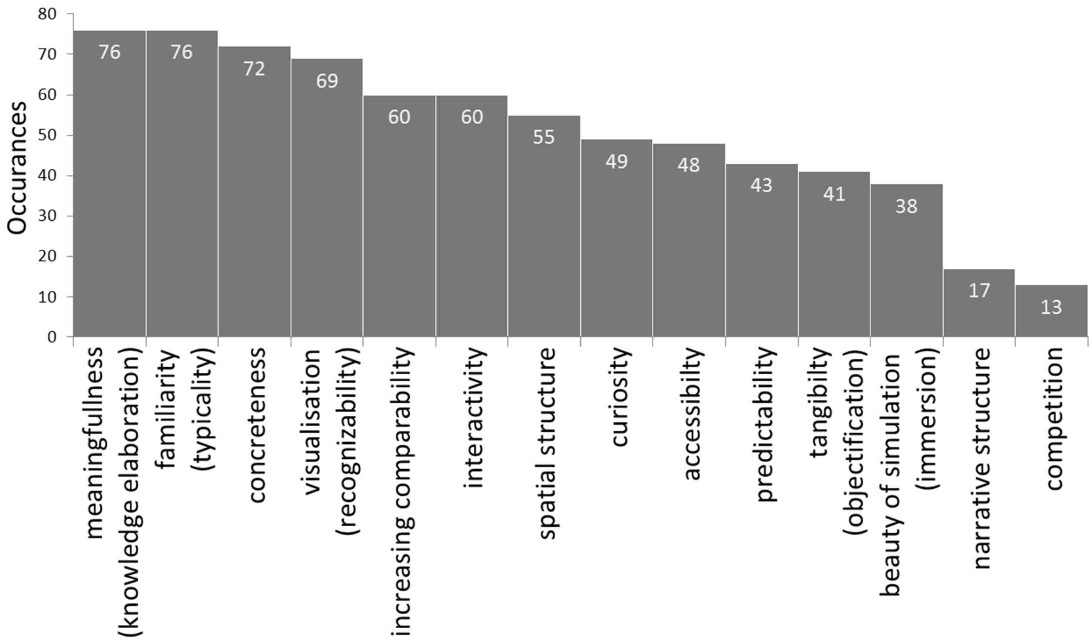


Figure 2-52. Accumulated context shift properties. The fundamental cognitive properties of meaningfulness (increased knowledge elaboration) and familiarity (typicality) can be found in all context shifts.

Also, 95% of the systems implemented a shift from an abstract to a more concrete usage context, and 91% build on improved visualizations. Narrative structures or competitive elements supplements are only used by 7% respectively 5% of the tested systems.

Based on the assumption that all investigated systems improved efficiency user intrinsically, as it is defined for Digital Transformatives, it can be concluded that, apart from the two omnipresent characteristics, none of the other properties can be the fundamental efficiency driver for such systems. This conclusion is invigorated by a hierarchical dependency structure, which can be found among the properties. The omnipresent properties are a basic part of all other characteristics.

Improved concreteness correlates with increased meaningfulness and familiarity. Improvements in visualizations also increase knowledge elaboration or typicality. Better comparability makes information more meaningful. Increased interactivity enhances knowledge elaboration. Enhancements achieved through spatial structuring are based on our familiarity with understanding spatial structures; the same accounts for narratives structures and tangibility. Improved accessibility corresponds to higher knowledge elaboration. Finally, curiosity and competition are very low level drivers of human action. Curiosity and competition makes us repeat certain actions, this way forms cognitive prototypes and increases familiarity. Predictability again, is based on such prototypes and on familiarity.

3 CONCEPT DESIGN OF DIGITAL TRANSFORMATIVES 105

3.1 CONCEPT OF COGNITIVE EFFICIENCY DRIVERS IN HUMAN COMMUNICATION 105

3.2 DIGITAL TRANSFORMATIVE CONCEPT - COGNITIVE PROTOTYPE BASED SYSTEM DEVELOPMENT FOR IMPROVED EFFICIENCY IN HUMAN-MACHINE COMMUNICATION 107

3.3 DIGITAL TRANSFORMATIVE MAIN CHARACTERISTICS 112

3.4 BASIC CONCEPT VALIDATION 115

3.4.1 Familiarity as a Driver for Human Efficiency 115

3.4.2 Memory - Basic Concept Test Case 116

3.4.3 Correlation between Familiarity and Performance in Team Sports 122

3.4.4 Conclusions of the Basic Concept Validation 128

3 Concept Design of Digital Transformatives

In the previous chapter the conceptual basis for Digital Transformatives was approached bottom up, from a cognitive perspective, and top down, based on an analysis of existing systems. Both approaches led to similar concepts with varying terminology and supplementing insights, indicating the common ground of these concepts. In the following, the views will be summarized and consolidated, first to a concept of efficiency enhancing mechanisms in human communication, and second to a concept of human machine communication, which also form the conceptual ground of Digital Transformatives. The bottom up approach started with investigations on mnemonic devices. Those investigations were further detailed through analyses on related cognitive linguistic mechanisms, such as cognitive prototypes, categorization, and conceptual metaphors.

3.1 Concept of Cognitive Efficiency Drivers in Human Communication

Mnemonic devices are cognitive techniques which may improve the memorization of information. Referring to the description given by (Voigt, 2001), a mnemonic device offers additional anchors (μ) between the new content to be remembered (A) and familiar information (B). This definition reflects a common understanding of mnemonic devices and can be summarized in a model as depict in Figure 3-1 (left). Hereby mnemonic techniques associatively increase meaning of new information by offering procedures for mapping this information on a familiar context. They help structuring information during the encoding phase to enhance the storage and recall of information in memory (Becker-Carus & Herbring, 2004; G. H. Bower & Clark, 1969; Ericcson et al., 1980; Raugh & Atkinson, 1975; Solso, 2005). Similar concepts can be found for conceptual metaphors. As graphically depict in Figure 3-1 (middle) metaphors use source concepts, as vehicles for explaining target concepts (Lakoff & Johnson, 1980; Ungerer & Schmid, 2006). The source concept inherits salient attributes to the target concept, which increases efficiency in communication.

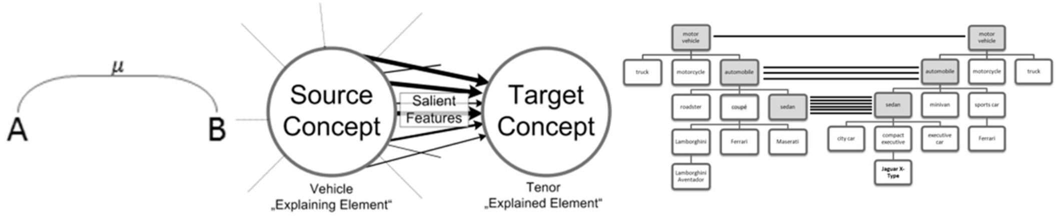


Figure 3-1: Schema of a mnemonic device (left): Familiar known information is linked onto new target information, predominantly based on salient features (middle). This allows us to use cognitive categories to improve efficiency in communication (right)

A rhetorical metaphor conducts a transfer of an expression from one subject domain to another, based on analogies or parallels between the both. A transfer only seems to be useful when the source subject domain is known (Strube et al., 1996). Cognitively, metaphors work on a level of similarity comparison. They are closely related to similes or analogies by working on items that share primary attributes (D. Gentner et al., 2001). Metaphors may also be seen as a species of categorization (Glucksberg & Keysar, 1990; Glucksberg et al., 1997; Honeck et al., 1987; Kennedy, 1990). Categories describe cognitive structures where concepts are organized based on common features or through similarity to a prototype (Coley et al., 1997; Hampton, 1995; Medin, 1998; Sternberg, 2008; Wattenmaker, 1995; Wisniewski & Medin, 1994). This way, categorization raises cognitive and communication efficiency in the same way conceptual metaphors raise efficiency (further detailed in *Cognitive Efficiency Catalysts in Communication* pp. 46).

The mapping of concepts is pervasive in many cognitive processes, and similarity comparisons are highly efficient inborn cognitive processes (Fauconnier & Turner, 2003; D. Gentner, 1983, 2003). As detailed earlier in this text, conceptual mapping is also fundamental for human communication instruments, such as categories, metaphors, and semiotics. Such cognitive mechanisms provide empirical ground for determining types of conceptual mapping that may improve efficiency. All of the investigated cognitive instruments build on conceptual mapping between base and target concepts for efficiently communicating new knowledge from speakers to listeners. Accordingly, the basic Digital Transformative schema, introduced in chapter 1.5.2, can be adapted to the cognitive communication schema shown in Figure 3-2. In this case, the new knowledge to be communicated is part of the speaker's knowledge, which is unknown by the listeners. As depict in the figure, conceptual mapping instruments make use of an area of shared cognitive concepts to transport knowledge more efficiently. An increase of efficiency demands a certain amount of similarity between base and the target concepts. The higher the transformational load, necessary to transform between base and target concepts, the less the communicational efficiency. Hereby efficiency correlates to similarity.

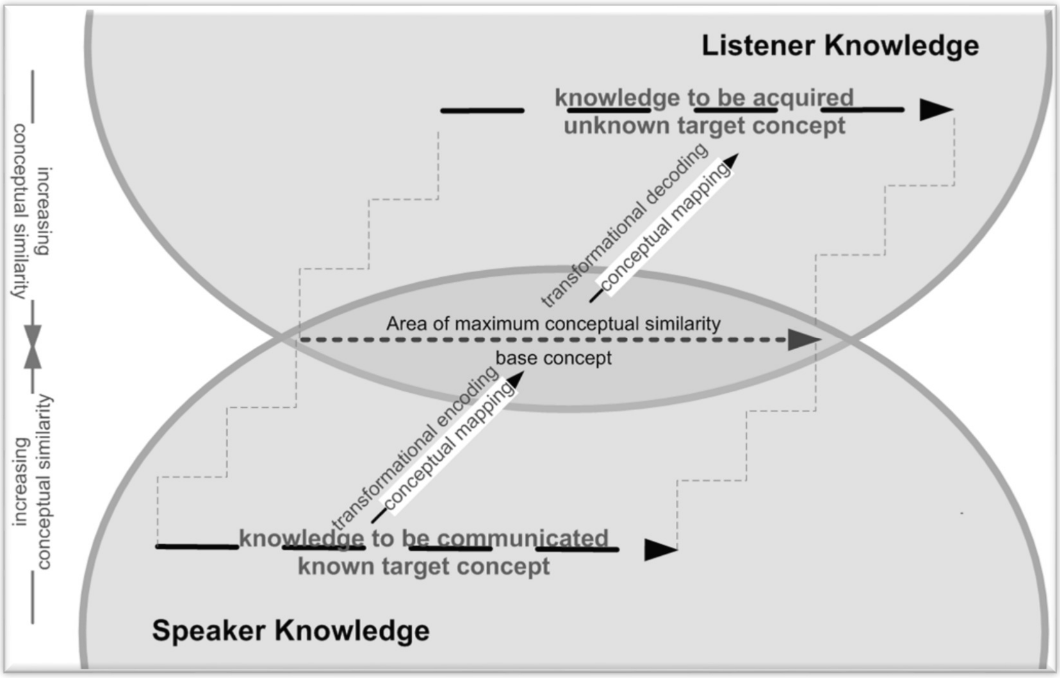


Figure 3-2. Refinement of the Digital Transformatives schema in human communication.

The main challenge of developing conceptual mapping instruments lies in determining the area of increased conceptual similarity. The area of increased conceptual similarity seems to be comparable to a shared basic level (compare *Cognitive Efficiency Catalysts in Communication* pp. 46).similarity and salience

3.2 Digital Transformative Concept - Cognitive Prototype Based System Development for Improved Efficiency in Human-Machine Communication

After consolidating investigations on mnemonic devices, metaphors, and categorization to a concept of cognitive efficiency drivers in human communication, the concept will be further refined towards a concept for human-machine communication. Therefore, the machine is seen as a communication partner, substituting the speaker of the previous defined schema in Figure 3-2.

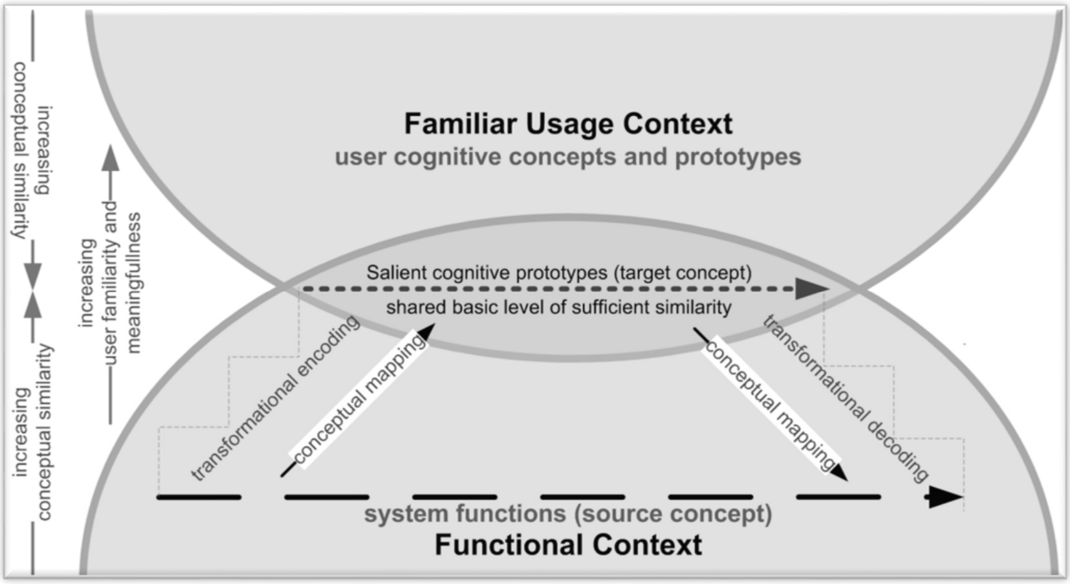


Figure 3-3. Concept for efficient human machine interfaces, derived from efficiency enhanced cognitive processes.

Consequently, a system could use the same efficient mechanisms for communication, as we know it from human-human communication. The user interface provides rich communicating channels and various appearances. Analogous to differing conceptual concepts of speaker and listeners in human communication, the conceptual system performance context often also varies from the users' conceptual performance contexts. As understanding is improved through information transformation onto proper concepts, performance should analogously be improved through adequate contextual mappings. Hence, the user interface should be implemented on a shared basic level, which corresponds to a conceptual user cognitive prototype context. As part of the digital system interface, output information should be encoded into this high performance user context, and user inputs need to be decoded back into the system context. Unlike, human-human communication the system interface can be completely adapted to the user conceptual context, during the design process, integrating the conceptual mapping for encoding and decoding information into the system logic, hidden from the user.

Feature 4. The user interface provides a bidirectional conceptual mapping between user context and system context through transitional encoding and decoding.

From investigations on human-human communication it can be derived that the shared basic levels lie on cognitive prototypes. New information is cognitively processed through similarity comparisons using cognitive prototypes as reference points. Multiple studies indicate that the frequency of features is relevant for prototype generation (Neumann, 1977; Posner & Keele, 1967; Reed, 1972; E. Rosch & Mervis, 1975). Prototypes are probabilistic clusters based on natural occurrences of patterns, which form around some kind of averages of a class of objects (Franks & Bransford, 1971; Neumann, 1977; Posner et al., 1967; Posner & Keele, 1967; Reed, 1972; Solso & McCarthy, 1981). Studies on the world color survey, or on password security, provide illustrative insights on this prototype formation, and underline this relationship of environmental occurrences and prototypical clusters (DataGenetics, 2013; P. Kay et al., 2009; P. Kay & Regier, 2003; Richard Cook et al., 2012). Prototype categories correlate with areas of improved cognitive efficiency (Heider, 1971a, 1971b, 1972; E. H. Rosch, 1973b; E. Rosch & Mervis, 1975; E. Rosch, 1975a, 1975c, 1978).

Feature 5. Prototype categories reflect probabilistic real world stimuli of high occurrence frequencies and improved cognitive performance.

As detailed in (*Performance, proceduralization, and category prototypes* pp. 34), the mechanisms of category prototyping may also be active in domains such as learning and automatization processes (also compare De Groot, 1978; Kirkham, Slemmer, & Johnson, 2002; LaBerge, 1975, 1976; Samuels et al., 1978; Sternberg, 2008; Vicente & De Groot, 1990).

Feature 6. The fundamental mechanisms of prototype categories are also active in process automatization through training.

Moreover, similarity comparison is a key mechanism for increasing efficiency of cognitive concepts mapping. Comparisons are directed from the more salient object serving as the base to the target object. The object more similar to a prototype is predominantly considered more salient. Salient base features are preferably mapped first. This is easily understood by reversing metaphors, which affects their interpretability (D. Gentner & Clement, 1988; Glucksberg & Keysar, 1990; Glucksberg et al., 1997; Ortony et al., 1985; Ortony, 1979; Tversky, 1977). In a next step the salient features of the base concept are transferred to the target concept. There is only little knowledge on how exactly this transfer happens. According to elaboration of knowledge in chunking theory the two concepts may in some way simply be interconnected (compare Performance, proceduralization, and category prototypes pp. 34 and *Pervasiveness of Similarity Comparisons* pp. 44).

Feature 7. Most salient or familiar features of a prototype are mapped first.

Feature 7 is supported by studies showing that people identify objects at a basic level more quickly than they identify objects at higher or lower levels (E. Rosch et al., 1976). Analogous children learn typical instances of categories earlier than they learn atypical ones (E. Rosch, 1978)

Feature 8. Conceptual context mapping is directed from the base to the target.

Feature 9. Objects similar to a prototype are predominantly considered more salient.

The epidemiological property extraction of context shifts of existing Digital Transformatives indicates that the user performance context lies in an area of increased familiarity and semantic elaboration. Other properties, as illustrated in Figure 3-4, were not necessarily found in all context shifts, hence may not be an essential user efficiency driver for digital systems.

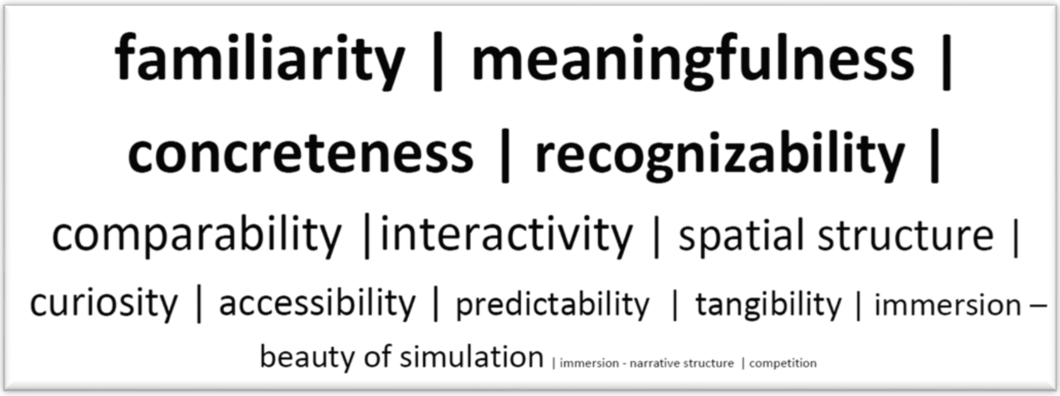


Figure 3-4. Context shift property tag cloud sorted by frequency. Familiarity and meaningfulness appeared to be common characteristics of all context shifts.

The importance of familiarity also has been heuristically emphasized in multiple best practice guidelines, most prominently in the Star user interface design guideline (D. C. Smith et al., 1990). Smith and others detail the main principles used during the development of the Star™ user interface. They state that any design decision should have been made in favor of providing something easy over hard, concrete over abstract, visible over invisible, copying over creating, choosing over filling in, recognizing over generating, editing over programming, or interactive over batch. Generally spoken, besides concrete over abstract one should always be aiming for familiar over unfamiliar. We see more things than we do not see, we copy or mimic more often than we create, we choose more often than we design, we recognize more than we generate, change things rather than create from scratch, we interact more than we plan, and because we have done hard things so often they became easy. In this context, one of their main design goals lies in pursuing familiar user's conceptual models.

Guideline 1. Concretized guidelines from the Star UI for pursuing higher familiarity: favor something easy over hard, concrete over abstract, visible over invisible, copying over creating, choosing over filling in, recognizing over generating, editing over programming, interactive over batch.

It has been argued in section 2.2 that familiarity may be a conceptual description for typicality of prototype categories and automatization processes.

Hypothesis 2. Familiarity corresponds to cognitive prototype categories and well-practiced processes which describe areas of increased performance.

Additionally, all analyzed systems also increase meaningfulness, which has been referred to semantic elaboration, in this work. We also know that conceptual mapping demands sufficient concept similarity to the source concept. The can be concluded to Feature 10:

Feature 10. Digital Transformative interfaces are situated in a context with maximum user familiarity, which corresponds to cognitive prototypes on a shared basic level of sufficient target similarity.

Many studies and research projects show improved usability through the use of metaphors in human computer interface design (Dix, 2004). Such studies also revealed negative effects of using metaphors. As already mentioned above, a metaphor can only be successful if its analogy is positive. A transfer of knowledge from one to another domain is only positive, when it enables users to actually apply it in the same way (Hesse, 1966). However, since two different domains by definition cannot be identically, there have to be negative analogies, too. Those negative analogies can easily be disturbing or distracting (Allwood & Eliasson, 1987, p. 170; S. A. Douglas

& Moran, 1983; Lansdale & Ormerod, 1994, p. 179,180; D. Norman, 1998, p. 180,181). Unmatchable features might lead to a misunderstanding or wrong interpretation of the target system. Negative analogies evoke misunderstandings, leading to improper use of the system (compare Preece et al. (1994), Gentner & Nielsen (1996) or Halasz & Moran (1982)).

Affordances might also be misleading and incorrectly interpreted by users (Gaver, 1991). The basic cognitive mechanisms of Affordances are also closely related to the concept of *the Law of the Instrument*.

Guideline 2. In combination with Feature 7: Be aware that negative analogies do not occur among the most salient features.

While mnemonic devices are particular techniques for improving memorization, DTs more generally, are meant to support existing tasks in the users' environment. Hence, the frame for each DT is set by certain actions necessary for achieving a given task. This frame provides the starting point for the development of every DT.

3.3 Digital Transformative Main Characteristics

In the last section it has been determined that efficiency increases in communication are related to cognitive prototype categories and semantic elaboration. Cognitive prototype categories are formed probabilistically based on frequency of real world stimuli. Very similar mechanisms can be observed in automatization processes. Our performance adapts to environmental requirements, in a way that we show highest performance at tasks or stimuli of high occurrence. Cognitive prototype categories are usually determined through typicality tests. Since typicality hardly is associated with automated processes, it has been argued that the overall principle corresponds to the idea of familiarity. The term "familiarity" meets a commonly understood concept, which may closely relate to cognitive prototype categories and processes in practice.

Apart from familiarity, the epidemiological property extraction also revealed semantic elaboration as a fundamental common characteristic of context shifts of Digital Transformatives. The practical analysis showed that it is hard to determine semantic or knowledge elaboration, even if we refer to the more colloquial term of "meaningfulness".

Sematic elaboration is often considered to be a working principle of mnemonic devices. It is not contradictory to the concept of familiarity. Familiar information is likely well connected to other concepts. However, elaborate knowledge often does not correspond to most efficient cognitive concepts.

The processing of information demands cognitive load. (F. I. M. Craik, Govoni, Naveh-Benjamin, & Anderson, 1996) asked probands to perform a time critical reactive visual task, and a memorization task simultaneously. In a control group the same tasks had been conducted sequentially. The tests revealed a significant performance decrease of the primary memorization task during the double activity. (Engle, Tuholski, Laughlin, & Conway, 1999) assumes that control of the focus of the double activity demands extra cognitive load. It can be assumed that the internal process of associating information demands extra processes for staying focused, which measurable increases the cognitive load (also compare (F. I. M. Craik et al., 1996; Iidaka, Anderson, Kapur, Cabez, & Craik, 2000; Oberauer et al., 2005; Thompson-Schill, D'Esposito, Aguirre, & Farah, 1997; Vincent, Craik, & Furedy, 1996).

Investigations made on cognitive load, connected with semantic elaboration, give evidence for increased efficiency at familiar actions and mindsets. Since familiar information should be better recognized and understood, the associative processing of this information is comparably less demanding. Consequently, semantic elaboration might mainly profit from familiarity. Moreover, the concept of knowledge elaboration seems to be fuzzier then the frequency based concept of familiarity. Therefore, familiarity can be seen as primary characteristic for Digital Transformatives.

Starting from an analysis of mnemonic devices a basic concept has been elaborated. The feature set of Digital Transformatives has been extended and new, more concrete hypothesis, have been added to the initial hypothesis. The current features and their dependencies, as well as the hypotheses are shown in Figure 3-5.

The assessment of Hypothesis 1 is essential for validating Feature 1. The evaluation of Hypothesis 1 first demanded an elaborate concept of Digital Transformatives, as it has been developed in the beginning of this chapter. Accordingly, Digital Transformative interfaces are situated in high performance usage contexts. System functions need to be encoded into such contexts, while user inputs are decoded back into function contexts. This major characteristic is expressed in Feature 4. Feature 8 and Feature 7 are defining sub-features of Feature 4.

The concept depends on high performance usage contexts, expressed in Feature 10. Hence, Feature 4 is dependent on Feature 10. According to cognitive research cognitive prototypes categories mark such areas of high human potentials. They are mainly formed through automatization, learning processes, and frequent occurring environmental stimuli (Feature 5, Feature 6). Salience is a major characteristic of cognitive Prototypes (Feature 9), and human beings living in similar environments tend to develop similar cognitive prototypes (Feature 2).

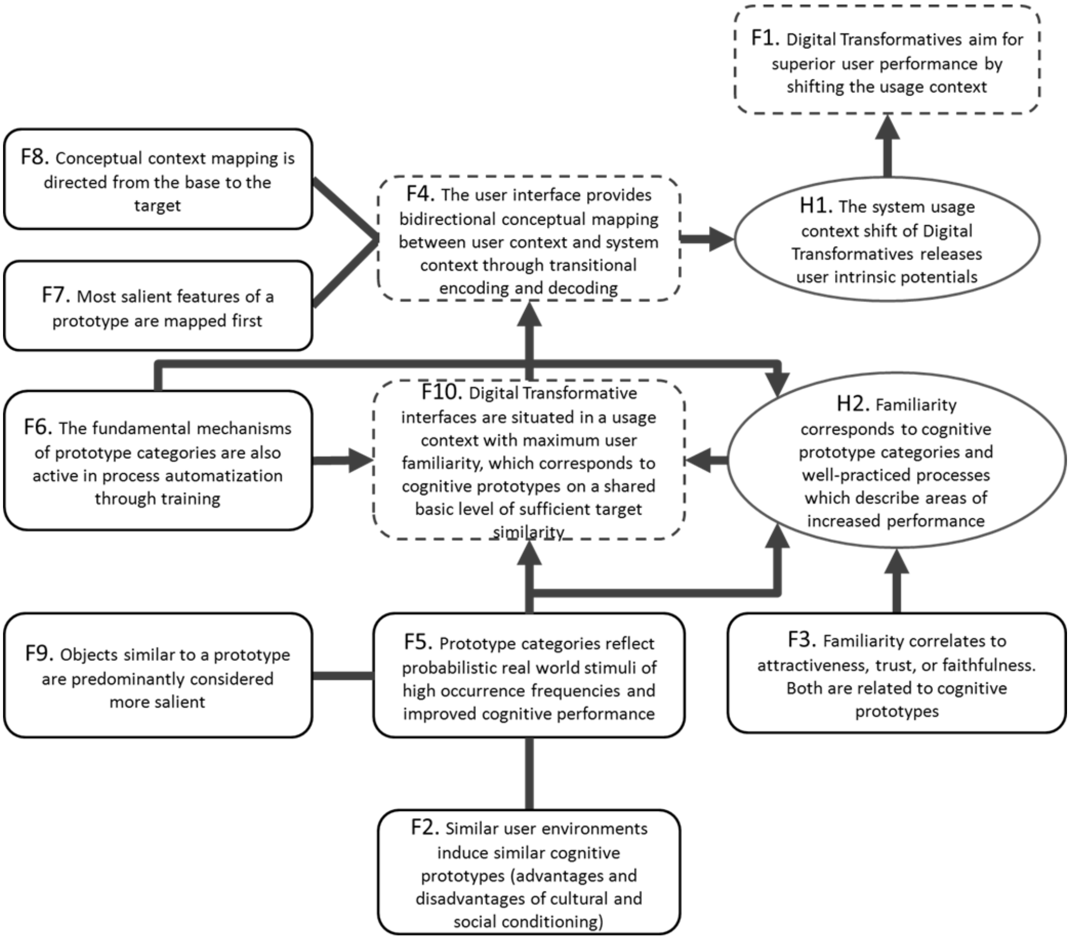


Figure 3-5. Initial features, hypotheses, and their evidential dependencies. Features with little evidence are marked with dotted lines. Dependencies are marked through arrows, and defining sub-features are connected through lines.

Moreover, the practical feature extraction, conducted in chapter 2.2, suggests the importance of familiarity, in relation to cognitive prototypes (Hypothesis 2). Although the previous section elaborated many indicators that familiarity is a proper superordinate concept for cognitive prototype categories and automatization processes, further validation is necessary to proof a correlation between familiarity and user performance. Hence Hypothesis 2 will be evaluated in the next chapter and further in chapter 4.2.1 subsection *User Rated Familiarity*.

3.4 Basic Concept Validation

This chapter aims at validating Hypothesis 2: *Familiarity corresponds to cognitive prototype categories and well-practiced processes which describe areas of increased performance*. In a first step, existing evidences in research will be described, the second part consists of a small scale prototypical test. This test is used as a pre-test to get a first indication.

3.4.1 Familiarity as a Driver for Human Efficiency

Familiarity can be found in many heuristics and guidelines on Human-Computer-Interface design. However, it is understood as a common colloquial characteristic, rather than being seen as a measurable key feature for designing a interfaces (Dix, 2004; Preim, 1999; Rogers et al., 2011; Shneiderman & Plaisant, 2010; D. C. Smith et al., 1990).

The possible correlation between familiarity and efficiency in human cognition has also been elaborated in the previous chapter. Cognitive prototypes and proceduralization may be connected to the process of familiarization. Increasing performance through practice, as it is described with the power law of practice, correlates with increased familiarity of certain procedures. The frequency of stimuli exposure also relates to cognitive recognition, and the processing performance of such patterns. Other studies give further evidence for the importance of familiarity as a performance driver.

FAMILIARITY RELATES TO PRODUCTIVITY IN WORK ENVIRONMENTS
(P. S. Goodman & Garber, 1988) showed that absenteeism had an impact on the rate of accidents. Data gathered from production crews in five underground coal mines was studied. The study focused on the individual workers and led to the assumption that absenteeism raises the probability of accidents, due to increased unfamiliarity, caused by short-term changes of the work environment. Based on this work (P. S. Goodman & Leyden, 1991) investigated 26 coal mining crews in two coal mines to study the effects of familiarity on group productivity. They created a model for determining a measurable change of familiarity when crew members temporarily absent. Short-term changes in crew constellations affected the members' knowledge about their co-workers, specific jobs, and work environment configurations. New workers were unfamiliar with the unique properties of the machinery, physical environment, job, and work habits of the original crew. In turn, original crew members were unfamiliar with the work habits of the new replacement worker. This way, (P. S. Goodman & Leyden, 1991) conclude that familiarity refers to the level of

knowledge a person has about co-workers and work activities. In their model familiarity is measured based on the number of shifts workers stay in a similar situation. Changes in job, crew, or section, lower the level of familiarity. (P. S. Goodman & Leyden, 1991) determined the altering level of familiarity over several shifts for each worker. From the average of its members a crew familiarity value was calculated and then related to the crew productivity. The productivity was measured in tons of coal mined per crew. Hereby, the study revealed that a decline of familiarity is also associated with a decline in productivity.

In another experiment conducted by (Gruenfeld, Mannix, Williams, & Neale, 1996) groups of three persons had to solve a murder mystery. Each group member received a set of interviews, a map, a list of characters, handwritten notes, and a newspaper article. The information distribution within the group was slightly altering, leading to several different perspectives. In order to solve the mystery, all perspectives had to be discussed and brought to a shared solution. The groups were constituted either of three strangers, two familiar member, or three familiar persons. In general the experiment showed that groups whose members were all familiar also were more effective at pooling knowledge, and integrating alternative perspectives. However, it also turned out that those groups of familiar members consolidated more harmonically, leading to a less diverse final perspective. Hence, they were less likely experiencing conceptual conflicts among the differing perspectives than groups of strangers (Gruenfeld et al., 1996).

Obviously, familiarity among group members is not the only factor affecting group performance. Group composition of diverse expertise, cultural diversities, social interaction, and other factors influence group performance (Espinosa et al., 2001; Hinds, Carley, Krackhardt, & Wholey, 2000).

Other studies showed that increasing familiarity with a certain working domain acquired through seniority, also correlates with higher performance (Banker, Datar, & Kemerer, 1987; Gordon & Fitzgibbons, 1982; R. Katz, 1982).

3.4.2 Memory - Basic Concept Test Case

For further evidence on the relation between familiarity and performance, it has been decided to perform a recognition and memorization test based on faces. Faces are good test cases because persons are common objects of interest. Moreover, they provide a continuous spectrum of samples covering various levels of familiarity, from relatives, which should be familiar, over friends, celebrities, to complete strangers. Additionally, those results can easily compared to existing prototype tests on face recognition and existing Digital Transformatives based on face recognition, such as the PassFace system described in section 2.2.3 (Davis et al., 2004; D. T. Levin, 2000; T. Luce, 1974; Malpass, 1992; Solso & McCarthy, 1981; Valentine & Endo, 1992).

As a first test environment the application *Memorize* [15] of the Sugar® platform (sugar labs, 2010) has been used. Sugar was originally developed for the XO-1 laptop, known from the *One Laptop per Child* initiative. Memorize is a computer implementation of the card game called *Concentration* (Glonnegger, 1988). The traditional game consists of a set of cards of matching pairs. In the beginning all cards are shuffled and laid on a surface with the faces flipped downwards. The game is played round-robin with a desired amount of players. At each turn the current player chooses two cards and flips them face up. If a matching pair is found the player keeps those cards, otherwise the two flipped cards are turned again.

For this test two sets of cards with comparable content were set up. One set consisted of familiar content, while the other set had to be unfamiliar. Since all persons of the test group knew each other, the familiar set consisted of faces of the group members. The unfamiliar set showed strangers, randomly downloaded from the internet. Test persons had to play both sets. Their performance was compared.

Additionally, two audio sets, with each holding nine pairs of voice recordings, have been prepared as well. All participants were provided an audio recording, introducing themselves by saying “Hello here is”, followed by their name. Those audio snippets have been used for the first audio set of the memory cards. As a second, more unfamiliar variant, all participants are introduced by a single speaker. The speaker said “Hello here is” followed by the name of the participant. Instead of showing the face, a card flip triggers a one-time playback of the according recorded audio.

However, incidentally all familiar photos or sounds could be more remarkable than the unfamiliar ones. Hence, in order to minimize remarkability influences apart from familiarity, the exact same tests were performed by a control group. The control group followed the same procedure, except it was unfamiliar with all presented persons. This way, influences on performance unrelated to familiarity, such as image composition, contrast, or other remarkable features, could be extracted.

The sets have been tested with 8 test group and 11 control group participants. The ages of the participants were distributed over ages of 4 to 65, as shown in Figure 3-6. The test group was exactly half and half, female and male. In the control group 6 out of 11 participants were female.

[15] Memorize 35: http://activities.sugarlabs.org/en-US/sugar/addon/4063 (Version 35)

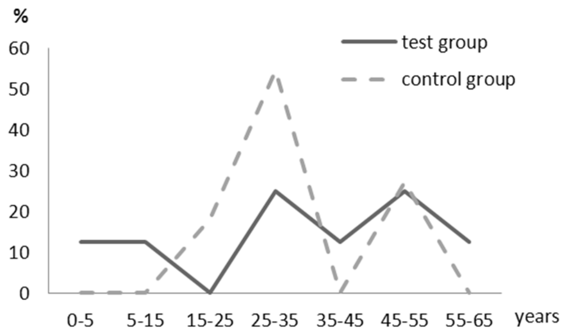


Figure 3-6. Age distribution of the test group and control group.

<u>MEMORY - SETUP, APPARATUS, AND TEST PROCEDURE</u>

The test was conducted casually at home, outside of a laboratory environment. A computer was set up in a separate room, and participants were asked to play the game one after the other. The participants were informed about the test situation. The tester sat next to them, clearly counting the moves needed. Participants were provided with their result right after each run. It was open to the participants to compare themselves to the others. It was made explicit that there was neither time pressure nor that results were made public, or compared to others. The rules of the game were explained to every participant before they started to play. They were also invited to try out the usage of the system at a sample game with different content, prior the actual test. The game interface is shown in Figure 3-7.

Figure 3-7. The game interface of the memory test environment. The screenshots display the interface for a visual set of cards (left), and an audio set (right).

The display is divided into three areas. The menu bar at the top, the score board at the left side and play field at the right. In the beginning of each test, all cards are flipped face down, a solid grey side facing up. A left mouse click on the grey backside of the card flips the card, displaying the content of this card. If two cards are flipped and they are matching, they stay in position faced up, marked with a yellow border line. Otherwise they are automatically flipped after a certain time, or at the next click.

Per test run, each participant had to play all four sets. All four sets were played one after the other without any break, in random order, alternating between aural and visual. The first eight test runs were conducted on one day. It was tested on a free schedule. Test persons were free to play whenever they desired. Three participants performed one additional test run the other day. The control group was tested using the same procedure. Tests with the control group were conducted on a different day, also in a leisure context.

MEMORY - EVALUATION

In total a sample rate of 3080 moves were recorded from both groups. Most target test persons really enjoyed the game. 6 of them wanted to play again. 3 actually played again the next day.

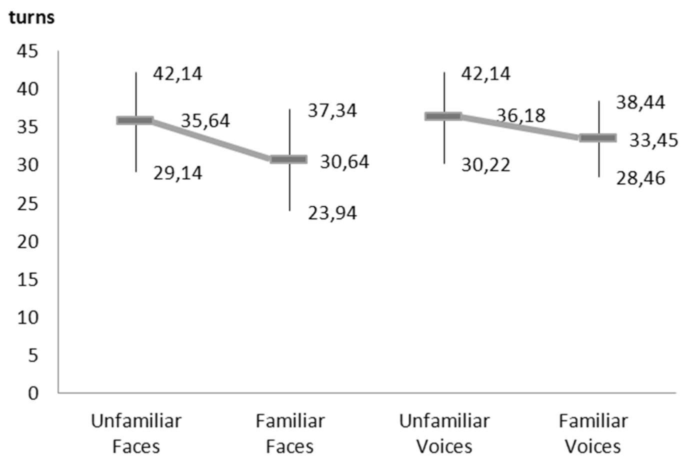


Figure 3-8. Average turns needed by the participants of the test group, who should be sensible to familiarity differences.

The performance was measured according the turns needed to find all matching pairs of a set. In total 22 test runs have been conducted by 8 participants.

In the aural test, an average of 36,18 (SD 5,96) turns were necessary for all participants to find all matching pairs of unfamiliar voices. In comparison to that 33,45 (SD 4,99) moves were necessary to reveal all pairs of familiar voices, which corresponds to an improvement of nearly 8,55 %. The standard deviation for unfamiliar voices was 5,96 and for familiar voices 4,99. Regardless, the tendency supports the proposed concept of improved performance through familiarity.

The aural results are affirmed by the faces test. Unfamiliar faces demanded 35,64 (SD 6,5) turns in average, compared to 30,64 (SD 6,7) turns for familiar faces. This corresponds to an improvement of approximately 14 percent. The standard deviation for unfamiliar faces lies at 6,5, and at nearly 6,7 for familiar faces, which corresponds to 18 percent of the standard deviation for unfamiliar faces. Figure 3-9 shows the results of the 22 control group test runs for visual and aural samples.

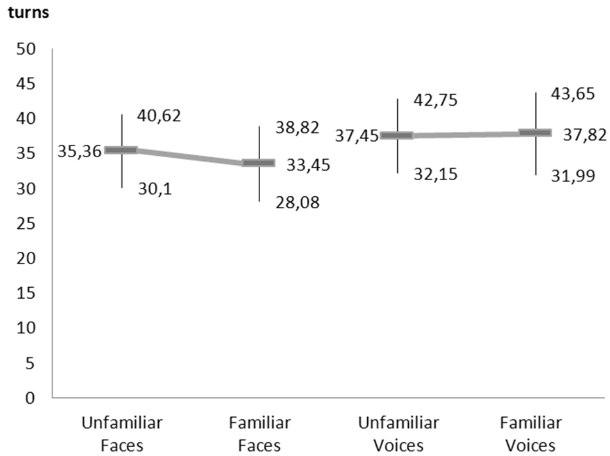


Figure 3-9. Average turns needed by the participants of the control group. All sets were unfamiliar to those members.

The visual performance of the control group was similar to the performance of the target group. The control group performed 5,4% better on faces familiar for the target group, although there was no familiarity advantage on such visual samples. The 14% performance increase observed at the target group still indicates an advantage of about 8,6%, which can be accounted to familiarity. While the visual samples seemed to hold more variables influencing recognition and memorization, the aural memorization tests of the control group showed less external influences. The control group performance was almost identical for all voice samples, while the target group performed approximately 7,5% better on the recognition and memorization of familiar voices.

In total the test group performed remarkably better on familiar test samples, than the control group, as shown in Figure 3-10.

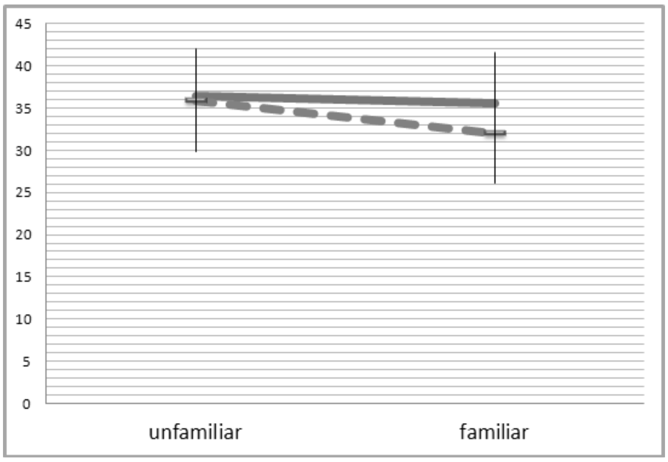


Figure 3-10. The results of the familiar test group (dashed line) and the control group (solid line) distinguished between familiar and unfamiliar.

Those descriptive statistics are underlined by inferential statistics. Therefore, a paired-samples t-test was conducted to check, whether the difference between the familiar and unfamiliar results of the test group and the control group are statistically significant. If familiarity has no influence a reliable difference should be determined for both groups. For the test group the difference was statistically reliable with $T(21)=2.99$, $p=0.007$, while for the control group the difference was clearly unreliable with $T(21)=0.63$, $p=0.53$. This clearly supports the influence of familiarity in the above task.

The test runs revealed two major concerns regarding the game performance. On the one hand, a user might be lucky finding a pair on the first flip, and hereby reducing the overall complexity. On the other hand, two users stated that they were rather looking for remarkable features than for familiarity. The influence of luck can be neglected over a certain number of repetitions. It could also be integrated into the game logic to reduce the factor of having luck. The influence of remarkable features might be reduced by selecting content which does not differ much in its remarkability, as well as by offering more test sets. Currently two different sets are provided.

The result of the test including voices is even more interesting. One could assume that the control group performed comparatively better on the voice memory, because the audio content would not be as feature rich as the visual content. Thus, the audio snippets were harder to differentiate, which makes them harder to remember. (Avons, 1999) conducted an empirical study presenting test persons different check patterns. The probands had to remember a sequence of such patterns in the right order. The performance was worse when patterns were more similar to each other. On the other hand, the double task of memorizing sound and content could have been a relevant factor. This double task could have led to an increased cognitive load, due to the additional challenge of focusing and assessing information importance. (compare F. I. M. Craik et al., 1996; Engle et al., 1999)

CONCLUSIONS
This test gives further indication on the validity of the basic concept, since it showed improved performance at familiar items.

Other psychological studies on face recognition are in line with the results of this test. Investigations of the PassFace mechanism showed clear tendencies for race, gender, and attractiveness (Davis et al., 2004). Individuals tend to be more efficient in recognizing and memorizing faces of people of their own race. This so called own-race effect is accounted to the increased exposure of members of the own racial group (D. T. Levin, 2000; T. Luce, 1974, 1974; Malpass, 1992; Valentine & Endo, 1992; Walker, Tanaka, & others, 2003).

There seem to be clear similarities between familiar items and cognitive prototype categories. According to that, our understanding of familiarity of faces corresponds to probabilistic clusters formed based on environmental stimuli. The familiarity clusters also seem to serve as reference points for recognition and memorization, located at areas of high performance.

Psychological studies also show that the sense of attractiveness is quite common. We tend to find average faces attractive (Langlois & Roggman, 1990; Langlois et al., 2000; Rhodes, 2006).

The test results support Feature 5 (*Prototype categories reflect probabilistic real world stimuli of high occurrence frequencies and improved cognitive performance*) and especially give positive validation for Hypothesis 2 (*Familiarity corresponds to cognitive prototype categories and well-practiced processes which describe areas of increased performance*).

3.4.3 Correlation between Familiarity and Performance in Team Sports

While the previous recognition and memorization test addressed analogies between familiarity and cognitive prototypes, the following test aims at more complex procedural tasks, which involve automatization through practice. Hereby, the studies on group performance described in *Familiarity as a Driver for Human Efficiency* (pp. 115), provide a good conceptual ground for a comprehensive performance evaluation based on team sports. Professional team sports are usually based on competitive assessment of group performances of complex physical, psychological, and social interactions. Many archives are publicly available, substantially documenting professional performances. Hence, such databases provide an interesting ground for investigating performances in groups. The findings of (Banker et al., 1987; P. S. Goodman & Garber, 1988; R. Katz, 1982), and associated studies on proceduralization and expertise, as detailed in chapter *From controlled to automated processes to habituation* (pp. 38), indicate performance gains in work environments through familiarity, and should be visible here as well.

MEASURES FOR PERFORMANCE AND FAMILIARITY

As one of the most popular sports of the last decades, football offers a rich database. Football offers various statistics on expressing and measuring team performance. Team sports, in general, offer a highly complex case with multiple factors affecting the final performance. On the one hand, in professionally played sports, like football, the differences between players are considerably low, so that psychological differences are more relevant and visible. The referee could also be an important factor. The fans are playing a big role. The strategy and individual players have to fit. Although, goals and points might not be the perfect measure for performance, they are the measure the game is optimized on, as part of the game rules, so they actually provide a perfect measure of choice. Although a team might be most elegant or advanced, as long as this does not result in a high number of points and goals, the team is not performing well, in consideration of the rules provided. Hence, the crucial measure for football is the scoring system based on points and goals (Deutsche Fußball Liga GmbH, 2012; Gesellschaft für DFB-Online mbH, 2012).

The measure for familiarity is more imprecise. A first approximation could be achieved by simply comparing the names in the starting lineup of consecutive matches, and count the players fluctuation. Maximum familiarity can be assumed from a team that always plays with the same players. Minimal familiarity may be indicated by a team playing every game with completely new players. However, among those extremes this method has many complex cases, which are hardly comparable. First, it is not differentiated between the familiarity of the substituted players. For example, if a core of 10 players plays all the time, and one player is substituted every alternating game, then the familiarity measure of such a team would be identical to a team, where the substituted players are completely new at every match. While one team would have 10 completely new players in 10 matches, the other team would only switch between two substitutes and consist of 12 players. The familiarity of the second team should be different to the one of the first 20 players. Additionally, and most importantly the consecutive match based analysis would not consider many days of practice in-between the matches.

A better approximation for familiarity should be provided by the numbers of players used in a season. Taking the general team fluctuation as a measure leads to the assumption that a team with a smaller amount of players, used in a season, should have a higher internal familiarity than a team with more players. It can also be assumed that a team with higher fluctuation in their starting lineup also has higher fluctuation in general – including the training session.

TEAM SPORT ANALYTICAL PROCEDURE AND RESULTS

As explicit above the analysis was based on archived data of the German Fussball Bundesliga. The Fussball Bundesliga is a professional German football league (Deutsche Fußball Liga GmbH, 2012; Gesellschaft für DFB-Online mbH, 2012). For familiarity and performance assessment the starting lineup of each team and according match results were demanded. Since the desired data was not available in one package, a html crawler has been implemented in Smalltalk, automatically requesting and collecting demanded data from the websites of (Fussballdaten Verlags GmbH, 2012). Data from 13752 games of the first, and 15276 of the second German Fussball Bundesliga has been collected and investigated, making a total sample rate of 29028. The games of the *1. Bundesliga* included information of 51 teams and roughly 44 seasons in the time from 1966 to 2010. 127 teams of the *2. Bundesliga* participated during the years of 1974 to 2010.

The Bundesliga score system has changed over years, and it differs among competitions. Therefore, three measures have been calculated. The first measure is defined by the won games of a team. The second measure is based on the old scoring system, rating a tie with one point and a win with two points. The third measure is given by the new rating system, where the winning team gets one point for a tie and three points for a win. All three measures resulted in analogous curves for almost corresponding ratings of team performance.

Comparing the dashed line corresponding to un-familiarity (players in use per game) to the performance line (points per game) indicates a clear correlation between familiarity and and performance. The worst teams of the *1. Bundesliga* used four times more players, while the best teams were approximately 22% more successful. Compare Figure 3-11

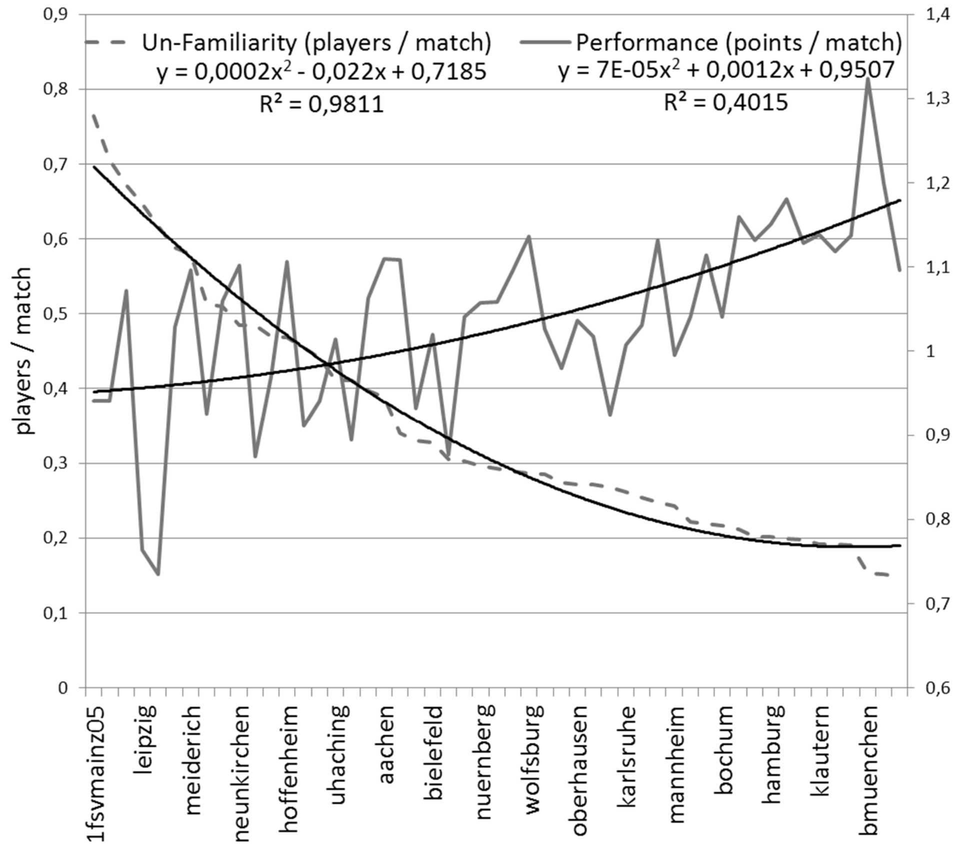


Figure 3-11. Correlation between performance and un-familiarity of teams of the 1. Bundesliga (not all shown teams are labeled).

In the *2. Bundesliga* good teams integrated a new player approximately every 5th game while the worst teams integrated a new player nearly every game. This comes with a performance improvement of 15 percent. The curves also clearly indicate higher variance for unfamiliar teams, while the success curves are more stable for familiar teams (see Figure 3-12).

If team fluctuation is linked to team familiarity, and success to performance, as it has been argued above, then a decreasing number of team fluctuations can be associated with increased success. And it follows that increased familiarity leads to better performance. The investigations support the phrase “Never change a team to win!”, or “Never change a winning team!”. However, the results have to be taken with care. The amount of used players includes some imprecision in reflecting team familiarity. Unfortunately, there are no control measures extracting such disturbances, as it was done in the previous memory test. One could argue that successful teams do not have the need to change the team very often, while bad teams try to improve through change. At the same time, physical, psychological, and social factors are inevitable, and continuously drive fluctuation in the team. Players get injured, older, have social problems with team mates, or other problems influencing their fitness and performance. All those influences lead to continues team changes.

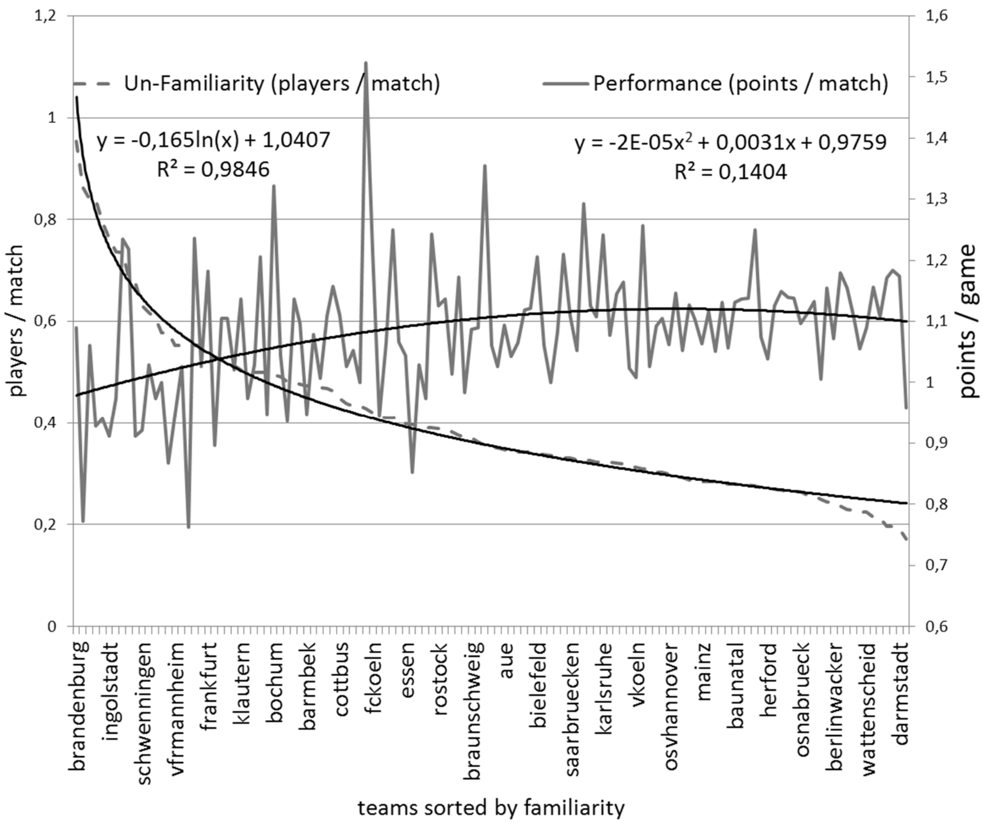


Figure 3-12. Correlation between performance and un-familiarity of teams of the 2. Bundesliga (not all shown teams are labeled).

<u>CONCLUSION</u>

The findings on proceduralization and expertise provide a cognitive basis for productivity gains in work environments through familiarity. Such findings are supported by the previously conducted team sports analysis. Practice of physiological procedures, and interaction with objects increases familiarity, and forms procedural clusters of high performance (compare to *Affordances from UI Concepts* pp. 99). This analysis is in line with studies on productivity in work, providing empirical evidence that the same patterns are valid for group interaction (Banker et al., 1987; P. S. Goodman & Garber, 1988; P. S. Goodman & Leyden, 1991; Gordon & Fitzgibbons, 1982; R. Katz, 1982). Familiarity with group members correlates to efficiency analogous to cognitive prototype categories.

As a regression analysis shows, the results also indicate a linear regression of performance with logarithmic growth of un-familiarity. The stability index for first league linear regression equals R^2=0,9071, and for the second league R^2=0,9205. The corresponding stability index for the logarithmic growth of un-familiarity is around R^2=0,98 for both leagues. Furthermore, with a stability index of R^2>0,9 the relation of familiarity and performance can be described with a potential progression, analogous to the formula for the law of practice (further described in section *From controlled to automated processes to habituation* pp. 38). Hence, familiarity seems to correspond to practice in automatization processes. This relation is visualized in Figure 3-13.

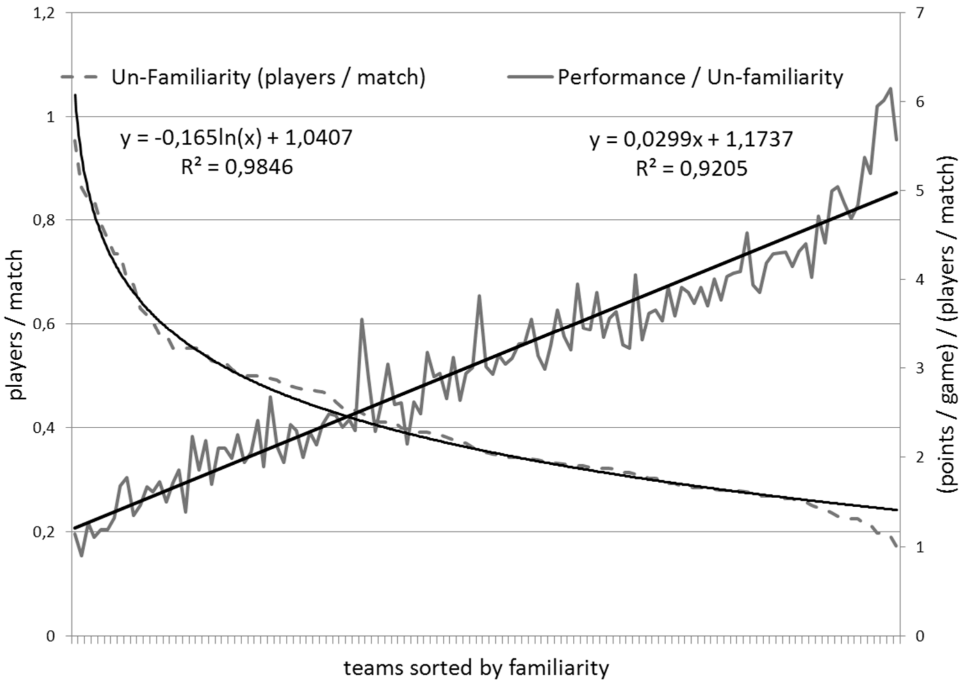


Figure 3-13. Unfamiliarity regresses logarithmic reciprocal to a linear progression of the relation between performance and un-familiarity, with a reliability of over 90%.

The results offer support for Feature 6 (*The fundamental mechanisms of prototype categories are also active in process automatization through training*) and give further positive validation for Hypothesis 2 (*Familiarity corresponds to cognitive prototype categories and well-practiced processes which describe areas of increased performance*).

3.4.4 Conclusions of the Basic Concept Validation

The previous sections gave further empirical evidence on Feature 5 (*Prototype categories reflect probabilistic real world stimuli of high occurrence frequencies and improved cognitive performance*) and Feature 6 (*The fundamental mechanisms of prototype categories are also active in process automatization through training*). Together with cognitive evidences expressed in Feature 3, those features refine validity of Hypothesis 2 (*Familiarity corresponds to cognitive prototype categories and well-practiced processes which describe areas of increased performance*). Hence, Hypothesis 2 is transformed into Feature 11:

Feature 11. Familiarity corresponds to cognitive prototype categories and well-practiced processes which describe areas of increased performance.

Feature 11 offers further validation for Feature 10 (*Digital Transformative interfaces are situated in a context with maximum user familiarity, which corresponds to cognitive prototypes on a shared basic level of sufficient target similarity*), which is also supported by cognitive evidences expressed in Feature 6, Feature 5, Feature 9, Feature 2. Thus, Feature 10 is based on a comprehensive empirical ground, further validating Feature 4, which provides the basis for Hypothesis 1. Consequently, Hypothesis 1 becomes Feature 12:

Feature 12. User context shifts of Digital Transformatives increase user performance and efficiency.

This provides us with a more refined feature graph (compared to Figure 3-5 on page 114) shown in Figure 3-14.

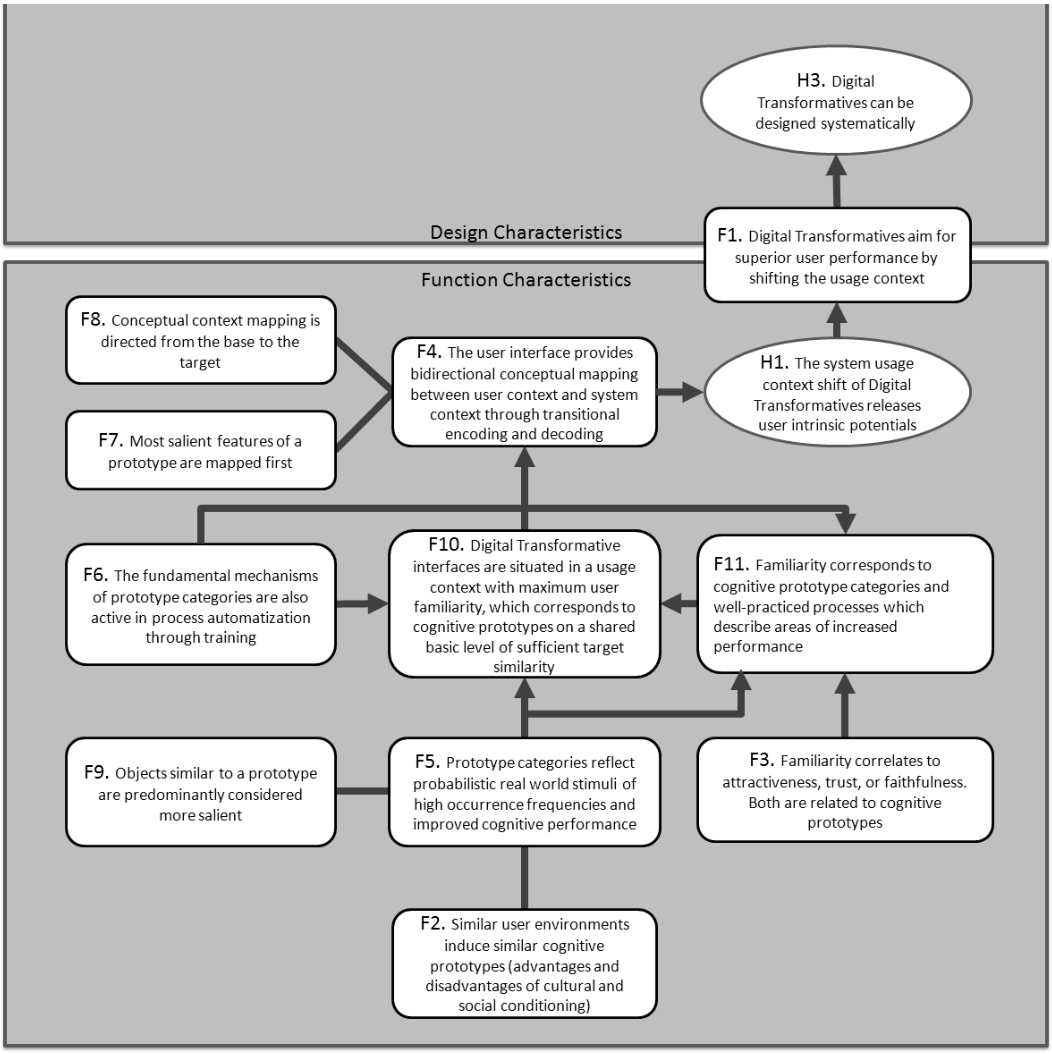


Figure 3-14. Digital Transformative Function characteristics feature graph.

After the basic concept of Digital Transformatives, and the principle of operation has been elaborated in the previous part of this work, the following part focuses on a determining a systematic design methodology for Digital Transformatives. The investigations are driven by the question, whether there is a systematic design methodology for creating Digital Transformatives.

Hypothesis 3: Digital Transformatives can be designed systematically.

4 DESIGNING DIGITAL TRANSFORMATIVES 131

4.1 DESIGN METHODOLOGY OF INTERACTIVE SYSTEMS 131
4.1.1 Interactive Systems Development 133
4.1.2 Interactive Systems Design Phase 135

4.2 INTERACTIVE SYSTEM DESIGN METHODOLOGY FOR DIGITAL TRANSFORMATIVES 152
4.2.1 Methods for Assessing Familiarity 154

4.3 SYSTEMATIC METHODOLOGY FOR FINDING A DIGITAL TRANSFORMATIVE CONTEXT 171
4.3.1 Prerequisites 172
4.3.2 DT Design Challenge 172
4.3.3 Salient Super Prototype Identification Approach 175
4.3.4 Sub-Action Modeling Approach 183

4.4 SUMMARY - OVERALL DIGITAL TRANSFORMATIVES DESIGN PROCESS 186

4 Designing Digital Transformatives

This chapter builds on Hypothesis 3 (*Digital Transformatives can be designed systematically.*). It will start with an analysis of general interactive system design. This analysis aims at finding anchor points that need modification in order to foster systematic design of Digital Transformatives. After those elements of the design procedure are determined, DT specific assessment and concept design methods will be elaborated.

4.1 Design Methodology of Interactive Systems

User centered interface development has taken an evolutionary path towards incremental iterative software design, where todays systems are often implemented as a steady optimization of an initial basic system.

In the context of quality assurance and performance optimization, Deming (2000) compared the "old way" of quality assurance of product suppliers, before the industrialization, to the "new way" (Deming, 2000; compare Shewhart, 1939). In the old days, product suppliers, such as, tailors, blacksmiths, shoemakers, or milkmen, knew their costumers personally. They knew the costumers needs first hand. Products or services were designed, produced, offered, and, based on direct feedback, adapted until costumers were satisfied. After industrialization, structures changed to serve mass markets. Wholesaler, jobbers, and retailers became an important part of sales processes, with an increasing risk of losing the personal touch. The former cycle was easily broken. Products were designed, created, and offered without knowledge about consumers. The conservative iterative individualized product development cycle (Figure 4-1 top) could be replaced by a straight process (Figure 4-1 bottom).

In the beginning, the disadvantages of limited knowledge about the target group might have been compensated by mass production. With increasing competition products demanded higher quality to be successful. A product designed for a specific consumer does not automatically address a whole group of different consumers. The former cycle had to be closed again. A fourth step was introduced, based on consumer research. The fourth step sought to test the product in service, find out what users thinks of it, and why others have not bought it (Deming, 2000).

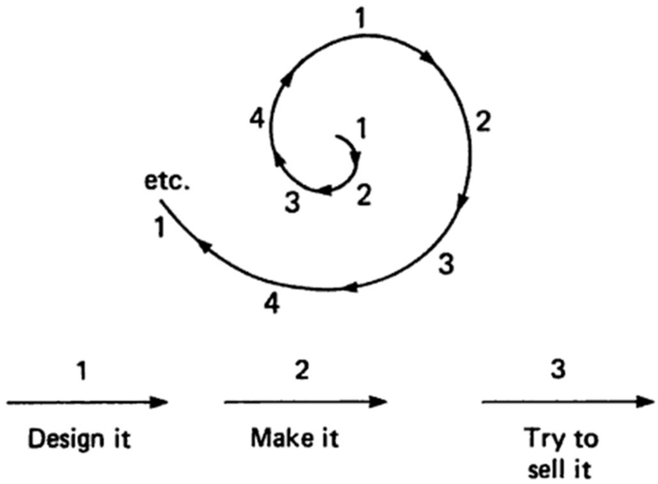


Figure 4-1. Iterative individualized product development and quality assurance before industrialization (top). Product development right after industrialization started (bottom). Back to four step assurance (top), after three step approach proofed to fail.

Software design has gone a similar way. Software mass production began with the advent of personal computers. First personal computer software systems designs were based on a hardware specification, followed by a functional specification, which was then used to create a logical user interface and command structure (Seybold, 1981). This common practice, however, started to change in the early 1980s, especially with the design of the *Star* desktop metaphor user interface. With the goal of building the "*ultimate professional workstation*" (Harslem & Nelson, 1982, p. 377) the development of Star began in 1977, while the actual software implementation started in 1978. Those first two years were spent with specifying the system, and building iterative prototypes, before the first line of product code was written. (D. C. Smith et al., 1990) further highlights the importance of defining users' conceptual models, during the initial design phase, rather than starting off with writing functionality, and putting a user interface on top of it (also compare Seybold, 1981). Since then user centered design became increasingly important for creating successful software.

During the years of 1981 and 1982, which could be considered to be an early stage in computer software design, Gould & Lewis (1985) wanted to find out about the commonness of the following three principles in computer system design: early focus on users and tasks, empirical measurement, and iterative design. They interviewed 447 people, consisting of designers, programmers, and developers. Each of the participants should name five major steps in the development and evaluation phase of a new computer system for end-users. Every fifth of the interviewed named iterative design, 40 percent listed empirical measurement, and 62 percent saw early focus on users as a major step in the design process. 26 percent did not mention any of the principles.

It can further be distinguished between human computer interface design, and classical development of non-interactive software. Classical software development follows a clear sequence of steps, with none, or just a few iterations. Non-interactive systems are closed; component behavior can be defined and determined in the conceptual phase. Interactive systems are not closed, in a sense that users can be seen as part of the system. User behavior cannot be predicted by the system designer. Thus, the development of human computer interfaces demands an iterative approach, which is described under the term of Usability Engineering (Jakob Nielsen, 1989; Preim, 1999; Rauterberg, Spinas, Strohm, Ulich, & Waeber, 1994).

4.1.1 Interactive Systems Development

(Gould & Lewis, 1985) proposed three major principles for designing computer systems for people. The first is described as *Early Focus on Users* (also compare Dix, 2004; Preim, 1999; Rogers et al., 2011). They propose the identification of potential users to get designers in direct contact with them already prior system design phase. In this phase, designers should acquire an understanding of the user through interviews and discussions. Users might even train designers their existing procedural patterns. This way, designers are able to analyze both, the challenges of the users, and their current environment. They further propose *participatory design*. Hereby potential users should become part of the design team. As a second major principle they see *empirical measurement*. They suggest to study learnability and usability of the developed system, already early in the development process. Users should be analyzed, performing certain tasks on prototypical systems, while performance, thoughts, and attitudes are recorded. As the last major principle they see *iterative design*. Starting from a basic system, several iterations of prototyping, behaviorally evaluating, and implementations should be conducted to improve the system.

(Gould & Lewis, 1985) further refine their principles into an initial design phase, followed by an iterative design phase. The initial design phase consists of a preliminary specification of the user interface, where the designers also collects critical information about users. Surveys should be conducted, or consultants interviewed, to get a general picture of potential users. Further, a more distinct picture may be gotten from direct user consultation. Additionally, *Behavioral Goals* should be defined. Behavioral Goals, for example, describe how many users perform a certain task in a certain amount of time. They provide measures to assure proper proceeding of system development. There should be a clear evaluation procedure for each behavioral goal. According to (Gould & Lewis, 1985) behavioral goals should at least consist of a description of possible users, a list of tasks to be performed, the environment they are performed in, and the measures of interest, such as learning time or errors. Also planning the organization of work initially, simplifies cooperation of multiple designers and developers in the later phases to the system development. In conclusion, (Gould & Lewis, 1985) emphasis the importance of the initial design phase, which establish behavioral goals and user access, and hereby pave the way for the continuous evaluation and modification of the interface, and following iterative steps.

(A. Taylor, 2000) asked 38 project managers about success criteria of IT projects. 1027 projects were covered consisting of development, maintenance, or data conversion projects. Half of the projects were development projects. As a result, managers clearly determined requirements definition as the most crucial project success criterion. A clear and detailed project plan was identified to be of minor importance for a successful project. This result is not very surprising, in consideration of most projects being iterative. Progress of a user centered iterative project is hard to foresee, making it necessary to modify project plans frequently.

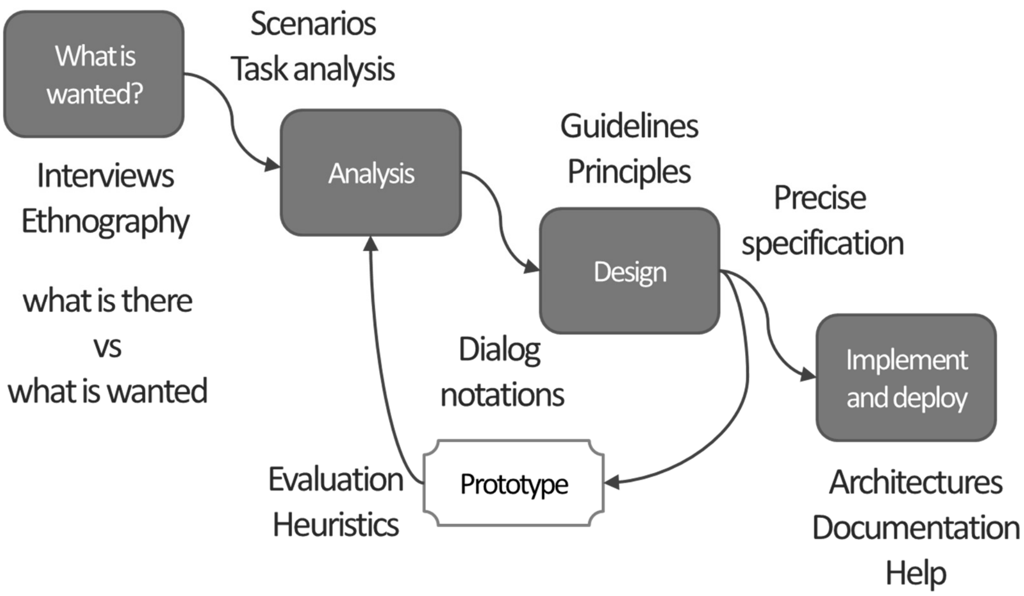


Figure 4-2. Interaction design process as a water fall model including an iteration cycle (from Dix, 2004).

The principle of iterative design can be seen as a fundamental concept in human machine system development. Its basic schema has been re-modeled, in slightly varying ways, multiple times already (Dix, 2004; International Standards Organisation, 1999; Preim, 1999; Rogers et al., 2011; Sommerville, 2001). Most of them include the four main phases, already described by Shewhart's product design and quality assurance cycle (Shewhart, 1939).

(Dix, 2004) illustrated the main phases of interaction design processes in a water fall model, including an iterative loop, as show in Figure 4-2 (compare Royce, 1970). A design process usually starts with a study of requirements. This phase should determine where we come from, and where we need to go. A follow-up analysis details the initial requirements analysis to determine key issues. It is also the starting point of the iterative phase, and serves as a basis for the subsequent design phase. The analysis reveals what needs to be done, while the design phase deals with the question on how to do it. Due to the complexity and unpredictability of human behavior the design concept has to be designed with the use of real users by implementing it prototypically. The prototype is re-evaluated, possibly leading to a re-conceptualization and further iterations. Finally, the system can be transferred into a production system.

4.1.2 Interactive Systems Design Phase

As Digital Transformatives aim at system design, typical steps of software engineering, such as architectural specification or unit tests, will not further be discussed here. In the following, all major stages affecting the initial design phase will be investigated in detail. This comprises the initial requirements analysis and specification, followed by concept design and prototyping and evaluation.

As a first step, the designers should become acquainted with the users by gathering general information about them, observing them at relevant tasks, interviewing them, or through active involvement in the design process.

Every iteration is followed by a formative evaluation, while the whole life cycle is followed by a summative evaluation to assess whether a system is ready for exploitation. Usually a formative evaluation results in a list of faults and problems, which might be addressed in the next design cycle. Hereby it may be noted that an iterative development typically reaches economical limits before it is driven to ideal perfection. Every iterative user centered design is based on prototype creation. The simplest prototype might by a scenario telling a design story. Stories can be written in plain text, or presented as a sketch based story board. They commonly describe certain relevant use case situations (Dix, 2004).

REQUIREMENTS ENGINEERING, SPECIFICATION, AND BRAINSTORMING

Requirements capturing should be a major part of every system design, and it is crucial for iterative human centric design. Requirements set the frame for the whole system design, and they are fundamental for conceptualization and evaluation in every step. Therefore all relevant aspects, regarding the user, the system environment, and target domain, should be collected (Dix, 2004; J. Nielsen, 1992; Rogers et al., 2011). In the following a brief overview will be given on requirements specification, for further details it is referred to corresponding literature.

In general it is distinguished between functional and non-functional requirements. While *functional requirements* address the systems functionality, *non-functional requirements* are constraints that might hinder the use of a system, such as performance or resource constraints. (Rogers et al., 2011) present a more detailed differentiation. They distinguish between functional requirements, data requirements, environmental requirements, user requirements, and usability requirements. Hereby *environmental requirements* are also described as *context of use*. The environment includes four main aspects to be looked at: the physical, social, organizational, and technical environment. The *physical environment*, for instance, could be influenced by lighting conditions or spacing of a work place, the *social environment* holds aspects of collaboration and communication, and so on.

Verification and validation of a system is being conducted based on high-level costumer requirements. Additionally, the system needs to be internally consistent and complete. A crucial question in the process of usability engineering comprises the measures of success for a system (Dix, 2004; J. Nielsen, 1992; Jakob Nielsen, 1994a; Whiteside, Bennett, & Holtzblatt, 1988). The challenge for designing a usable system is the identification of finding the right criteria, which finally lead to a positive judgment of the usability of a product. The clear definition of design goals, and the proper rating of their level of completeness, reached in the end, can be very challenging. Therefore, a requirements specification should provide guidance for agreeing on such measures. In the design process of user centered systems, the requirements specification should also include a usability specification, giving usability guidance. "*The major feature of usability engineering is the assertion of explicit usability metrics […]*" (Dix, 2004, p. 204)

(Whiteside et al., 1988) describe techniques for creating a usability specification exemplified on the design of a programming a video cassette recorder control panel. They propose to determine all relevant interaction attributes, and define six terms for each of those attributes. The six terms are used to specify each attribute, also declaring success measures. Table 4-1 exemplifies such terms, by the example of the usability specification of a video cassette recorders backward recoverability:

	Description	Example Attribute: Backward recoverability of a video cassette recorder
Measuring concept:	Concrete description of the attribute	Undo an erroneous programming sequence
Measuring method:	How will the attribute be measured?	Number of explicit user actions to undo current program
Now Level:	Current state available	No current product allows such an undo
Worst Case:	Lowest acceptable solution	As many actions as it takes to program in mistake
Planned level:	Design goal	A maximum of two explicit user actions
Best case:	Best possible with current tools and technology available	One explicit cancel action

Table 4-1. Six terms for addressing all relevant interaction attributes of a usability specification by (Whiteside et al., 1988).

The usability objective is described first, followed by a definition of its measuring methodology, and by an outline of the success criterion, consisting of the now level, worst case, planned level and best case. Usually it is aimed for an improvement of the current state. The characterization of the current state is followed by the lowest acceptable measure, which is the worst case. The planned level identifies the actual design target considered to be feasible, and best case describes the best state possible, with currently available tools and technology.

According to (Whiteside et al., 1988), the current level is defined through an existing system, competitive systems, the task without the use of a computer system, an absolute scale, an own prototype, user's own earlier performance, each component of a system separately, or a successive differentiation between best and worst values observed in user tests. Also the following possible quantitative measures are presented.

Quantitative usability measures of a given system (after (Whiteside et al., 1988) adapted by (Dix, 2004))
Time to complete a task
Percent of task completed
Percent of task completed per unit time

Ratio of successes to failures
Time spent in errors
Percent or number of errors
Percent or numbers of better competitors
Number of commands used
Frequency of help and documentation use
Percent of favorable/unfavorable user comments
Number of repetitions of failed commands
Number of runs of successes and of failures
Number of times interface misleads the user
Number of good and bad features recalled by users
Number of available commands not invoked
Number of regressive behaviors
Number of users preferring your system
Number of times users need to work around a problem
Number of times the user is disrupted from a work task
Number of times user loses control of the system
Number of times user expresses frustration or satisfaction

Table 4-2 Quantitative system usability measures (after (Whiteside et al., 1988) adapted by (Dix, 2004)).

ISO 9241 additionally emphasizes the three categories of usability: *Effectiveness*, *Efficiency*, and *Satisfaction* (International Standards Organisation, 2008). The measures might be applied for all three these categories.

Capturing Requirements

In general, the requirements analysis starts off with an initial capture of system behaviors and processes, producing big amounts of raw data. Every environmental situation demands its own capturing technique. In a next step, the raw data needs to be structured and organized. Therefore, unstructured aural, visual, or textual information is transformed into models or other representations. Such models provide an abstract perspective on analyzed systems and processes, which is helpful for further analysis and interpretation. In the following some capturing techniques will briefly be described, and a set of models are presented.

(Rogers et al., 2011) name five major methods for requirements capturing: questionnaires, interviews, focus groups or workshops, naturalistic observation, and the study of documentations. Questionnaires are a common tool for capturing information. They usually consist of simple yes or no answers, multiple choice or even free text input. Designing questionnaires is a well-researched field, with an extensive pool of guiding literature. Also very common are interviews, may they be oral, written, formal or informal, structured, unstructured, face to face, or remote. A popular variant are group interviews, which usually demand a structuring frame, often induced through workshops, or the formation of focus groups. Such interviews might as well be oral or even online using techniques such as blogs, or forums.

A deeper insight into process workflows is given through observation of users in their real environment, and task performances measurements. Observations may be completely passive, or participatory, where the observer is actively involved in various processes. In turn, users can also be involved into the prototyping process.

Some empirical guidance for the data-gathering techniques, described above, can be found in (Rudman & Engelbeck, 1995). The overview, shown in the following table, was developed during the design process of a complex graphical user interface for a telephone company.

Technique	Good for	Kind of data	Advantages	Disadvantages
Questionnaires	Answering specific questions	Quantitative and qualitative data	Can reach many people with low resource	The design is crucial Response rate may be low. Responses may not be what you want
Interviews	Exploring issues	Some quantitative but mostly qualitative data	Interviewer can guide interviewee if necessary. Encourages contact between developers and users	Time consuming. Artificial environment may intimidate interviewee
Focus groups and workshops	Collecting multiple viewpoints	Some quantitative but mostly qualitative data	Highlights areas of consensus and conflict. Encourages contact between developers and users	Possibility of dominant characters
Naturalistic observation	Understanding context of user activity	Qualitative	Observing actual work gives insights that other techniques can't give	Very time consuming. Huge amounts of data
Studying documentation	Learning about procedures, regulations, and standards	Quantitative	No Time commitment from users required	Day-today working will differ from documented procedures

Table 4-3. Overview of requirements data-gathering techniques, from (Rudman & Engelbeck, 1995).

Each technique has its advantages and disadvantages. For example, Questionnaires might deliver valuable quantitative information, with comparably little effort. However, the final results are heavily depending on the questions-answers schema. Since they are created by the interviewee, they hardly give completely unforeseen answers. Interviews give more freedom for the interviewed, and may reveal completely unexpected new aspects. However, a dialog or discussion may also easily drift into a wrong direction, and their evaluation is more laborious. (Olson & Moran, 1995) point out that the decision about the right capturing method is also influenced by economic factors. Tradeoffs are often made regarding time consumption and the gained level of detail. For example, one has to decide whether it is better to gather data from thousands of participants, using a questionnaire, or interviewing only a sample of them.

(Maiden & Rugg, 1996) presented a framework for supporting the requirements engineering process. The framework named ACRE gives guidance based on six major factors influencing requirements acquisition. Hence the quality of results is influenced by the purpose why requirements are captured (*purpose of requirements*), the type of knowledge acquired depending on the method used (*knowledge types*), incomplete or incorrect knowledge of a stakeholder (*internal filtering of knowledge*), knowledge that cannot be communicated (*observable phenomena*), the contextual influences of the acquisition, such as political or financial (*acquisition context*), and the right mixture and sequence of methods applied (*method interdependencies*). They investigate those six factors on a representative sample of requirements techniques available such as observation, unstructured interviews, structured interviews, protocol analysis, card sorting, laddering, repertory grids, Brainstorming, rapid prototyping, scenario analysis, RAD Workshops, and ethnographic methods. Observation, scenario analysis, RAD, and ethnographic methods are rated to be most valuable for the acquisition of behavioral requirements. Unstructured interviews, structured interviews, protocols, laddering, repertory grid analysis, and rapid prototyping are rated to be a less useful. Card sorting is considered to be least helpful for behavioral requirements acquisition. (Maiden & Rugg, 1996) also provide an overview of conditions and resource constraints of each method, as detailed in the following tables:

Knowledge type	Observation	Unstructured Interviews	Structured interviews	Protocols	Card sorting	Laddering	Repertory grid analysis	Brainstorming	Rapid prototyping	Scenario analysis	RAD	Ethnographic
Behavior	✓✓	✓	✓	✓	-	✓	✓	✓	✓	✓✓	✓✓	✓✓
Process	✓	✓	✓	✓	-	✓	✓	✓	✓	✓✓	✓✓	✓✓
Data	-	-	-	-	✓✓	✓✓	✓	-	✓	✓	✓✓	✓

Table 4-4. Effectiveness of methods for acquiring different types of knowledge, from (Maiden & Rugg, 1996).

constraints	Observation	Unstructured Int.	Structured Inter.	Protocols	Card sorting	Laddering	Repertory grid analysis	Brainstorming	Rapid prototyping	Scenario analysis	RAD	Ethnographic
Meeting is needed	-	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	x
Prepare session time	✓✓	✓✓	-	✓	✓	✓✓	-	✓✓	-	-	✓	✓✓
Session acquisition time	✓	✓✓	✓✓	✓✓	✓✓	✓✓	✓✓	✓✓	✓✓	✓✓	x	x
Obtain requirements time	-	✓	✓	-	✓	✓	✓	✓	✓✓	✓	x	x
# requirements engineers	1	1	1	1	1	1	1	1	1	1	1	1
# stake-holders	1	1	1	1	1	1	1	1	1	1	6	2
Friendliness to stake-holders	-	✓✓	✓	-	✓✓	✓✓	✓✓	✓	✓	✓	x	x
No technological overheads	✓	✓✓	✓✓	-	✓✓	✓✓	✓✓	✓✓	x	✓✓	x	✓

Table 4-5. Conditions and resource constraints for different requirement acquisition methods, from (Maiden & Rugg, 1996).

After requirements raw data is captured and documented usually is transformed into models. Models help understanding existing systems in the analysis and design process (Sommerville, 2001). Multiple model notations have been developed for representing raw data in a new perspective, some of them are shown in Figure 4-4. Depending on the chosen representation a model can deliver a contextual perspective, a perspective showing the behavior of the system, or a structural perspective.

The boundaries of the system are best assessed through *Context Models*, which represent the environment. Since they do not show the relations to other systems or processes, they are oftentimes supplemented by *Process Models*. *Data Flow Models* visualize how data is processed, stored, or exchanged within a system. In the notion of the Data Flow Model, rounded rectangles represent processing steps, arrows represent data flow, and rectangles are data storages or sources. A *State Machine Model* can be used to represent an event based perspective on a system. Such a model shows a behavioral representation of event triggered state transitions within a system. *Data Models* deliver a perspective on the data structure and semantic relations between data entities. They are usually used in combination with a Data Flow Model. Most common data models are *Entity-Relation-Attribute Models*. Another valuable perspective on an existing system might be gained by creating *Object Models*. Usually object classes are identified to categorize single instances. Attributes, inheritance, or relations further define those objects. An example of a *UML Object Model*. Another model for representing behavior is provided by a *Sequence Diagram*. Those object behavior models might be based on a certain scenarios, showing action and interaction sequences of entities, and collaboration between stakeholders.

It is emphasized that a system models are abstractions rather than being an alternative representation of real systems. Abstraction comes with information reduction. According to Sommerville limited support for non-functional system requirements is one of the major weaknesses of such representations. Moreover, they lack guidelines for design support, might lead to a documentation overload, and are often hard to understand (Sommerville, 2001). Hence, for user-centred design, requirements data representations such as scenarios, use cases, essential use cases, and task analysis may be better suited (Rogers et al., 2011).

Task Analysis

A Task analysis is a special technique for investigating existing systems or situations. A task analysis might be supported through scenario development (Rogers et al., 2011). Usually user tasks are previously analyzed and matched with requirements. Additionally, possible new tasks are identified. Such new tasks are also part of the requirements and are typically gathered through empirical observation existing procedures and task performances. Obviously it is hard for a designer to foresee all tasks and requirements, prior system implementation. Throughout an advancing design process new tasks are revealed and the design is adapted (Dix, 2004) In the following some task analysis techniques will be described.

The *Hierarchical Task Analysis* (HTA) introduced by (Annett & Duncan, 1967) is one of the most common methods for performing task analyses (Rogers et al., 2011). HTA is a goal based approach, and is being applied in many different fields (Astley & Stammers, 1987).

All involved tasks for reaching a certain goal need to be determined and organized hierarchically. Additionally, the semantic or procedural relations between tasks are described. Described actions might be of physical or observable nature. I we analyze the chopping of wood for example[16], one needs to get ready to chop wood, put a piece of wood in position, chop the peace, and test if it is properly chopped. The sequence above only describes one case. If the piece of wood was not chopped properly it would have to be chopped again. All different kind of situations, cases, or procedures are covered by the HTA using so called plans. There are multiple notions for listing the task hierarchy and related plan. A simple notion proposed by (Annett & Duncan, 1967) would look as follows:

0. In order to chop wood
1. Get ready to chop wood
1.1. get axe
1.2. get wood
1.3. find proper chopping surface
2. Put piece of wood in position
3. Chop piece of wood
3.4. Hit piece of wood central
4. Test chopped piece of wood
5. Hit piece of wood again
5.5. Turn axe around and chop upside down
5.6. Aim for indent

Table 4-6. Example of a HTA task hierarchy. Notion proposed by (Annett & Duncan, 1967)

Plans describe different procedures for reaching the desired goal. Plans also cover various alternative cases. Every level of the hierarchy has its own plan. Tasks are recursively broken down into sub tasks (Table 4-7).

[16] A quote by Albert Einstein: "People love chopping wood. In this activity one immediately sees results."

plan 0:	do: 1-2-3-4. If piece of wood is not chopped properly do: 2-4-5
plan 1:	do: 1.1-1.2-1.3
plan 3:	do: 3.1
plan 5:	do: 5.2. if axe sticks in wood do: 5.1

Table 4-7. Example of a HTA plan. Notion proposed by (Annett & Duncan, 1967).

One of the challenges in the HTA is the stopping rule, when a task should not be broken down into a subtask analysis anymore. (Annett, Duncan, Stammers, & Gray, 1971) formulated the $P*C$ rule as a stopping rule. The task analysis should be stopped when the probability of failure (P) multiplied by the cost of failure (C) exceeds an acceptable level. The formula can be seen as guideline rather than being accurately applied. The exact calculation might also be comparably time consuming, since the probabilities of failure are often unknown and have to be approximated (Neville A. Stanton, 2006).

The original method of the HTA has been further developed and complemented by multiple researchers in several fields. For example, heuristic support is provided through sets of questions, which support the HTA process of finding sub-goals in various fields such as training design (Table 4-8), interface design (Table 4-9), or job design (Table 4-10). Other support is provided through sub goal templates.

Training Design
What is the goal of the task?
What information is used for the decision to act?
When and under what conditions does the person (system) decide to take action?
What is the sequence of operations that are carried out?
What are the consequences of action and what feedback is provided?
How often are tasks carried out?
Who carries the tasks out?

Table 4-8. Questions supporting the HTA process of finding sub goals in training design (Piso, 1981)

Interface design
What are the sensory inputs?
How can the display of information be improved?
What are the information processing demands?
What kind of responses are required?
How can the control inputs be improved?
What kind of feedback is given?
How can the feedback be improved?
How can the environmental characteristics be improved?

Table 4-9. Questions supporting the HTA process of finding sub goals in interface design (Hodgkinson & Crawshaw, 1985)

Job design
How does information flow in the task?
When must tasks be done?
What is the temporal relation of tasks?
What are the physical constraints on tasks?
Where can and cannot error and delay be tolerated?
Where is workload unacceptable?
Where is working knowledge common to more than one task element?
Where do different tasks share the same or similar skills?

Table 4-10. Questions supporting the HTA process of finding sub goals in job design (Bruseberg & Shepherd, 1997).

Practical experiences with the HTA showed that basic usage may be acquired relatively quickly, while a complex usage of this methodology demands expert guidance, and some month of experience (N. A. Stanton & Young, 1999). (Ormerod & Shepherd, 1998) also report that the use of sub goal templates improves the acquisition of the HTA by novices. Computerized sub goal template support led to even better results. Frameworks for HTA have been developed on pen and paper basis, and as digital variants.

Conceptualization and Prototyping

Design concepts for interactive systems often evolve from ideas, often generated with brainstorming techniques, and assessed according to previously determined requirements, experience based heuristics, and guidelines. Those concepts are usually early represented in a prototypical form. Prototypes are crucial for any iterative system design. Several different types of prototypes might be used in an iterative process. Simple throw-away prototypes potentially improve the specification process, already during the requirement specification phase. Moreover prototypes can also evolve to the final product. Throw-away prototypes are usually created using rapid prototyping techniques. (Gould & Lewis, 1985) state that an iterative development phase demands fast, flexible prototyping, and highly modular implementation, to permit early testing and easy adaptation. For rapid prototyping several techniques are available such as the use of scenarios, storyboards, paper prototypes, Wizard of Oz technique, or even simplified simulations.

A very simple but also helpful technique, often not considered as prototyping, is the development of *scenarios*. Scenarios are concrete descriptions of situations in a certain context. Scenarios, or *use cases*, are also often used for requirements acquisition, or for determining task definitions. In the context of this work, scenarios are only seen as a platform for the designers to express their ideas, and for the users to express their needs. Capturing requirements based on use cases demands the three major steps of a normal iteration: first the scenario should be conceptualized, then implemented, and finally evaluated. All three steps might be accomplished by both, designers and stakeholders (Rogers et al., 2011). (Dix, 2004) sees scenarios as rich stories of interaction, useful for considering the usage of a system in more detail, and for creating a concrete basis of communication with users, developers, or other stakeholders. They might be iteratively adapted and help for deeply considering concepts.

Storyboards may be seen as a more visual representation of scenarios, known from film industry. However, storyboards do not focus on certain situations. Thy usually provide snapshots of system-user interaction, and showcase concrete concepts of interface and interaction design, including action sequences.

In graphical user interface design *Paper Prototypes* are often used, as graphical interface mockups, providing a first look and feel of an interface. Usually elements of interaction are hand drawn on paper. They can be cut out and freely arranged. This way it is possible to discuss and quickly modify several interface designs. Even interaction might be tested by human simulated interaction behavior.

The idea of human simulated computer behavior is also known from the *Wizard of Oz* technique. *Wizard of Oz* is commonly tested with potential end-users, which interact with an interface controlled by a designer. The designer mimics expected results, while the user has the impression of interacting with an artificial system.

Moreover, rapid prototyping tools, such as *HyperCard,* allow for rapidly designing and testing graphical user interfaces. *HyperTalk* gave further possibilities to add scripted computer behavior to those prototypes, created with HyperCard (D. Goodman, 1987). This way, prototypes could seamlessly be transformed into functional systems, which were typically inefficient and limited in terms of performance, resource management, and responsiveness.

Systems which iteratively evolve from an early prototype to a final product are usually more complex. At a certain point in the design process they are more economical as throw-away prototypes. The focus in an iterative design process often lies on non-functional features; functional features, such as safety or reliability issues, are often neglected. Designers should be aware of such issues and consider those features in their concepts, even if they are not implemented in early prototypes (Sommerville, 2001).

EVALUATION

Typically three major steps are performed per iteration during a HMI design process: a conceptualization phase is followed by an implementation phase, which is then evaluated. The results of the evaluation phase kick start the next iteration, demanding a re-conceptualization, re-implementation, and so on (Dix, 2004; Jakob Nielsen, 1989; Rogers et al., 2011). The importance of proper evaluation is fundamental for any further development and evaluation. Budget and time often do not allow for a quantitative evaluation, at every iteration.

It can be differentiated between *model based*, *heuristic*, and *empirical* evaluations (Preim, 1999). The evaluation of a system helps validating concepts and identifying specific problems of a system. Several methods are available ranging from usability laboratories to field analysis. Evaluations may involve users, or may be limited to experts and the designers themselves (also compare (Jakob Nielsen, Mack, & Shirk, 1996; Jakob Nielsen, 1994b))

A number of evaluation methods, may be applied early in the design process, even before usable prototypes exist. A popular example for such a method is the *Cognitive Walkthrough* of an expert. Therefore, experts are provided with a specification or prototype of the system, a task list, a list of user actions necessary to perform each task, and general information about the users' knowledge and experience. Based on this information, the experts evaluate the system, answering the four specific questions: Do the functions of the system meet the users expectation? Are all functions visible for the user? Are users able to utilize the provided functions properly? Will users understand provided feedback?

Evaluation might also be based on cognitive and design models. A well-known model of this kind is the *GOMS* (goals operators, methods and selection) model. Such models are meant to give guidance on design decisions.

The above evaluation techniques are helpful for early system evaluation. Later in the design process user based evaluations are inevitable, and necessary for a successful system design. While the testing environment of first prototypes might situated in a laboratory, at some stage the system should be tested in the real user environment. Taking into account, that fundamental changes in the system design becoming increasingly expensive with raising system complexity, it should be considered to take the extra effort, and start early testing in the real environment (compare participatory design and ethnographic methods).

Other evaluation aspects such as the choice of participants, amount of tests, and setup of experiments are not relevant for this work, yet. Procedural aspects are more relevant for the later framework design.

CONCLUSIONS

Digital Transformatives are highly systems, which closely interact with its users. Since user behavior cannot be predicted by system designers, an iterative procedure is required in the development process (Jakob Nielsen, 1989; Preim, 1999; Rauterberg et al., 1994). Hereby, sufficient user involvement seems to be essential. Early focus on users, already prior the design phase, is advantageous, and may be applied in different variants, ranging from a passive task based user analysis, to more active involvements, such as interviews, discussions, or participatory design (Dix, 2004; Gould & Lewis, 1985; Harslem & Nelson, 1982; Preim, 1999; Rogers et al., 2011; Seybold, 1981; D. C. Smith et al., 1990).

The basis for every design is a requirements analysis and specification. It can be distinguished between functional requirements, data requirements, environmental requirements, user requirements, and usability requirements (Rogers et al., 2011). A crucial question in the process of usability engineering comprises the measures of success for a system. It is hard to clearly define design goals and rate their level of completeness reached in the end (Dix, 2004; J. Nielsen, 1992; Jakob Nielsen, 1994a; Whiteside et al., 1988). Such measures can be deduced from the requirements. The measures' boundaries are defined by the interval of acceptance, as illustrated in Figure 4-3.

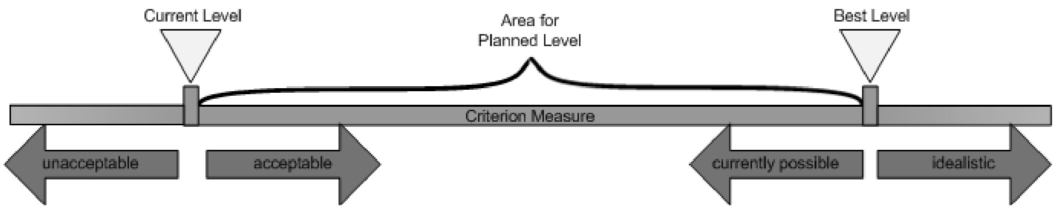


Figure 4-3. Interval of acceptance of measurement levels in a usability specification (adapted from (Dix, 2004; Whiteside et al., 1988)).

The interval of acceptance ranges from the current level to the best possible level. Every planned interaction feature should improve the current state, while not being to idealistic. A usability specification template, promoting the definition of the interval of acceptance, is given by (Whiteside et al., 1988). For each interaction attribute they propose to define a measuring concept, a measuring method, the current level, the worst case, the planned level and the best possible case.

Multiple requirement capturing techniques are available, such as structured and unstructured interviews, focus groups and workshops, naturalistic observation of users or task performances, the study of documentations, card sorting, brainstorming, and scenario analysis (Rogers et al., 2011). (Maiden & Rugg, 1996) assessed those regarding their effectiveness for capturing behavioral requirements. For further investigations, the raw requirements data is further structured, and transformed into model representations such as, context models, behavioral models, data models, or object models (Sommerville, 2001). Those models offer various perspectives on existing processes and dependencies, giving a solid basis for further interpretations and implications on an improved system design (compare Figure 4-4).

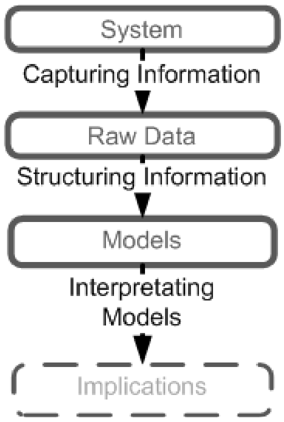


Figure 4-4. General steps of Requirements Engineering.

However, most of those models only offer limited support for non-functional, interactive systems (Sommerville, 2001). A common approach for user-centric requirement captures, of existing systems and situations, is given by task analyses, making it overly interesting for the design rational of Digital Transformatives. Task analyses are conducted in multiple ways. Either users are observed, or they are actively asked for task descriptions, for example through scenario development. The Hierarchical Task Analysis (HTA) introduced by (Annett & Duncan, 1967) is one of the most common methods of that kind (Rogers et al., 2011). HTA is a goal based approach and is being applied in many different fields (Astley & Stammers, 1987). Tasks are structured hierarchically, in a model of tasks and sub-tasks, which are connected through semantic and procedural relations. The overall design process of iterative systems is summarized schematically in Figure 4-5.

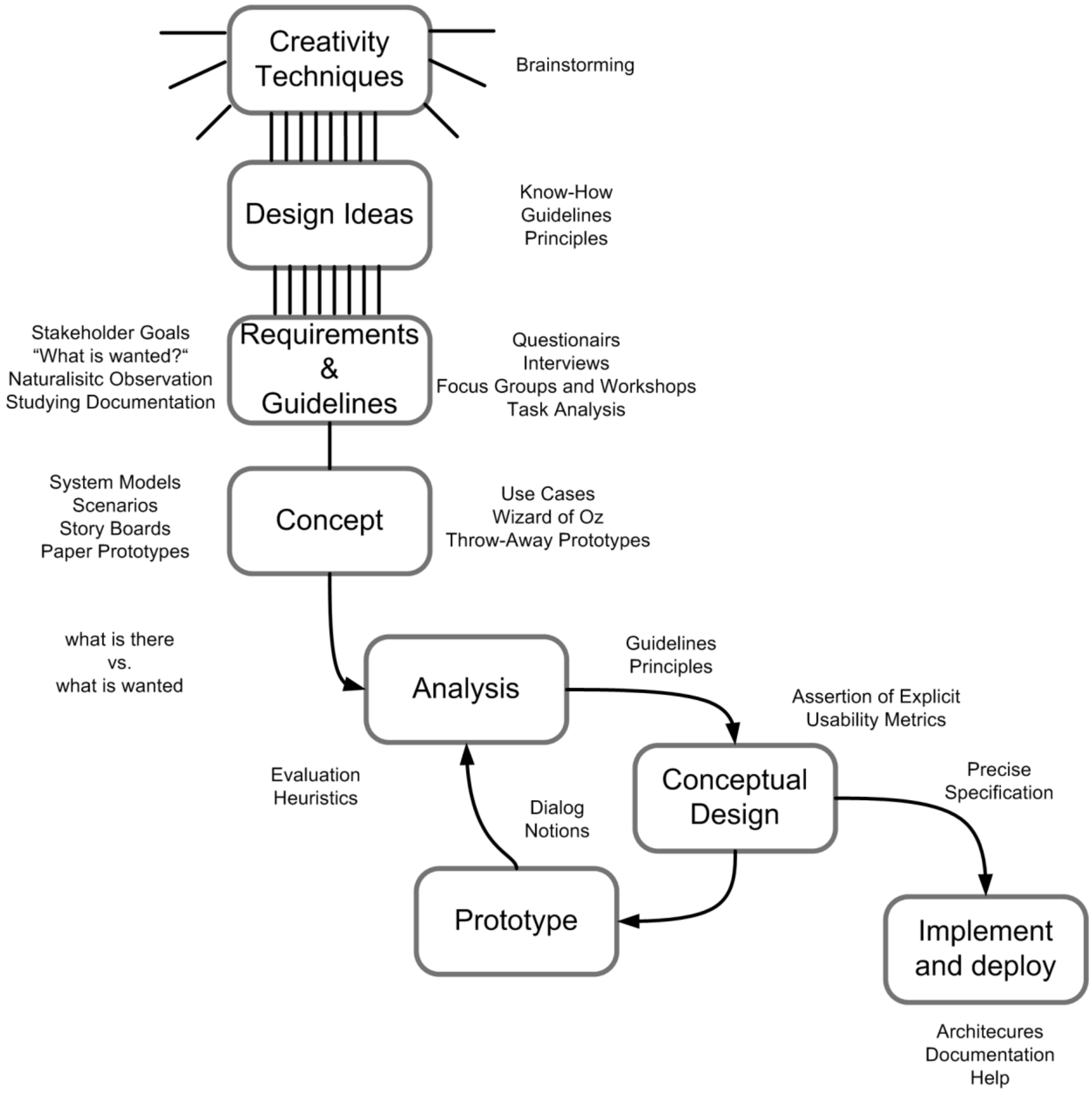


Figure 4-5. Iterative interaction design process as a water fall model including an iteration cycle (adapted from Dix, 2004).

Most human centered system designs start with creativity sessions, leading to concept ideas. A concrete prototypical representation of those concepts allows for an early start into the iterative cycle. For rapid prototyping, several techniques are available, such as the use of scenarios, storyboards, paper prototypes, Wizard of Oz technique, or even simplified simulations (Dix, 2004; Rogers et al., 2011). An iterative development phase demands fast flexible prototyping and highly modular implementation, to permit early testing and easy adaptation (Gould & Lewis, 1985).

In an iterative design process evaluations set the basis for further system refinements, moreover, they help fulfilling requirements and assessing system quality. Multiple types of evaluations for interactive systems are being used; heuristic evaluation, cognitive walkthroughs, formal usability inspections, feature inspections, consistency inspections, are some of them (Jakob Nielsen et al., 1996; Jakob Nielsen, 1994b).

4.2 Interactive System Design Methodology for Digital Transformatives

Digital Transformatives (DT) are based on system concepts, which shift the program performance context into a user familiar context, aiming for cognitive prototypes on a shared basic level (compare Feature 10). Consequently, the main difference between the design procedure of Digital Transformatives and other systems may only be found in system concept design phases. In a common interaction design process this might start with the initial design question “What is wanted?” (compare Dix, 2004).

Design concepts for interactive systems usually evolve from ideas gathered through brainstorming sessions. Such conceptual ideas are assessed based on previously determined requirements, experience based heuristics, and guidelines. Consequently, the basic principles of Digital Transformatives need to be implemented on the brainstorming and requirements level, which is influencing the concept design.

Guideline 3. Basic principles of Digital Transformatives should be implemented on the brainstorming and requirements level, which is influencing the concept design.

Within the design phase, requirements are used to assess conceptual design ideas. Hence, the development of a Digital Transformative may be conducted by considering the basic DT features as essential requirements:

- Feature 1. Digital Transformatives aim for superior user performance by shifting the usage context.
- Feature 4. The user interface provides a bidirectional conceptual mapping between user context and system context through transitional encoding and decoding
- Feature 10. Digital Transformative interfaces are situated in a context with maximum user familiarity, which corresponds to cognitive prototypes on a shared basic level of sufficient target similarity

Such requirements ought to be implemented through familiarity assessment, which, hence, has an outstanding role in the design of DTs. This leads to the following DT specific adaption of the interactive design process.

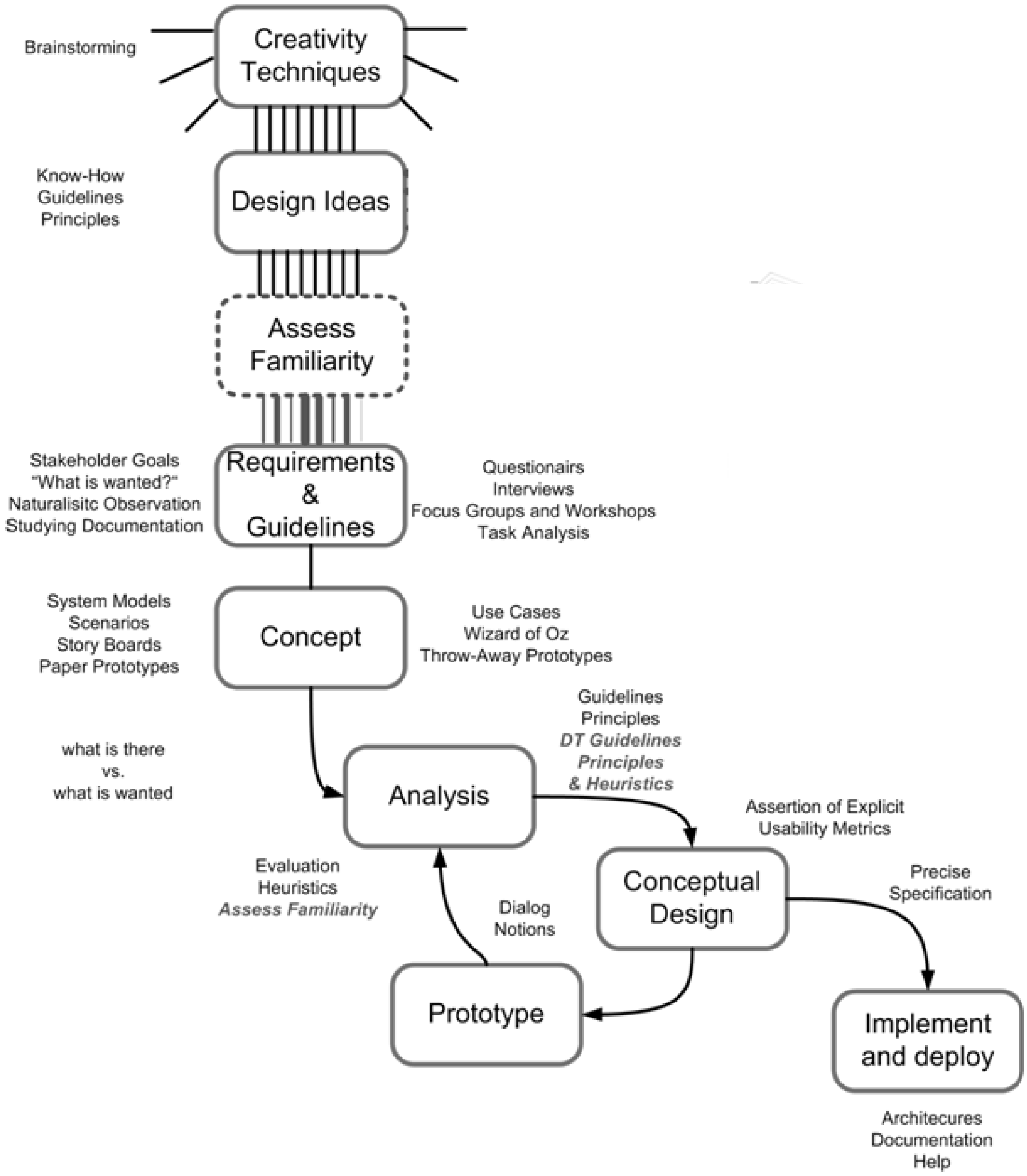


Figure 4-6. First adaption of the interactive system design process to foster the design of Digital Transformatives.

The design process is extended by a block for assessing early design ideas according user familiarity. The same familiarity assessments also apply later in the iterative development phase, during evaluation.

Guideline 4. Assess concepts and implementations by user familiarity.

4.2.1 Methods for Assessing Familiarity

There are many indicators that familiarity corresponds to cognitive prototype categories and well-practiced processes, which describe areas of increased performance (compare Hypothesis 2 and chapter 3.4 *Basic Concept Validation* pp. 115). The main challenge remains in finding such cognitive and procedural prototypes of increased familiarity. The investigations conducted in this work suggest multiple ways to determine familiarity. In the following, four major methods for assessing familiarity will be introduced:

- Heuristic expert estimation
- Probabilistic Environmental Observation
- Learning Curve Analysis
- User Rated Familiarity

HEURISTIC EXPERT ESTIMATION
The *heuristic expert estimation* can be seen as the most effortless assessment of familiarity. However, it is fully based on empirical knowledge, hence, may be biased easily. It also builds heavily on generalization, and might be the most inaccurate of the methods described here.

Guideline 5. Use heuristic expert estimations if you quickly require tendencies.

PROBABILISTIC ENVIRONMENTAL OBSERVATION
The *probabilistic environmental observation* makes use of Feature 5 (*Prototype categories reflect probabilistic real world stimuli of high occurrence frequencies and improved cognitive performance*) and Feature 6 (*The fundamental mechanisms of prototype categories are also active in process automatization through training*). Further details are described in *Concepts, Prototypes, and Categories* (pp. 23) and *Performance, proceduralization, and category prototypes* (pp. 34). Section 3.2 elicits how such features relate to Digital Transformatives. Those features allow for an indirect assessment of user familiarities, by creating a probabilistic map of procedural and cognitive influences, and stimuli of the user environment. Already (da Signa, 1892) highlights

the environmental influence by indicating that signs used for aiding memory are often remarkable occurrences surrounding us. Validation for this method may be given through the comprehensive research in the field of color perception, and the world color survey, as detailed and tested in the following.

Color Perception and the World Color Survey

Color studies had great influence on the proposition of the concept of cognitive prototype categorization. While language seems to play a key role in cognitive prototype categorization processes (Davidoff, 2001; Saunders & Van Brakel, 1997; Whorf & Carroll, 1956), the basic color terms survey of (Berlin & Kay, 1969) raised awareness for language independent cognitive structuring processes.

Since cognitive prototypes formation is heavily linked to the frequency of occurring features. Berlin and Kay's basic color terms may be based on different color frequencies in our environment (Berlin & Kay, 1969). Consequently, the findings of Berlin and Kay imply that we do not perceive the full color spectrum evenly. There should be color peaks in the visual perception of our environment; some colors and color tones occur significantly more often than others. Most likely, those should be similar to the basic colors shown in Figure 4-7.

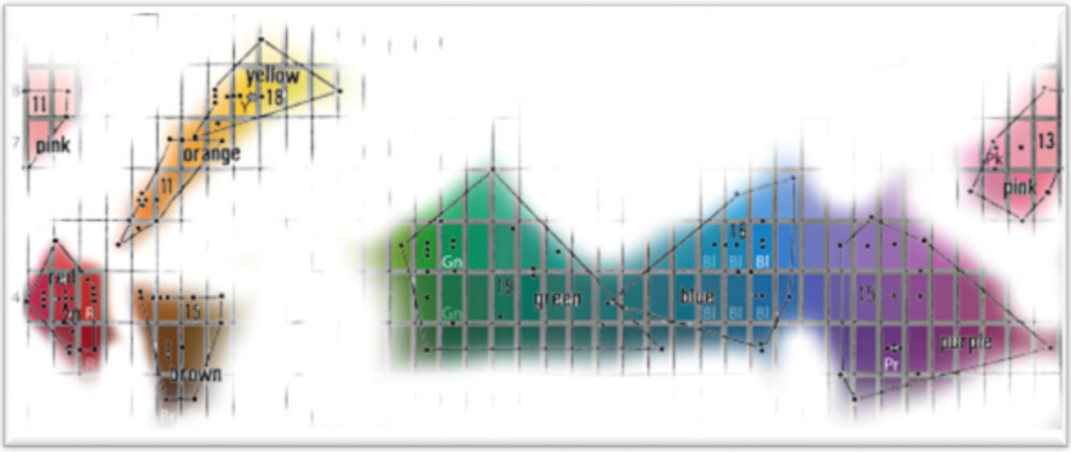


Figure 4-7. Universal focal colors determined by (Berlin & Kay, 1969).

Such a correlation between color occurrences and basic colors would provide us with a comprehensive, and simplified test bed, for investigating mechanisms of cognitive prototype composition.

Most controversially, colors terms were claimed to be universal, which would, for example, give hint for some sort of innate perceptual cognition (Davidoff, 2001). The clear universality, as it was claimed in Berlin and Kay's first study, however, can be seen as refuted. The results of succeeding studies give evidence for language independent cognitive prototyping and categorization.

Rosch found indication that focal colors are formed prior color naming (E. H. Rosch, 1973a). She concluded this finding from the insight that children at the age of 3 years showed preference for focal colors (Heider, 1971a). In another test she also showed that the Dani of Papua New Guinea remembered focal colors more accurately than non-focal colors, although their language only consists of two basic color terms (Heider, 1972). (E. H. Rosch, 1973a) mainly advocated the physiology of color vision for non-language based color names.

Contrarily, advocates of linguistic relativity theory continue to provide evidences for language driven categorization. On the one hand, numerous indicators, strengthening linguistic relativity theory, are provided (Davidoff, 2001). On the other hand, contradictions in studies, which support the universality of basic color terms are highlighted. For example, (Saunders & Van Brakel, 1997) emphasize that Dani color recognition performances were much worse than the performances of Americans (Heider, 1972). Those performance differences give evidences against the universality of color terms, and they found further support by studies of (Davidoff et al., 1999).

Moreover, (Saunders & Van Brakel, 1997) undermine the universality of colors by investigating this topic inter-disciplinarily from various domains, such as cognition and perception, physics, color metrics, developmental psychology, psychophysically, and neuro-physiologically. None of the domains seems to provide clear evidence supporting the universality of color names.

Especially the differences between the first basic color terms study and the comprehensive WCS provide multiple indications on the environmental influences on cognitive color categorization. The original study was ethnically versatile, but conducted with test persons mainly coming from San Francesco bay area. In contrast the WCS was based on different ethnical groups actually living spatially apart, thus being influenced by different environments. In the first study, although informants were linguistically different, their color categorization results were partially even more similar than the results of different informants, speaking the same language. Contrarily, the results of the WCS differed much more, where not only languages differed, but also the actual environment.

The above implications suggest the comparison of perceived color frequencies with the color term centroids, determined by the WCS. This way, one could verify the environmental influence, and additionally investigate cognitive prototype formation, based on the comprehensive test data of the WCS. Another promising advantage of color data lies in its simplicity. Most studies, regarding cognitive prototype generation, build on much more complex and indefinite feature spaces, such as shapes, faces or objects (Neumann, 1977; Posner & Keele, 1967; Reed, 1972; E. Rosch & Mervis, 1975; Solso & McCarthy, 1981). Moreover, the results of the WCS, including the raw data available online, provide a sound cognitive feature space (Cook et al., 2005; Paul Kay & Cook, 2011; Richard Cook et al., 2012). Given the environmental input stimuli, one would have a sound basis for investigating the development of cognitive structures and the development of prototypes. Consequently, one only needs to measure all colors perceived by humans in certain areas of the world, and compare those to the cognitive concepts determined by the WCS.

Test of Environmental Influence on Color Perception

Unfortunately the author was not able to find data about probabilistic color frequencies, perceived by humans. Hence, it has been decided to write a little program for determining the individuals color frequency distributions.

The test is based on the hypothesis that camera snapshots, taken by individuals, offer a quantisized representation of human perception. It is likely that such snapshots provide an interest biased representation, which might be advantageous for this test case, since we tend to blend out unimportant features. To get a raw data color frequency distribution, it would ideally be necessary to install a camera on a subject, which is filming a personal view all day long, for several days. However, snapshots should give an approximation. The increasing propagation of smart phones, offering enhanced snapshot camera functionality, led to a sound source for individual snapshot data, since smart phones are usually at hand at all hours.

Based on the above hypothesis a color frequency analysis application has been implemented. The software was developed in the Smalltalk based environment Pharo 1.4, running on all major operating systems, such as *Mac OS X*™, *Microsoft Windows*, and several Linux derivates (Apple Inc., 2012; A. Black et al., 2009; AlanC. Kay, 1996; Microsoft, 2012; pharo-project.org, 2012; Wikipedia.org, 2012c, 2012d).

In order to gather data, the application needs to be installed on test persons' personal computers. Once the application is installed, users are presented with a user interface as shown in Figure 4-8.

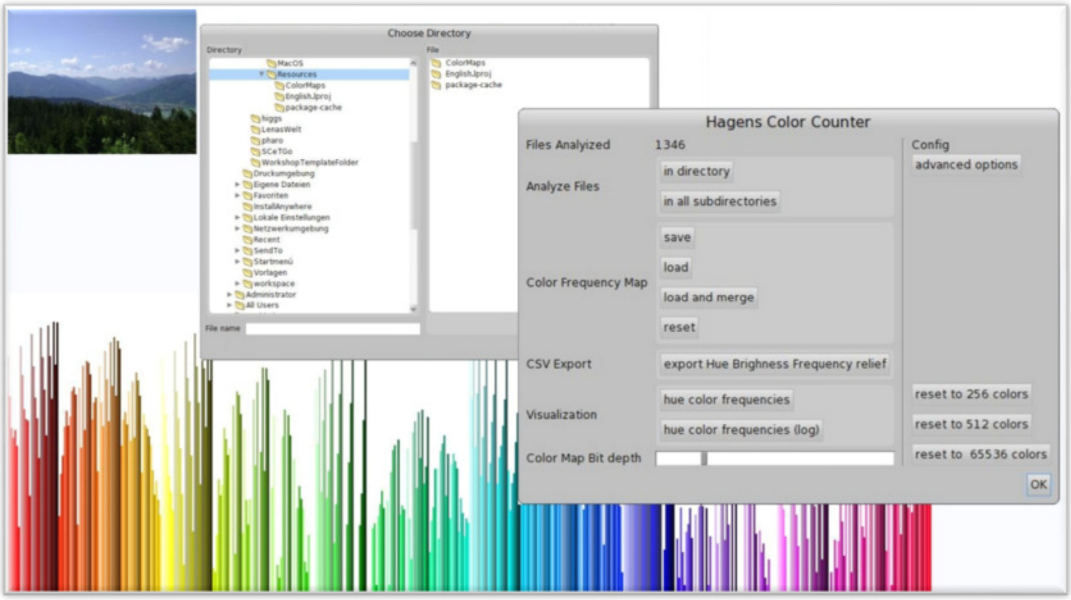


Figure 4-8. A screenshot of the Color Counter application.

The user interface gives access to a set of functions, such as analyzing images in a directory, analyzing images in all subdirectories, crawl online Flickr images, visualize color frequency distributions, load, export, or merge color frequency maps, and some more. Participants of the test only need to select a dictionary holding images of their smartphone or snapshot camera, to start the analysis. After finalization or user interrupt, the data can be saved, and provided to the author for further investigations.

In a prototypical proof of concept test the application was handed out to 6 test persons between the age of 24 and 68, 33% of them were female. The image data spanned the time from 2006 to 2012, and was mainly recorded in Germany. In total 14462 pictures were analyzed.

In order to minimize data load, a series of quantization mechanisms were applied. As detailed above, snapshots themselves already can be seen as a quantization of lifetime and color perception. Light frequencies are captured via CCD, which includes further quantization of the wave length spectrum into red, green, and blue color triplets, and a reduction of resolution from infinite to a certain quantity of pixels. Finally, the snapshots have been scaled down uniformly. The short extent, either x or y axis, was scaled to 240pixel, while the other axis was scaled accordingly in respect to the original aspect ratio. This way, the analysis was more independent from the orientation and resolution of the original snapshot. If the pictures were not scaled uniformly, snapshots of cameras with higher resolution had more impact, since they deliver more pixels. In the same way, with a scaling in only one direction, panoramic picture had comparatively less influence than a vertical snapshot. Additionally, a color space quantization down to 8bit color depth was performed, using standard conversion algorithms.

Figure 4-9 shows the results of the first test. Colors are sorted by hue, following the CIEXYZ standard (Moroney et al., 2002). Since the representation of the Munsell color space is in CIEL*a*b*, results may not directly be comparable, and there should be a minor loss in quality.

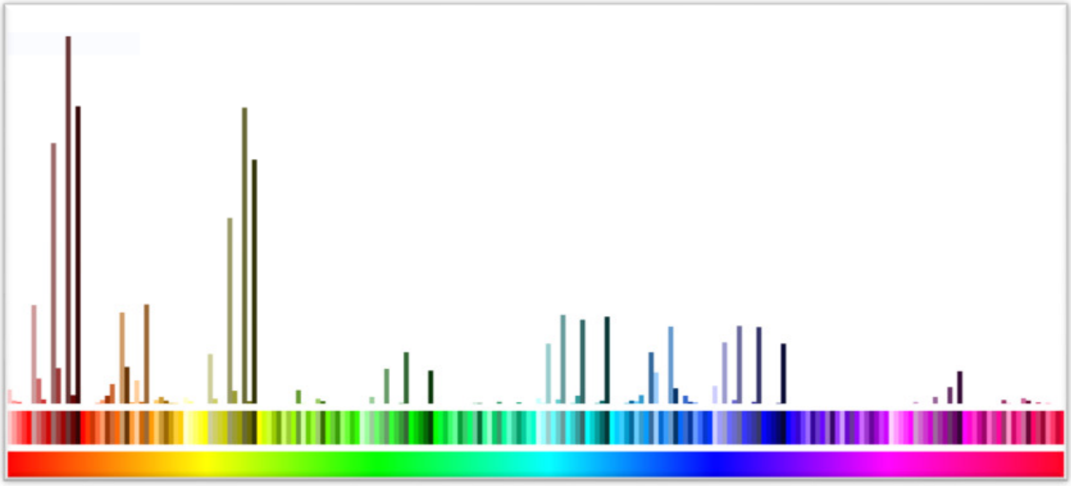


Figure 4-9. Quantized approximation of the accumulated perception of color frequencies in daily life (logarithmic scaling).

Black and white, along with several variations of gray, happened to be most frequent, but are not shown here. The visualization shows clear peaks for certain color tones of the color spectrum, with the two highest at a darker shaded red and a pastel green. The next level of peaks is defined around brown, dark yellow, orange, and blue colors. Purple and pink also form small clusters. This way, the peaks quite well tend to correlate with the focal color hierarchy determined by (Berlin & Kay, 1969) Figure 2-8.

A look at the 3D relief, displayed in Figure 4-10, gives further insights.

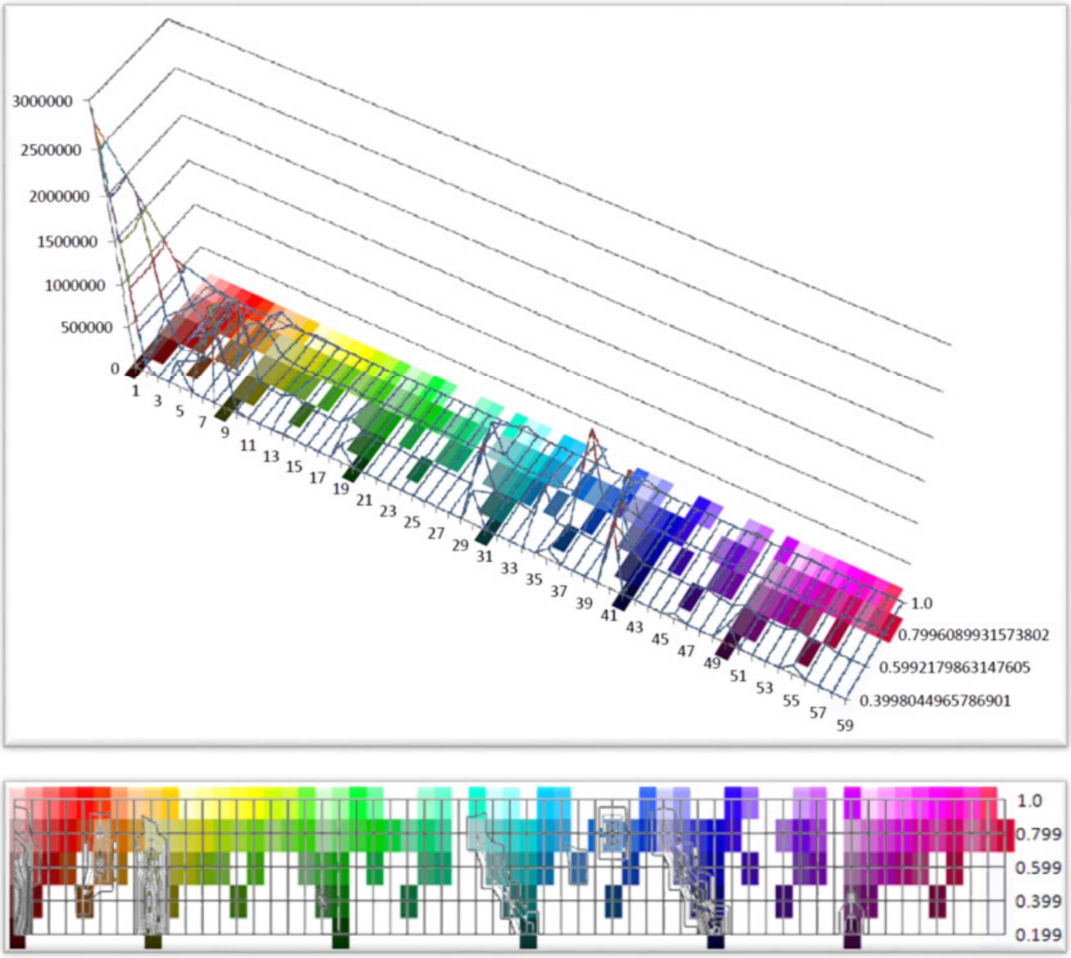


Figure 4-10. 3D representation (top) and relief (bottom) of perceived frequencies color on a 8 bit RGB (hue; brightness; frequency) color space; all logarithmic scaling.

It shows three peaks in the red, yellow, and pastel green area. Yellow, brown and orange form a clear peak, but seem to be close together, and frazzle into green. Also blue appears very frequent, scattered over three further peaks, which range towards green on the one side, and towards purple on the other side.

The diagram shows 10 peaks at red, orange, brown, yellow-green, pastel green, green-blue, blue, dark blue, purple, pink. A visual comparison of such colors to the colors determined by English informants in the original color survey from (Berlin & Kay, 1969) is given in Figure 4-11.

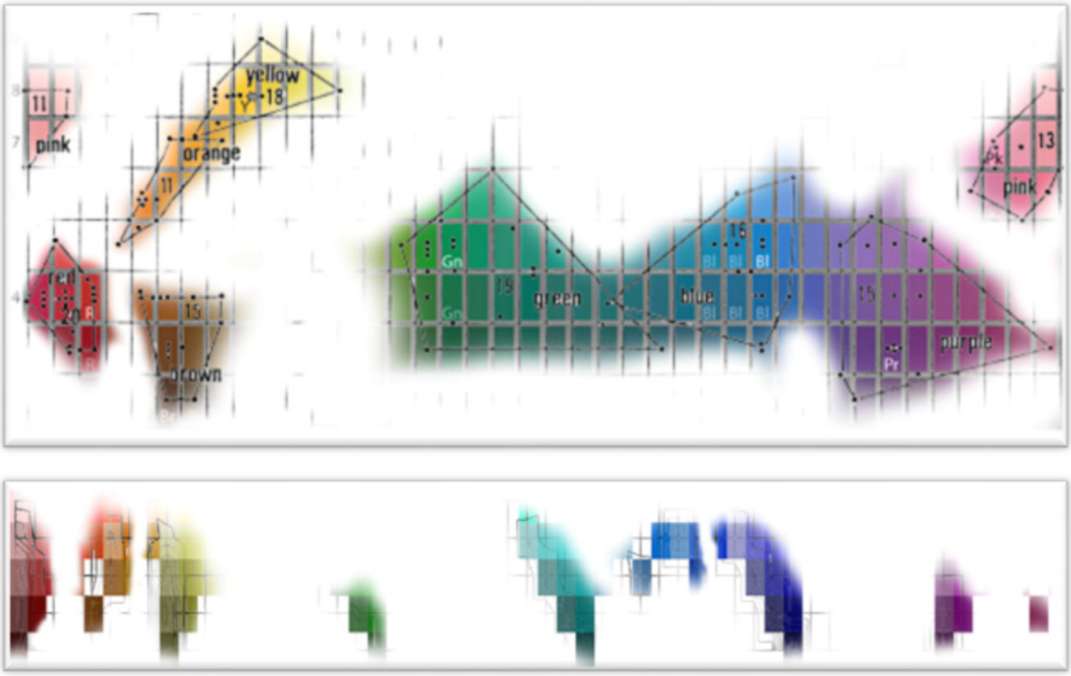


Figure 4-11. Clusters of color frequencies stimulating perception in an environment in Germany, compared to English focal colors determined in San Fransisco bay area by (Berlin & Kay, 1969).

The results of this test reveal a new view on universal color terms. The results provide indicators for a non-uniform distribution of colors in our visually perceived color space. Indeed, color frequencies show a distribution which seems to be similar to the distribution determined by the WCS, and the color term hierarchy (P. Kay et al., 2009; P. Kay & Regier, 2003). The findings above give further evidence that the focal colors determined in Berlin and Kay's first study never were universal, since they were based on informants who were exposed to a very similar environment. The WCS was carried out in different environment all around the world, so in respect to environmental color exposure, it indeed can be considered universal. The distribution of focal colors, or term centroids, showed stronger variations, but still seems to underlie a universal distribution, similar to the basic color terms of the first study.

The results give further hint for the environmental influences of color frequencies on the cognitive formation of basic color terms. A more elaborate test, with high resolution data, and conversion to the CIEL*a*b color system, could provide important data for understanding cognitive prototype categorization. In combination with the sound WCS database archives, we would be provided with a multi-dimensional set of exemplary conversion input values and according output values, providing a sound test bed for further hypothesis on cognitive structuring processes, and possibly revealing involved hidden mechanisms and influences. The following relation expresses this advantage.

If cognitive structures (captured by the WCS term centroids) are formed based on environmental stimuli (captured by the previous test) in combination with further cognitive processes (unknown) for structuring, then the combination of both known data sets help us learn more about the yet unknown cognitive structuring processes.

Conclusion

The findings on the world color survey, and other findings on cognitive prototypes, suggest that cognitive prototype structures correlate with stimuli frequencies. Such stimuli may be induced through the environment or through practice. Moreover, those areas of high performance often correspond to areas of high familiarity, as described in Feature 11, evaluated in section 3.4, and further assessed in the succeeding section on *User Rated Familiarity* (pp. 163). Hence, those features allow for a comparably accurate assessment of user familiarities, by creating a probabilistic map of procedural and cognitive influences.

Guideline 6. If relevant user environmental data can be captured easily, conduct probabilistic analyses of occurrence frequencies of procedural and cognitive stimuli. They offer an accurate measure for cognitive prototypes and familiarity.

LEARNING CURVE ANALYSIS

The *learning curve analysis* is based on research of process automatization or proceduralization, as detailed in section *From controlled to automated processes to habituation* (pp. 38). There is substantial evidence that novel processes are consciously controlled, and gradually transformed into automated high performance processes. The nature of performance gains through automatization has been mathematically expressed in the formula for the power law of practice.

Formula 3. Formula for determining familiarity adapted from power law of practice by (Gordon D. Logan, 2002).

Learning rate results in quicker acquaintance but also earlier stagnation (Gordon D. Logan, 2002). A zero slope is a classic indicator for a task being automatized, compare (Palmeri, 1999; Schneider & Shiffrin, 1977). The insights on the power law of practice lead to the following feature for Digital Transformative:

Feature 13. Familiarity correlates inverse proportionality to the slope of learning curves.

The feature can be used to determine procedural and cognitive familiarity, since the power of law and practice can also be found in pattern recognition, and other cognitive processes, as described in *Performance, proceduralization, and category prototypes* (pp. 34) (compare Chase & Simon, 1973; De Groot, 1978; Vicente & De Groot, 1990). Consequently, the power law of practice may be used for determining cognitive and procedural familiarity. Therefore two or more reference points are demanded. They may be captured by measuring performances of certain procedures or stimuli response mechanisms. Those reference points can further be used with Formula 3 to assess the current state in the learning progress, and the level of automatization.

Guideline 7. If user relevant performances are at hand or easily measurable, capture learning curves. Stagnating learning curves are indicators for areas with high procedural or cognitive user familiarity.

<u>User Rated Familiarity</u>

User rated familiarity is based on findings from cognitive prototype research. On the one hand, Rosch and Mervis found that subjects could reliably rate, to which extent a stimulus would be a cognitive prototype (E. Rosch & Mervis, 1975). Additionally such rating predicted performance in a number of tasks. Such findings are supported by further studies on cognitive prototypes, indicating increased performance at prototype categories, which were formerly determined through user rating

Hypothesis 4. Users are able to accurately rate their own familiarity, which reflect their performances.

This hypothesis will be tested through a modified version of the first memory test, conducted in chapter 3.4.2. The test should also provide additional validation for the basic concept of Digital Transformatives

The first memory test indicated the advantages of familiar over non-familiar memory items. For further verification, several flaws of the first test will be diminished in this test. Two aspects seem to be prevailing. Firstly, the sample rate should be increased, and secondly the factor of luck has to be reduced. Besides verifying whether familiarity leads to better performance, it is also of major interest, to investigate further details on the correlation between the both. Additionally, it is important to know to which degree it is possible for users to rate familiarity, and whether this can directly be linked to performance gains. The requirements stated above lead to changes in the application logic of the initial Memory, as it is described in the following paragraph.

Setup, Apparatus and Test Procedure

To get a higher sample rate the test has been conducted as an online test, potentially reaching out for a bigger mass of persons. A disadvantage of online tests is the anonymity of users, which makes it harder to guarantee a proper task execution. Online tests also might lack a personal feedback channel. Hence, every test showed a voucher for some candy, ready to be picked up at the interviewer's office, after test completion. This way, it can be distinguished between a test group personally known by the interviewer, and an anonymous group which.

The following analysis will first only focus on the personally known test users. They knew about the test case, but they were not informed about the goal of this test. 21 subjects participated at the full set of tests, with 28.5 % of them being female. The youngest participant was nine years old; the oldest had an age of 54. A majority of 70%, of the test persons, were between 25 and 35 years old, the rest were distributed almost equally over the other ages.

The subjects were invited to conduct the test at their own leisure, when and where they wanted. Only prerequisite was the availability of a web browser and internet access, to open the test webpage. In order to reduce the factor of luck, the game logic of the first test was adapted. Originally the positions of all cards are randomly revealed by the user. A very lucky player could uncover all matching pairs in a row. To reduce this factor, the game started with open cards which were flipped after a certain period of time. This way, the factor of luck was reduced only to those cards, which were not memorized in the initial phase.

The new test was implemented using the Seaside web development framework and Pharo programming environment (Ducasse, Renggli, Shaffer, & others, 2010; pharo-project.org, 2012; Seaside.st, 2012).

The tests consisted of 4 runs. Similar to the memory game of the first test, faces of different levels of familiarity were used. Therefore a pool of 42 cards was generated, each card tagged with one of the three categories of generally known personalities, currently popular celebrities, and random strangers.

The first run started off with 9 cards, consisting of three generally known persons, three not typically known faces, and three strangers, randomly selected from each category. On startup they were shown at arbitrary positions in a five by five square, as shown in Figure 4-12. An implementation of the *Lehmer random number generator algorithm* was used to select the cards and choose their positions (Park & Miller, 1988).

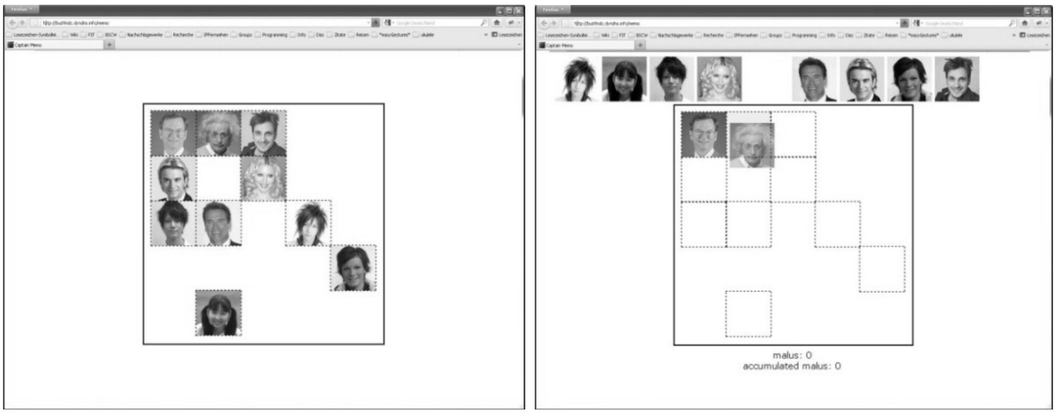


Figure 4-12: User determined familiarity test based on memorizing faces. Screencasts showing the memorization phase (left), and the assignment phase (right).

After 10 seconds the faces slowly faded out, leaving an empty square behind. After the memorization phase, they showed up above the assignment area as “draggables”. Via drag and drop, players were asked to put the faces back into their former positions. They were free to do so at their own speed and in free order. A counter below the square showed the number of wrong assignments. Every wrongly assigned card was put back to the draggables, right assignments stayed in place.

After all items were put back into place correctly, the test continued with a familiarity assessment interface. The familiarity assessment interface makes use of a comparative assessment tool created within this work (compare chapter 5.2.3). Therefore, the formerly shown faces had to be positioned relatively to each other, on a vertical from top (familiar) to bottom (unfamiliar). Cards could be placed freely on the screen via drag and drop, as shown in Figure 4-13.

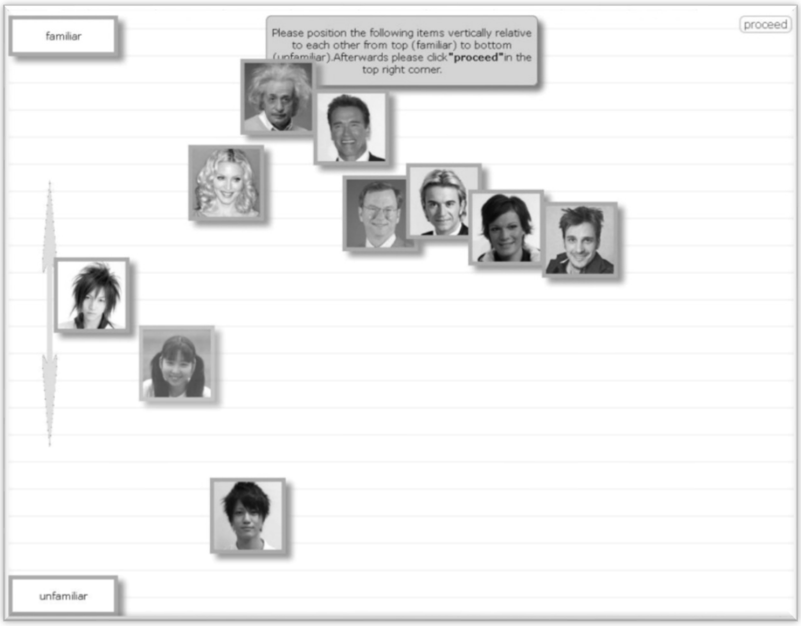


Figure 4-13. Drag and drop familiarity assessment on a vertical scale from bottom, for unfamiliar, to top, for familiar items.

Unrated cards were initially lined up on a horizontal line in the middle of the scale. In the example shown in Figure 4-13, “Einstein” was rated slightly more familiar than “Schwarzenegger”, followed by “Madonna”. One can also easily determine a group of familiar items and a clearly unknown person.

During the test, the procedures of memorizing and assessing had to be repeated three more times, with different faces and constellations. While the first memorization test was conducted on a white screen, the others showed faces from the same pool, placed on a map of Europe. The map should provide a reference frame, reducing the cognitive load of the user, to free more capacities for the actual test. The third run included only generally known persons, while the fourth run consisted only of strangers. Each subsequent assessment task only showed the faces of the previous memorization task, and not all that have been evaluated so far. If an item was already assessed previously, its position was restored, in order to make familiarity between single runs comparable.

Each memorization test ended with performance statistics. Test persons were shown their own score, and the general distribution of all other participants, for each particular test. They were also shown their personal history, if they conducted this test before (compare Figure 4-14).

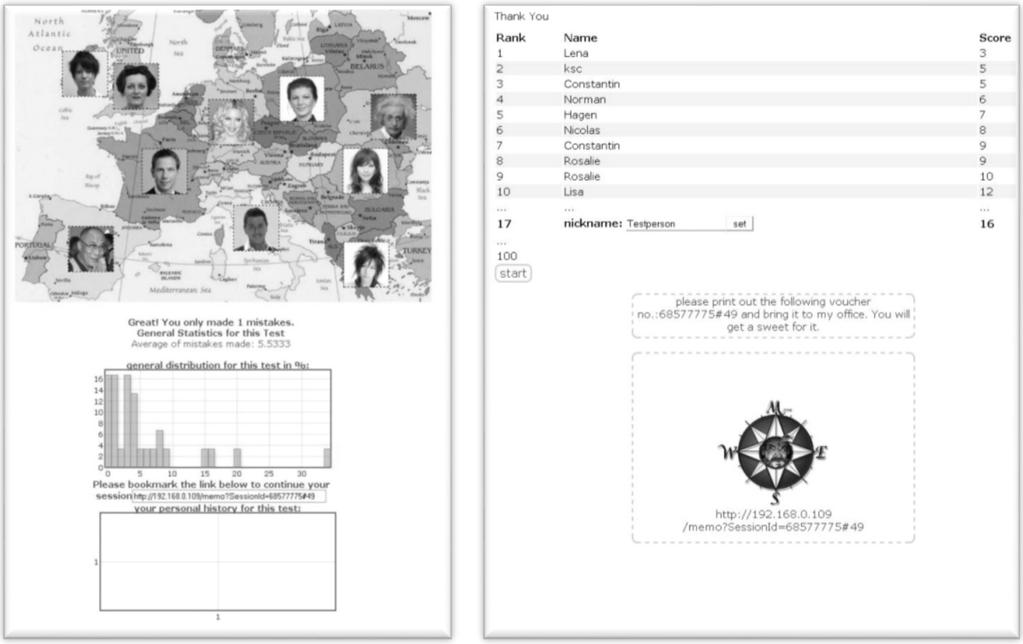


Figure 4-14. The second test with generally known and unknown faces including the statistics shown after each run (left). The final high-score and voucher for the users (right).

Further incentives were provided at the end of all four tests, such as a high-score and a candy voucher. Each voucher included a unique id, helping the author to allocate, otherwise anonymous, test runs to users. Voucher conversion also gave chance for a little informal interview.

Evaluation

The general feedback of the users was positive. The usage and tasks were understandable, and, for most users, even considered to be fun. Two participants reported initial problems with feedback for wrongly set items, which led them to drag the same card on the same position twice in a row. This flaw was unfortunate for the high score, but was easily filtered out in the evaluation results.

While in the first memory test (chapter 3.4.2) a distinct control group was necessary to validate results, the design of this test allowed users to be their own control group by assessing each item's familiarity. The test confirmed the tendency given by the initial memory evaluation. The familiarity assessment interface (as shown in Figure 4-13) allowed for a relative distinction of user familiarities. Within the familiarity assessment interface users were asked to vertically position test items according their familiarity, from top to bottom. This way, they created a reference frame of relative familiarity, with the most familiar face describing the top edge, and the most unfamiliar face placed at the bottom of the spectrum (compare Figure 4-15).

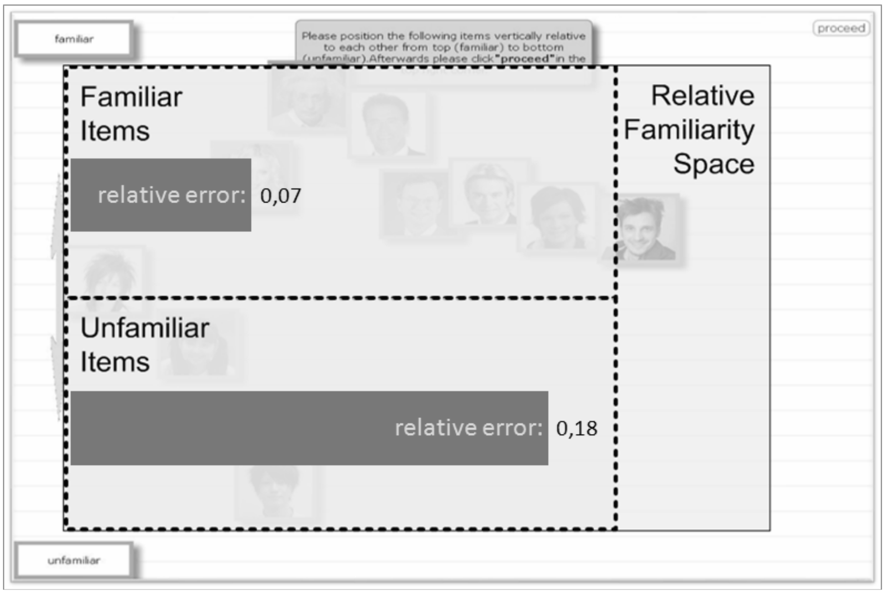


Figure 4-15. Comparative familiarity assessment of items in a memory test. Separated between familiar and unfamiliar (left).

For a separation between familiar and unfamiliar items, this frame can be split into half, as depict in Figure 4-15. Since already assessed items were restored in later assessment views, the four assessments could easily be merged into one, and became comparable. Finally the average error rate of all familiar rated items was compared to the one of the unfamiliar items.

In total 625 items have been rated in 21 tests. As indicated by the first test, users performed worse on unfamiliar items. With an average error rate of 18% wrong assignments, those items were 2.64 times more error-prone than familiar items (7% wrong assignments in average). An independent-samples t-test showed a clear statistically significant difference between the performance values of familiar and unfamiliar rated items with T(624)=-5,75, p=0.0016E-05.

A finer quantization into five blocks is shown in Figure 4-16. It details a more accurate correlation of familiarity and performance. Hereby good performance corresponds to a low error rate. For the further analysis, screen positions of familiarity ratings were transformed into number from 0, for the top rated familiarity, to 100 , the least familiar rated items.

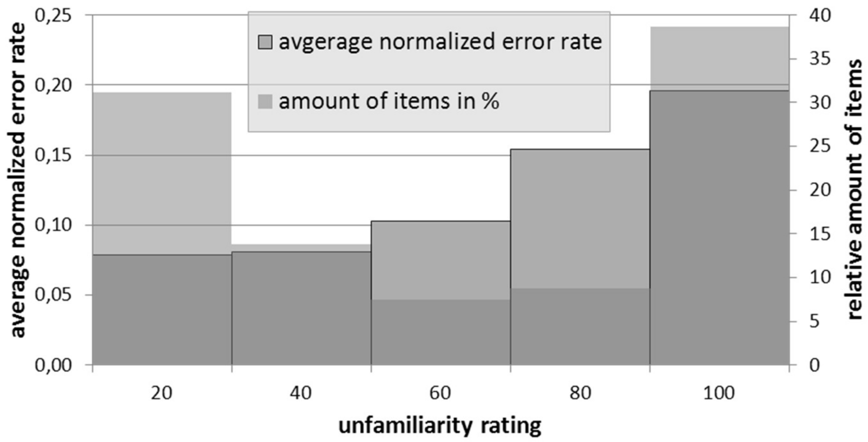


Figure 4-16. Average normalized error rate related to unfamiliarity.

Items, rated to be most familiar, reached a performance gain factor of 2.5, in average. Items rated between 0 and 40, led to best performances, while items rated 60 or worse corresponded dramatically with reduced performances. Almost 70% of the items were either rated to be highly familiar or highly unfamiliar; approximately 16% were rated between 40 and 80. The way of rating confirms the feedback given by test persons after the tests. Multiple participants reported that they only rated between familiar, unfamiliar and something in-between. The distribution of number of errors, among those three blocks, is displayed in Figure 4-17. Familiarity values rated better than 60 led to approximately 78% of right answers, compared to 53% for items with a familiarity above 60.

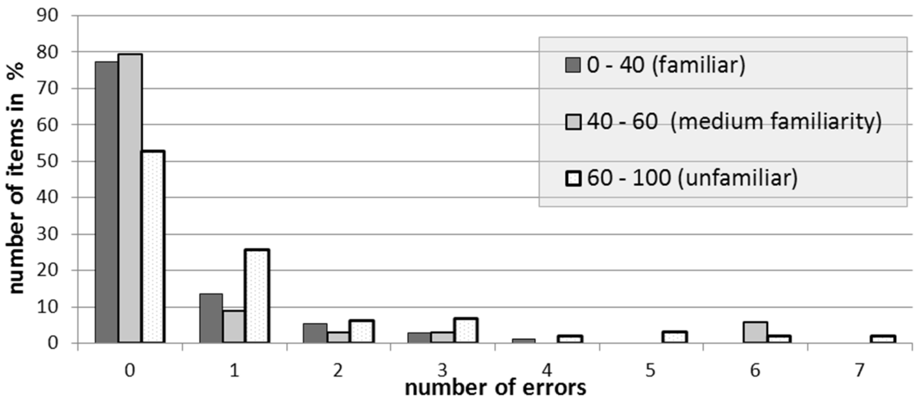


Figure 4-17. Error distribution of items rated with familiarity of 0-40, 40-60, and 60-100.

The diagram shows, how many consecutive tries were necessary in average, to make a right assignment. Hence, a long tail is an additional indicator for higher uncertainty. The tail for the 60-100s block is clearly longest. The decline of the 40-60 is steepest in the beginning, but gets a long tail through a peak at 6 errors. Such a peak cannot be found for items rated between 0-40, indicating that those items were set with most certainty, and arrangements usually not completely forgotten.

Test persons also tended to start the memory task with the most familiar items. This behavior is in line with findings of Rosch, where children learn typical instances of categories earlier than they learn atypical ones (Rosch 78). Further support is given by findings where people identify objects at a basic level more quickly than they identify objects at higher or lower levels (E. Rosch et al., 1976). The same mechanism seems to be fundamental in concept mapping, were salient features are mapped first (D. Gentner & Clement, 1988; Glucksberg & Keysar, 1990; Glucksberg et al., 1997; Ortony et al., 1985; Ortony, 1979). Figure 4-18 shows that the input order proportionally corresponds to error-rate and un-familiarity.

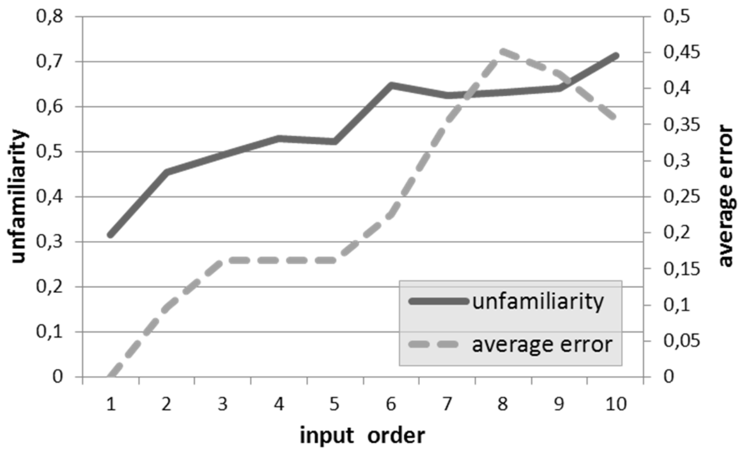


Figure 4-18. Familiar items were first set corresponding to lower error rates.

This perspective might also give an explanation for the peak at 6 errors for the 40s-60s block, shown in Figure 4-17). In consideration that familiar items are set first, items of this block were most likely assigned right after the familiar ones, and before the unfamiliar ones. At this time there were still many open options, leaving a higher chance for wrong assignments. The shorter tail for very familiar items, in Figure 4-17, also shows that only very little of such items were completely un-memorized, otherwise there would have been a peak beyond four errors.

Conclusion

The test gave further confirmation on the basic working principle of Digital Transformatives. It shows a clear correlation between user-rated familiarity and performance. In these tests, users were provided with a continuous comparative assessment spectrum. Although many users only felt confident with a differentiation between familiar and unfamiliar, the results suggest a more fine granular distinction. A quantization into five levels of granularity showed an exponential relation between familiarity rating and performance, reflecting the power law of practice (compare section *From controlled to automated processes to habituation* pp. 38). This test also provides further support for Feature 7 (*Most salient or familiar features of a prototype are mapped first*), indicated through a clear tendency to start the assignment and assessment task with the most familiar items.

It can be stated that users are able to rate their own familiarity beyond the binary granularity of unfamiliar and familiar. Further the own rating exponentially seems to map on performance. This test underlines findings on areas of improved performance and user rated cognitive prototypes as well as automatization (compare Heider, 1971a, 1971b, 1972; Kirkham et al., 2002; LaBerge & Samuels, 1974; McNamara & Kintsch, 1996; E. H. Rosch, 1973b; E. Rosch & Mervis, 1975; E. Rosch, 1975a, 1975c, 1978; Vicente & De Groot, 1990).

Hypothesis 4 can be transformed in Feature 14:

Feature 14. Users are able to accurately rate their own familiarity, which reflect their performances.

Consequently user familiarity ratings provide us with an accurate measure for determining areas of high user potentials.

Guideline 8. If environmental and performance measures are not available, let possible end-users rate familiarity. Familiarity ratings provide accurate measures for areas of high user potentials.

4.3 Systematic Methodology for Finding a Digital Transformative Context

In the previous section, assessment methodologies have been elaborated, helping with the evaluation of design ideas towards the working principle of Digital Transformatives (DTs). Such methodologies increase chances for selecting concepts from a pool of ideas, which are maximizing the use of cognitive and procedural areas of high efficiency. However, the methodologies are dependent on the quality and quantity of generated design ideas. This quality and quantity mainly depends on the designers' know-how and creativity. While Know-how may be acquired through experience, creativity is an unreliable factor. Therefore, some designers intuitively may find many useful design solutions, while others never fill the pool with DT relevant candidates. Hence, the investigations on the working principle of DTs provide an informative basis for creating systematic heuristics, which reduce the dependency on designer creativity and their unpredictable nature.

4.3.1 Prerequisites

In principle, Digital Transformatives (DTs) function analogous to cognitive categorization or metaphors in communication (compare *Cognitive Efficiency Catalysts in Communication* pp. 46). Therefore, re-engineering their working principle should provide the basis for a systematic methodology for finding proper DT contexts. An essential requirement for successful conceptual mappings is similarity between the base and the target concept. According to research on similarity measures of prototype categorization similarity of two concepts increases with a greater amount of shared salient features (compare Hampton, 1979, 1995; E. Rosch & Mervis, 1975; Tversky, 1977). As elaborated in chapter *Pervasiveness of Similarity Comparisons* (pp. 44), similarity comparisons are asymmetric from base to target concepts. Most salient features are mapped first (D. Gentner & Clement, 1988; Glucksberg & Keysar, 1990; Glucksberg et al., 1997; Ortony et al., 1985; Ortony, 1979; Tversky, 1977).

In order to transform such cognitive processes on interactive system design, a proper source concept representation is demanded. The source concept relates to the current state of a task, which are typically captured through questionnaires, interviews, focus groups, workshops, observation, or task analysis (compare chapter 4.1 *Design Methodology of Interactive Systems* pp. 131). The captured data is transformed into models for further interpretation. Hereby, the Hierarchical Task Analysis offers an action model representation, which can easily be transformed into concepts and features.

4.3.2 DT Design Challenge

The seed of most user-system-interactions is a set of source actions, induced by a task – initially triggered through interest. Due to former experiences, system designers are able to imagine such actions and choose adequate user interface (UI) elements. Hence, a given user action triggers cognitive concepts in the designer, which activate related UI-solutions.

As detailed in chapter 2.1, and in the section *Probabilistic Environmental Observation* (pp. 154), our mental structure of concepts is heavily influenced by the occurrence of extrinsic stimuli. Some concepts are better developed, and more efficient, than others. A visualization of such a cognitive concept structure is shown in Figure 4-20 (left). The 3D surface is defined through concepts holding two attributes; efficiency is mapped orthogonally[17]. Peaks mark cognitive prototypes of relatively high performance. The closer two peaks together, the more similar their concepts.

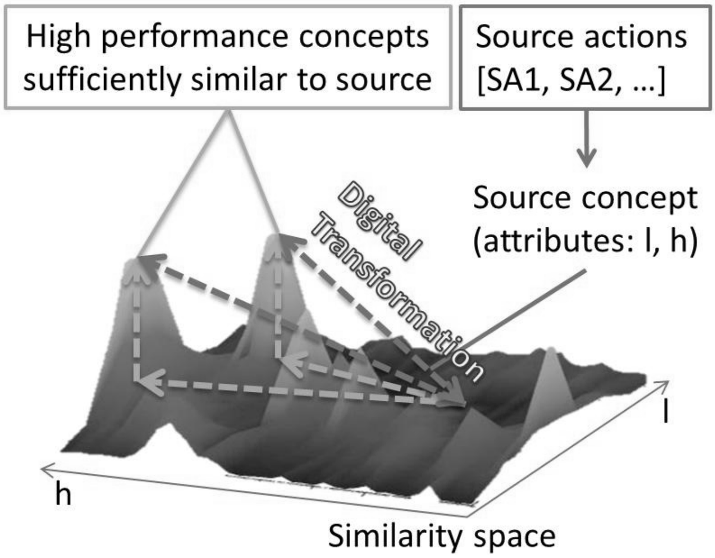


Figure 4-19. Visualization of the cognitive design principle of Digital Transformatives.

While common user interfaces address concepts directly related to demanded actions, Digital Transformatives rather implement interfaces on highest performance concepts, within a certain similarity range to the source actions. Hereby, similarity is necessary to allow for a proper digital transformation.

Guideline 9. Find super salient cognitive concepts in sufficient similarity proximity to original function related concepts.

Coming back to the designers' thoughts activated during the design process: the challenge during the DT design process of user interfaces lies in determining the closest high performance prototype concepts, beyond obvious prototypes. It is challenging, because our thoughts always tend to flow to cognitive prototypes contextually situated in close similarity to our problem base.

For example, let us consider the design of a new innovative craftsman coffee cup; a cup for drinking the morning coffee and for sinking a nail, if no hammer is at hand. A major system interface property concerns the possibility to hold this cup. The reader may try to think of the right handle for such a cup, at this point.

[17] The distribution is borrowed from the world color survey, with the two attributes hue and lightness

At first glance, a cup without any handle offers a first solution. It can be grabbed like a bowl for drinking, and used like a stone for hammering. One might also immediately think of a straight handle, as we know it from hammers, or a D-shaped handle as it is known from typical cups. The three possible solutions are shown in Figure 4-20 (top-right).

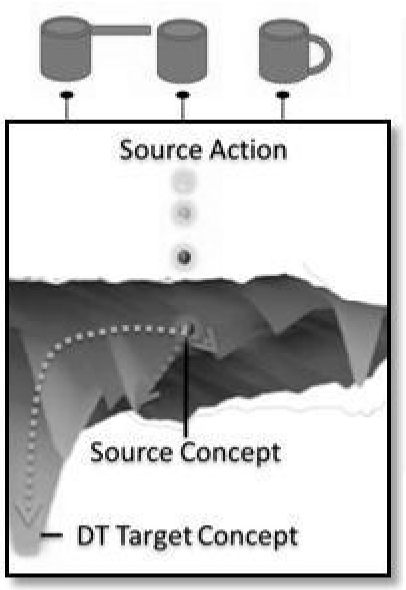


Figure 4-20. A visualization of the challenge of finding similar high performance cognitive concepts similar to the source concept.

While those solutions immediately come to our mind, it is hard to get beyond them. What other solution seem useful? We tend to think of the obvious (similar) concepts (prototypes), and we easily get stuck with those apparent ideas.

If we imagine the concepts in our brain as an upside-down Alpine scenery, as shown in Figure 4-20 (bottom-right), Dents mark areas of lowest effort. Our source action is a ball, dropped onto this surface. The drop down point determines the activated source concept. Once on the surface, the ball starts rolling downhill until it reaches a balanced state, caught in a local dent. In the same way the ball is attracted by adjacent dents, we inevitable think of cognitive prototypes similar to activated source concepts. Related experiences are observed in experiments on the recognition and memorization of ambiguous images (compare Figure 2-10 on p. 29). For example, in (Chambers & Reisberg, 1985) most test persons had problems seeing another interpretation, after a first one was identified.

The utilization of similar cognitive prototypes helps designers for finding proper interface solutions efficiently. Conversely, those solutions are found quickly, but other, much more salient concepts, in similar user action contexts, may lie beyond the apparent ones. Digital Transformatives aim for those concepts. While looking for those design solutions increases the one time design effort of a few developers, it also holds the potential to decrease the daily usage effort of legions of users.

Current user interface design methodologies offer no satisfying solution for the previously described DT design challenge. Some creativity techniques, such as Brainstorming can be used to create a big pool of design solutions, also including possible DT candidates. However, due to the random nature of the revealed solutions, it is very likely that good candidates are overseen in the evaluation process among many stronger, improper, possibly completely unrelated concepts. Thus, in the following, two approaches will be introduced. On the one hand, the challenge will be addressed from a cognitive side by introducing the *Salient Super Prototype Identification Approach*. On the other hand the *Sub-Action Modeling Approach* extends existing system design methodologies towards a methodology for systematically determining DT relevant design solutions.

4.3.3 Salient Super Prototype Identification Approach

The salient super prototype detection approach is mainly based on findings described in section 2.1 *Relevant Cognitive Mechanisms* (pp. 14). Starting from an evident understanding of our cognitive structures, a methodology will be developed for solving the previously sketched DT Design Challenge.

SIMILARITY BASED COGNITIVE STRUCTURES OF PROTOTYPE CATEGORIES

Multiple studies give evidence that our cognitive concepts are predominantly structured through similarity. Similarity comparisons in perception and recognition are inborn fundamentals, found throughout many cognitive mechanisms, such as recognition, abstraction, or categorization (D. Gentner & Christie, 2008; D. Gentner, 2003; Penn et al., 2008). We compare and memorize perceived information in reference to cognitive prototypes (Chambers & Reisberg, 1992; Franks & Bransford, 1971; Neumann, 1977; Peterson et al., 1992; Posner et al., 1967; Posner & Keele, 1967; Reed, 1972; Solso & McCarthy, 1981). As empirically demonstrated throughout various experiments, we prototypically categorize as we recognize objects. We do not store exact representations, but more efficiently, build on existing knowledge, and gradually refine its structure (Chambers & Reisberg, 1985; Chase & Simon, 1973; P. Kay et al., 2009; Neumann, 1977; Posner & Keele, 1967; Reed, 1972; Solso & McCarthy, 1981; Vicente & De Groot, 1990 and many more). Several studies underline the improved efficiency at prototype categories (Heider, 1971a, 1971b, 1972; E. H. Rosch, 1973b; E. Rosch & Mervis, 1975; E. Rosch, 1975a, 1975c, 1978).

The previous visualization, shown in Figure 4-20, reflects the structure found throughout our cognition, based on similarity and prototypability. Starting from concrete actions, we need to determine best prototypes with close similarity to the original tasks. Our knowledge is structured by similarity, and it is nested, as research on categories, pattern recognition, or chunking shows.

Natural prototype categories form around some kind of averages of a class of objects, integrating most typical features. Such salient category features develop based on occurrences in our perceived environment (P. Kay & Regier, 2003; Paul Kay & Cook, 2011; Neumann, 1977; Posner & Keele, 1967; Reed, 1972; E. Rosch & Mervis, 1975; Sternberg, 2008). This way prototypes may also form around probabilistic averages, resembling objects which do not exist outside cognition (compare (Solso & McCarthy, 1981)). We prefer cognitive operations on so called basic levels, which are situated between the most concrete and most abstract known concept.

To get a feeling for basic levels, one may simply think of a random object. What did you think of? We likely do not have a concrete item in front of our inner eye, nor do we imagine something completely abstract. Further, the reader may now think of a car; now think of a sports car; and finally think of some type of a Lamborghini sports car. It is likely that one did not think of the same object every time, although the previous terms are referring to commonly nested category prototypes. In all cases one could have thought of a Lamborghini, which is a sports car, a car, and an object. And one probably never thought of a very concrete object.

Basic levels change with context and expertise. The term car is commonly known, while knowing a certain Lamborghini type demands advanced expertise. Similarity comparisons also allow us to efficiently build on existing knowledge. By defining a Lamborghini as a sports car, it already inherits many salient features, including those of cars and objects.

Chunking plays a key role for increasing performance in automatization processes. Chunking is also a major element of prototype categorization, and its effects have also been researched in the area of Gestalt psychology (see chapter 2.1). As demonstrated in the last paragraph, prototypes are nested depending on the area of attention. Being able to recognize objects as a whole, or to change the focus of attention to sub-parts, is important for automatization processes, as well as for prototype categorization. Recognizing chunks at different levels of complexity is a key performance driver, as empirically proofed in multiple experiments (G. D. Bower, 2008; G. H. Bower, 1970, 1972; Chase & Simon, 1973; De Groot, 1978; Gobet & Simon, 1996a; Larkin et al., 1980; Lesgold, 1988; Reitman, 1976; Samuels et al., 1978; Vicente & De Groot, 1990).

Cognitive concepts are nested, and prototypes are available at various levels of complexity, practically experienced in our hierarchical understanding of categories. The underlying basic characteristic for organizing knowledge is similarity. Figure 4-21 illustrates how various concepts may be structured according to the previously summarized evidences on cognition. Similarity is referenced in the horizontal dimension, and typicality is mapped vertically. Additionally, different levels of complexity are visualized through zoomed image sections.

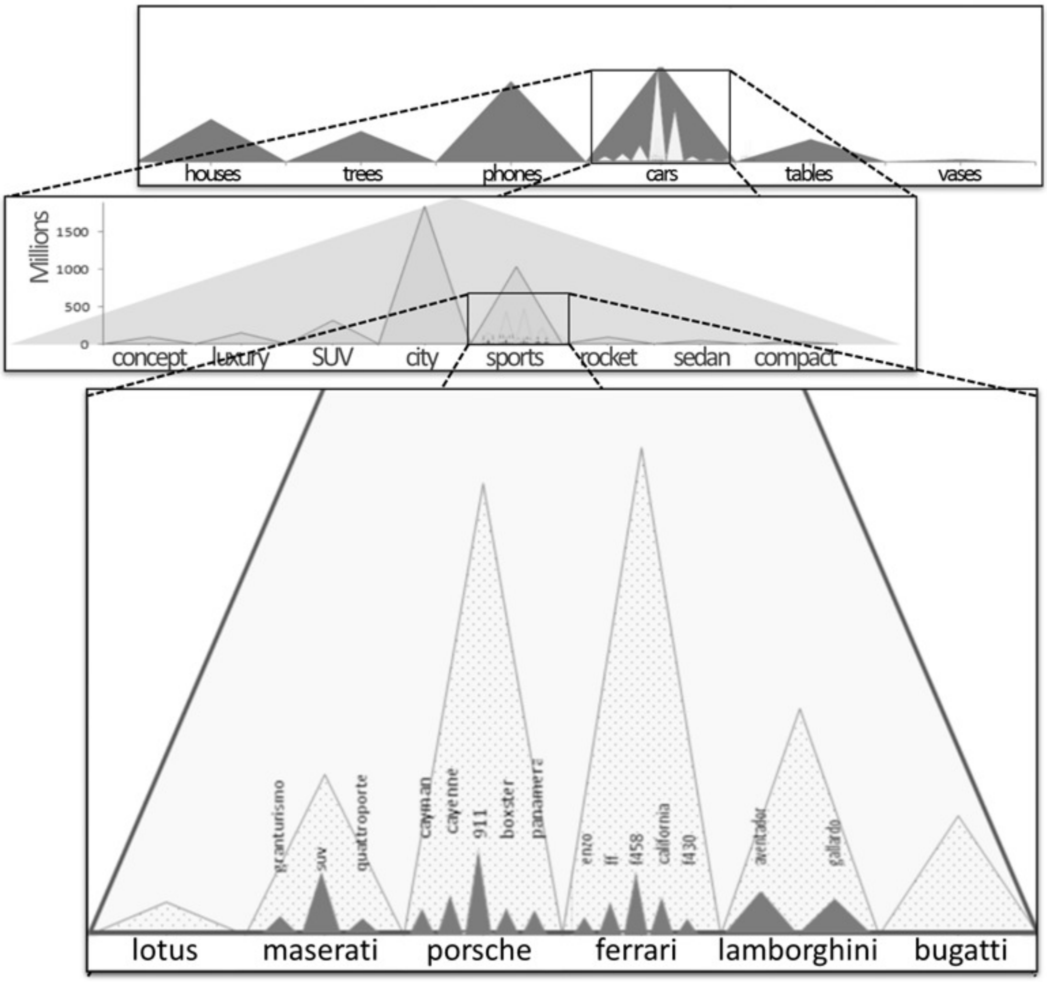


Figure 4-21. Accumulated English search queries at google.com. The query frequencies may represent a model for shared cognitive concepts[18].

For example, a Porsche 911, and a Boxter are similar to each other, sharing salient features typical for a Porsche car. Furthermore, a Porsche and a Ferrari are similar, sharing salient features of sport cars, which share features of cars, and so on. The

[18] Extracted from suggestqueries.google.com – March 2013

other way around, the term “car” might predominantly make us think of the prototype concept of a city or a sports car, and the prototype concept typically associated with the term “sports car” may be very similar to a Porsche or Ferrari.

While the above example depicts a rather common categorization, this organization exists beyond conventional taxonomies, in all kind of patterns of procedural and declarative knowledge. Cognitive prototype categories are vaguely defined in various dimensions and may even be changing dynamically over time. This aspect is well exemplified by tests on the recognition of shapes with Dani (E. H. Rosch, 1973b; Ungerer & Schmid, 2006).

User Determined Prototype Concepts

The previously elaborated understanding of our minds similarity based cognitive structure allows us for the systematic determination of demanded concepts. Since a new DT system should be addressing a shared end-user concept model, it is best to capture cognitive concepts directly from the users by asking them for concepts similar to the source concept.

Guideline 10. Capture information directly from possible end-users. System design concepts should be based on a shared cognitive model of possible end-users.

Due to the nature of prototypes, salient concepts will be named first and most frequent (E. Rosch et al., 1976). According to the cognitive structure shown in Figure 4-21, if we ask for cars similar to a Maserati, Ferrari, and Porsche are likely to be named first and most often. As the figure illustrates, concepts are structured based on categories, which are concepts and part of super categories themselves. Hereby, Ferrari and Porsche are considered to be concepts similar to Maserati. Hence, they are part of the same category, sharing the general features of a sports-car.

As research on prototype categories shows, alternatively to asking for similar concepts to the source concept, we could also ask for examples of the super category. Hence, if we were asking for examples of a “sports car” most quickly and often named terms should also be Ferrari and Porsche. Super prototype concepts, such as “sports car” in this case, hold most salient features shared by its members in average. Inversely most salient sub-concepts shape the understanding of their super concept.

Guideline 11. Traverse superordinate concepts to find cognitive prototypes of close similarity.

The biggest challenge lies in determining the super-category term. Fortunately, knowledge about categorization also provides us with a technique to request super concepts from a source concept. Super concepts are defined through most salient features of a class. Hereby, they describe the most distinctive features shared by all sub concepts. Consequently, one first needs to collect concepts similar to a source concept, and then use those to determine a super concept which comprises all sub concepts. This methodology for determining super concepts, or concept category terms, further helps us with finding super salient prototypes of close similarity.

From research on basic levels we know that most salient prototypes may not even lie on the most concrete level.

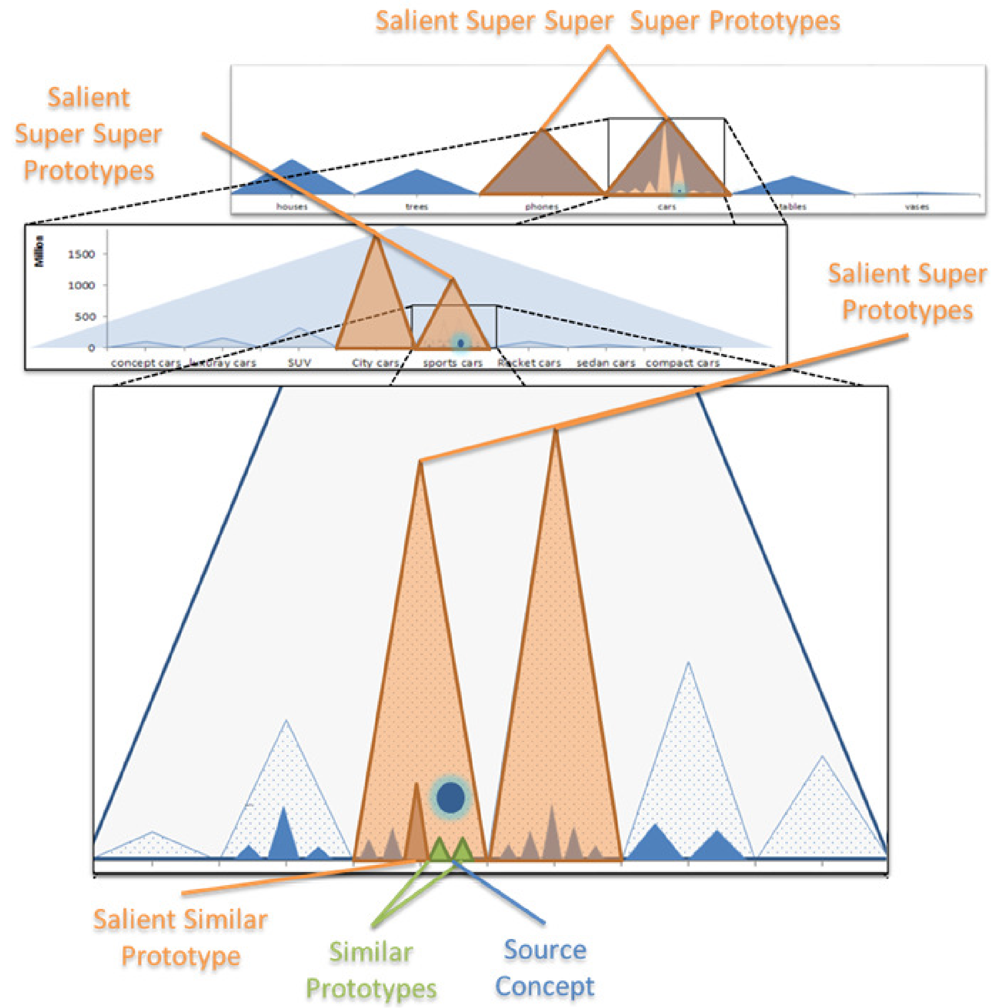


Figure 4-22. Visualization of relation between source concept (marked as a blue glowing sphere) and demanded salient prototypes and super prototypes.

There might be much more salient prototypes in further super levels. In order to traverse higher order super prototype categories, one could iteratively proceed with the previously described procedure to determine super prototypes. Each iteration gives access to new salient similar prototypes, as depict in Figure 4-22.

The visualization also shows that every iteration also broadens the base of including sub concepts, this way increases dissimilarity. This is obvious, if we consider that similarity features are inherited from super concepts to sub concepts – although in reality this is not exactly the case. Hence, super concepts generalize the context gradually, as the zoomed views in the figure illustrate. Finally, if necessary, every super prototype category can also be used to iteratively determine concrete sub concepts, as shown in Figure 4-23. Again, the most salient concepts of each category will be named first and most often.

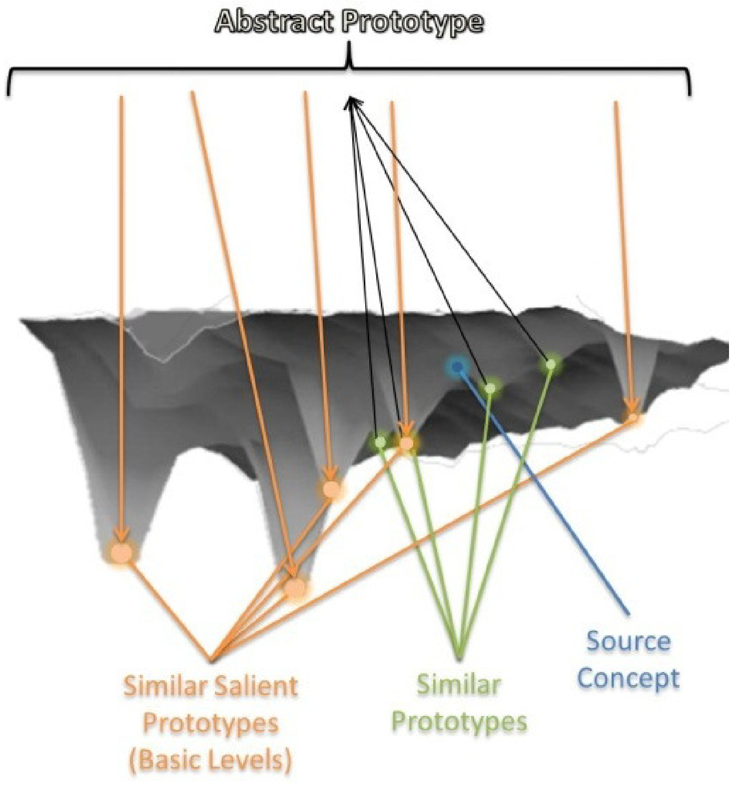


Figure 4-23. Visualization of determining super salient prototypes.

Consequently, this gives us the methods to span a salience-similarity space of concepts relative to a source concept. Concepts are requested directly from the users. Salience is determined through immediacy, efficiency, or evaluation of familiarity. Similarity is gradually extended by requesting super concepts, as described above, and tentatively dig deeper to gain salient concrete concepts.

SALIENT SUPER PROTOTYPE IDENTIFICATION PROCEDURE

The above description will be concluded in a procedural rule set for designing DTs, as follows.

Requisites:

a) Design should be based on a shared concept model of possible end-users. Therefore it is necessary to capture information from possible end-users.
b) According to the DT Design Challenge expressed earlier, we search for highly salient prototypes, with sufficient similarity to the source concept for a digital transformation.

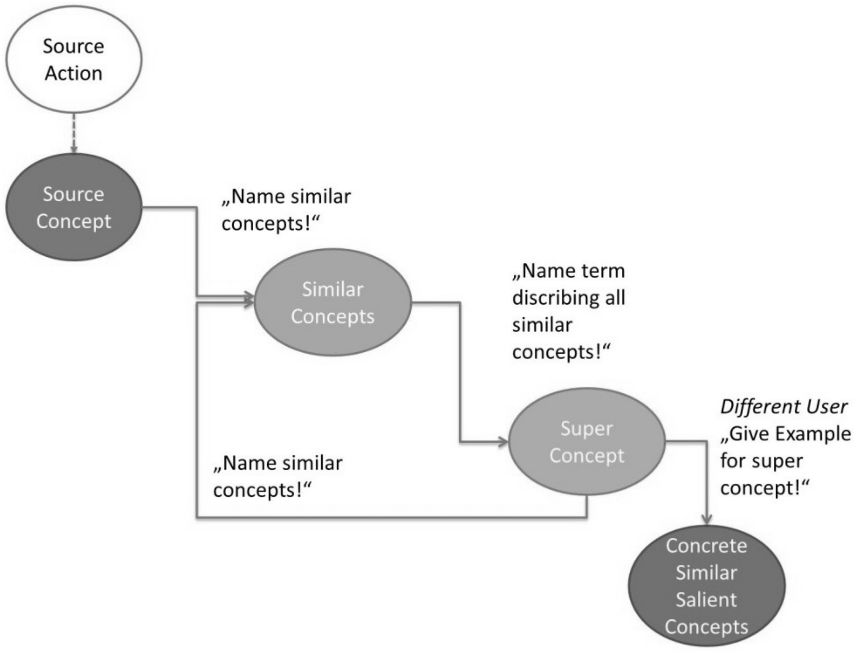


Figure 4-24. Schematic procedural overview of the salient super prototype identification approach.

Salient super prototype procedure (compare Figure 4-24):

1. Iterative determination of salient super prototypes. Initially starting from source concept, iteratively proceeding with super concepts.
    a. Ask end-users for similar concepts,
    b. Get super concept for requested similar concepts by asking for a term or concept describing all of the requested similar concepts,
    c. Iteratively ask users for typical examples of each super concept, to scan for more concrete salient prototypes

Ideally each step should be performed without knowledge about the previous step. Consequently they should not be conducted consecutively. Consecutive steps can be performed by distinct user groups. In any case, sort results by frequency or immediacy, because most immediately determined concepts ought to be most salient and efficient.

Finally one should get numerous relevant prototypes with differing salience and similarity. Those may be sorted through end-users, for further evaluation of familiarity and efficiency, as described in chapter 4.2.1.

The procedure applied on the above example could have the following results:

Starting concept is Maserati

1. Iteration (increasing salience, dissimilarity and generalization)
    a. Similar Concepts to Maserati: Lamborghini, Bugatti, Ferrari, Lotus
    b. Super term comprising all concepts: “sports car”
    c. Ask other users for examples of “sports cars”: Ferrari, Porsche, Lamborghini
2. Iteration (increasing dissimilarity and generalization)
    a. Similar Concepts to sports cars: rocket cars, concept cars, luxury cars, racing cars
    b. Super term comprising all concepts: “cars”
    c. Ask other users for examples of “cars”: City Cars, Sports Cars, SUVs
        i. Further iteration, ask for examples for City Cars: Toyota Camry, Ford Focus, Mercedes Benz (decreasing salience and generalization)
3. Iteration (increasing dissimilarity and generalization)
    a. Similar Concepts to cars: Bikes, Planes, Trains …

The concepts revealed in step c), for every iteration offer relevant concepts. The super concepts determined in step a) may also provide valuable concepts. As the example practically demonstrates: with every iteration dissimilarity and generalization increases. In this case, this also correlates with increased salience. On the other hand, every iteration within c) reduces salience. Indirect measures for salience are frequency, and immediacy through response times. Additionally, users can be asked to sort items by familiarity for giving further salience indices.

The result of this procedure provides system designers with potential prototype concepts, featured with salience values and similarity to the source concept – ready to start further design processes.

4.3.4 Sub-Action Modeling Approach

For the Sub-Action Modeling Approach, every action is successively divided into sub-actions, which can be seen as the features of the main task, in the same way features actually are defining sub-concepts of a main concept (compare JS Bruner et al., 1956; J. Fodor, 1994; Hampton, 1997b; Kruschke, 2003; Love, 2003; Sternberg, 2008). This way, the similarity measure of concepts can be mapped onto action models with sub-actions. Consequently, it will be assumed that the similarity of two actions is higher with a greater amount of shared sub-actions. A sub-sub action is equivalent to a feature of a feature of a concept, and so on. This assumption allows us to determine similar actions in the same way one could determine similar concepts, simply by looking for actions which share sub-actions.

Guideline 12. Ask for super salient neighbors which share salient sub actions (features) with the functional action (source concept).

The above arguments are brought into the context of Digital Transformative design as illustrated in Figure 4-25.

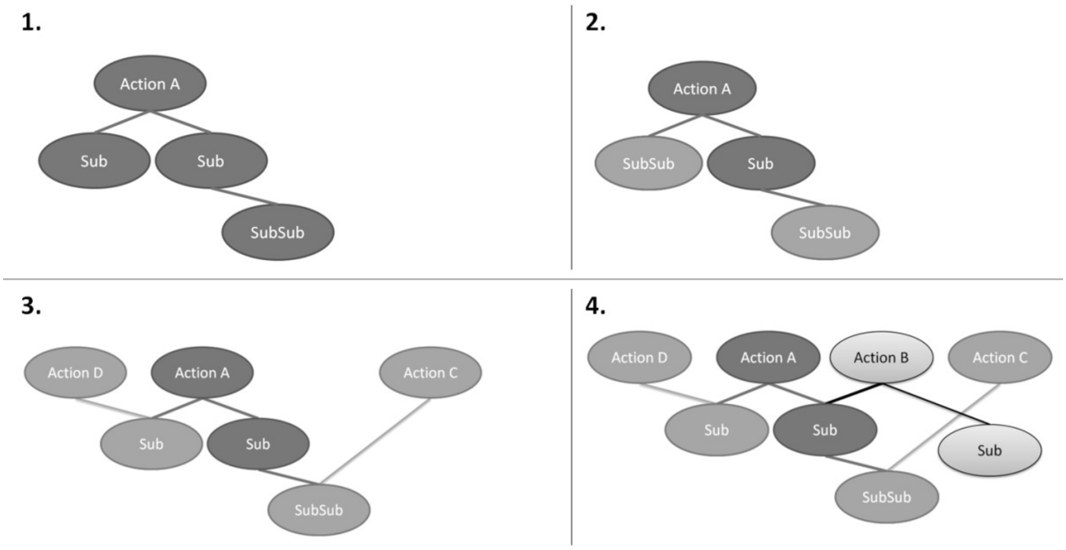


Figure 4-25. Sub-Action Modeling Approach. Four steps to find target contexts of high familiarity and sufficient similarity to the source context.

Just as any other interactive system, Digital Transformatives are created with the goal of supporting certain user actions or tasks, such as writing a document, giving a presentation, or navigating through unknown terrain. In Figure 4-25 this overall system goal is called Action A. The initial requirements analysis usually also models the current state. For the design of Digital Transformative, such a model ideally is represented through a hierarchical action model, successively describing the action through cascaded sub-actions as shown Figure 4-25-1. As elicit previously, cascaded action models behave analogous to cognitive concept models; in this context sub-actions are seen as features. This allows us to use the cognitive concept similarity measure for determining action similarity. Consequently, similarity of two actions is increasing with the amount and salience of sub-actions they are sharing. In the example given in Figure 4-25-3 Action D and C are similar to Action A, simply because they share sub-actions. Additionally, it can be assumed that Action D is more similar to Action A than to Action C, since it shares a direct sub-action, while Action C only shares a sub-action of a sub-action. If two actions share a first level sub-action they likely share a higher amount of sub-actions than actions, which share sub-actions at a deeper level. However, from research on cognitive concepts we know that, besides the amount of shared features, the salience of features also is important as a weighting factor for similarity.

Such insights show the potential of using sub-tasks for finding similar actions, as it is demanded for the Digital Transformative context. Additionally, the Digital Transformative user interface needs to reside in a user familiar context.

As emphasized in the previous section, it is important to involve possible end-users in the process described here. Therefore, end-users could be provided with sub-actions of the source action, given the task of naming Actions, which are including such sub-actions. Those actions would be similar to the source action, and obviously familiar to the user. This way they meet both major requirements of being familiar and similar. Additionally unfamiliar actions, such as Action B shown in Figure 4-25-4, would not be named.

Finally, one could assess the determined action contexts through an additional familiarity assessment as conducted in 4.2.1. For creativity reasons it might be advantageous to provide involved users with sub-actions out of context.

SUB-ACTION MODELING PROCEDURE

The procedure as elaborated above, and illustrated in Figure 4-25 can be summarized as follows

1. Starting from a source action (Action A), a hierarchical sub-action graph is modeled (if end-users are participating in the modeling process they should not be used for finding similar actions).

2. A group of end-users is presented only with sub actions of the initial analysis. The sub actions should be presented out of context of the source action. One should be aware of the level of cascading depth, since direct sub-actions might lead to very similar new actions, while actions of a deeper level could result in actions which are too different.
3. From the presented sub actions, end-users reversely model supper actions. They are simply asked to name actions which include the given sub actions. The sequence of named actions should be recorded, since it gives a first indicator for familiarity of this action. Familiar actions are named first.
4. Unfamiliar actions will not be modeled

4.4 Summary - Overall Digital Transformatives Design Process

In this chapter an adaption of the common interactive design process has been elaborated, fostering the design of Digital Transformatives. A schema of the changed procedure is shown in Figure 4-26.

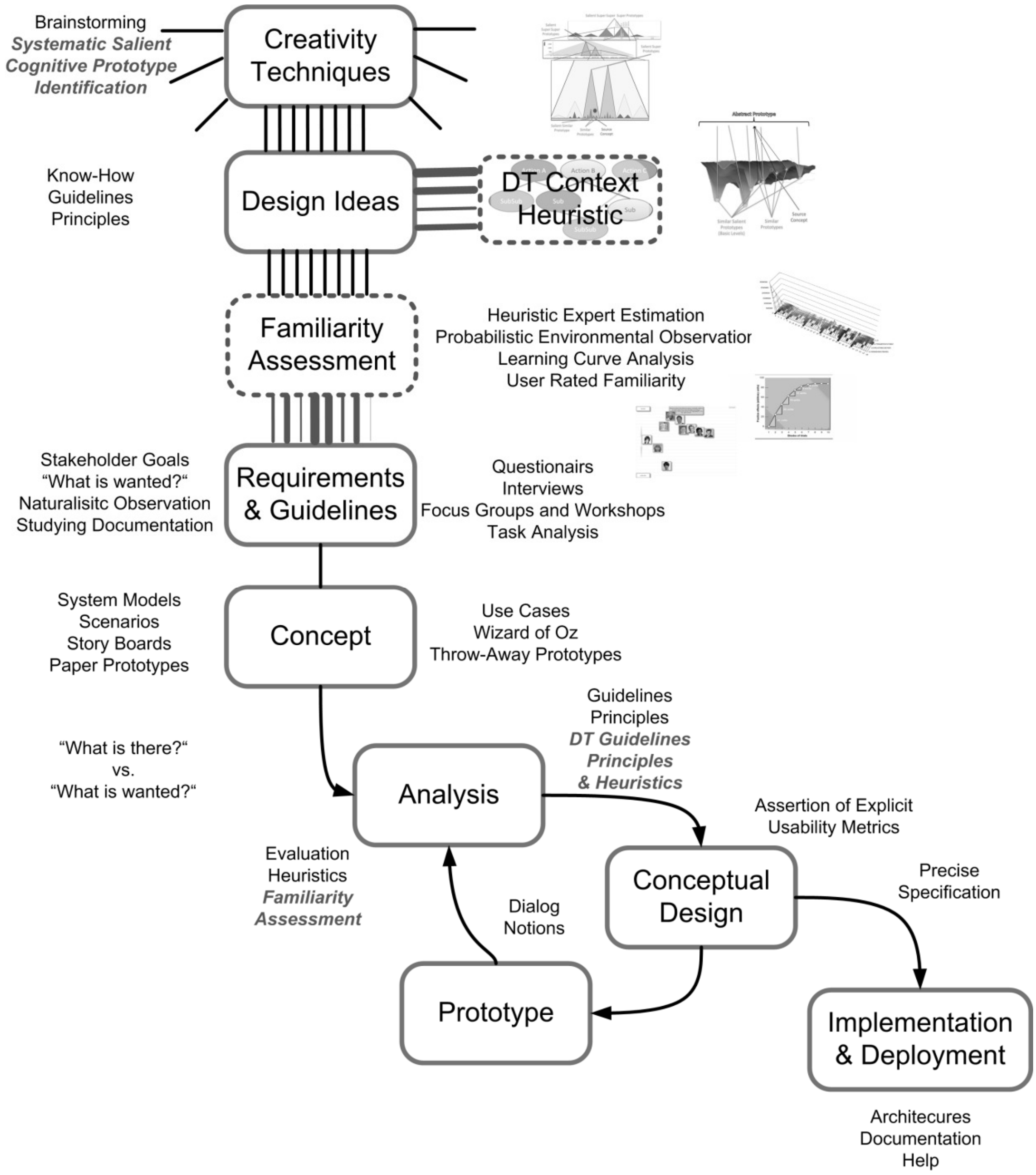


Figure 4-26. Schema of a possible Digital Transformatives design process.

The minimal change for increasing the chance of finding DT contexts is given by extending the requirements analysis of design ideas by sufficient familiarity assessments. Possible assessments methods have been proposed, ranging from inaccurate heuristic expert estimation over probabilistic environmental observation and learning curve analysis to most accurate user rated familiarity.

It has further been argued that such familiarity assessments still highly depend on the creativity and empirical potential of the system designer. In order to reduce designer specific dependencies of the design process a systematic methodology has been developed that potentially allows for determining proper DT user interface contexts. The development was based on the concept of Digital Transformatives and underlying cognitive principles, providing concept ideas which offer maximum user familiarity and sufficient similarity to the source context. The method is described in the previous section and should result in a set of DT design ideas, weighted by efficiency. Besides the elements influencing major steps in an interactive system design process, other elements need to be slightly extended through DT Design principles and familiarity assessments.

5 ITERATIVE USE CASE PROTOTYPES 189

5.1 GENERATION 1 – CONCEPT DESIGNS AND SCIENCE CENTER TO GO 190
5.1.1 Science Center To Go 193

5.2 GENERATION 2 – FIRST APPROACHES BASED ON FAMILIARITY 206
5.2.1 Captology and Persuasive Technologies 206
5.2.2 Drinking Garden – a Digital Beverage Coaster 210
5.2.3 Assess the Assessable 218

5.3 GENERATION 3 – COOPERATIVE COGNITIVE PROTOTYPING TOOLS 239
5.3.1 DT Web Module Sub-Action Modelling 240
5.3.2 DT Web Module Salient Super Prototype Modeling 242
5.3.3 Reading Speed Test 247
5.3.4 Conclusion – How to Create Tools for Capturing Cognitive Concept Structures 250

5 Iterative Use Case Prototypes

The methodological design was accompanied by steady use case and prototype development of Digital Transformatives (DTs). Some of the designed prototypes, which influenced the overall research process, will be briefly described in this section. Figure 5-1 illustrates how those implementations integrate into the overall iterative research methodology described in chapter 1.1. The design can be subdivided in three main phases (generations).

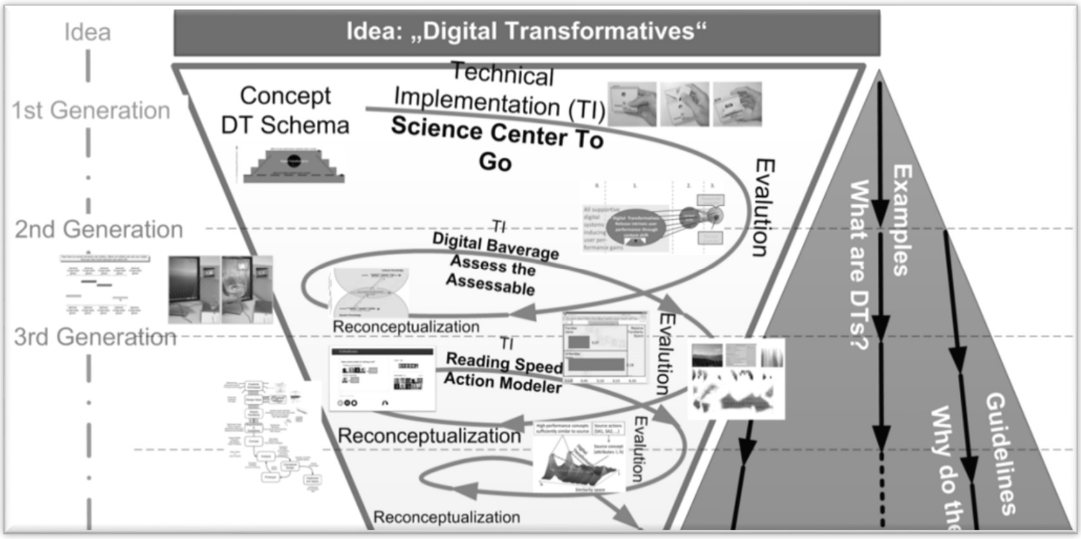


Figure 5-1. Illustration of the implementation of the overall research methodology.

In the first phase the DT defining schema existed, but the working principle was not elaborated. This phase mainly includes conceptual designs, clarifying the basic DT idea and conducting vague tests. It also comprises the development of the Science Center To Go, which is the most comprehensive system of all prototypes. The second phase started after the major working principle of DTs was revealed. Prototype implementations of that phase already aim for increased user familiarity and context similarity, such as an implementation for comparative assessment of items.

Additionally, the concept development in this phase showed the importance of cognitive automatization processes for Digital Transformatives, which directs focus on the research field of Persuasive Technologies. From a software design perspective, the concept of Persuasive Technologies in general is complementary to the concept of

Digital Transformatives, however several attributes are similar. A prototype was implemented which combines both approaches.

The third phase was influenced by the development of a systematic design methodology. The methodology has been tested to a small extent. The results of those tests led to a concept and implementation of a DT presentation tool. Moreover, tools for cooperatively supporting the DT design process have been implemented here, such as the Action Aodeler, or the Speed Reading Game.

The main prototypes of all three implementation phases will be shortly described, and set into relation to the DT concept. Finally the DT Framework will be briefly described, which is meant to offer a concrete online platform for further future development and refinement of this work.

5.1 Generation 1 – Concept Designs and Science Center To Go

The first concept showcase addresses the touristic sector. Tourists visiting a foreign city are usually unfamiliar with their new environment. On the one hand, it is good to be unfamiliar because only unknown spaces can be discovered. On the other hand, unfamiliarity might also be stressful, and makes tourists feel unconfident. In this case a Digital Transformative could be designed that keeps the original task of discovering a new place, but let's tourists still feel familiar with their environment. This goal might be reached by transforming the new city context into a known city context. Therefore, an Augmented Reality (AR) system could integrate landmarks of a familiar city environment into a new cityscape. This transformative concept is referred to the "One World One City" (OWOC) application.

For example, tourists from Cologne are visiting Chicago the first time. They have an hour transit at Chicago Union Station. If they knew the area as good as they knew Cologne, they would probably stroll around, and go shopping for the meanwhile. Since they do not know the area, they feel uncertain if they could walk around freely, which might bring them to far away from the train station, and in the worst case lets them miss the train. OWOC tries to solve this problem by offering a mapping of their home town onto the new place. The tourists map both cities by defining that cologne cathedral were located at Chicago Union Station. Their mobile phones, or AR goggles, offer an alternative Mixed Reality view, overlapping landmarks of both cities as illustrated in Figure 5-2.

Figure 5-2. Greetings from Chilogne: Illustration of the One World One City Digital Transformative design concept.

This way, the known cityscape of cologne provides familiar orientation anchors, embedded into the new Chicago environment. Chicago looks like Cologne with new facades.

Conventional systems, however, build on digital maps which evolved from classical maps. Geographical maps typically give a schematic overview, depicting the relation of space between elements of an area, such as streets, buildings, or districts. Although maps surely give important support, their usage demands the effort of determining the current location and orientation on the map. Digital maps reduce this user effort, by providing search functions and automatic localization, based on the users' current location. All those reductions are generated through algorithms; however, they hardly exploit the users' potential.

In our daily environment, we permanently orient ourselves, relatively effortless, without any maps. This is possible because we are familiar with most salient occurrences, and know how those relate to each other in space. Thus, landmarks of familiar environments form cognitive prototypes as reference points for our daily orientation. This assumption leads to a Digital Transformative concept, as shown in Figure 5-3.

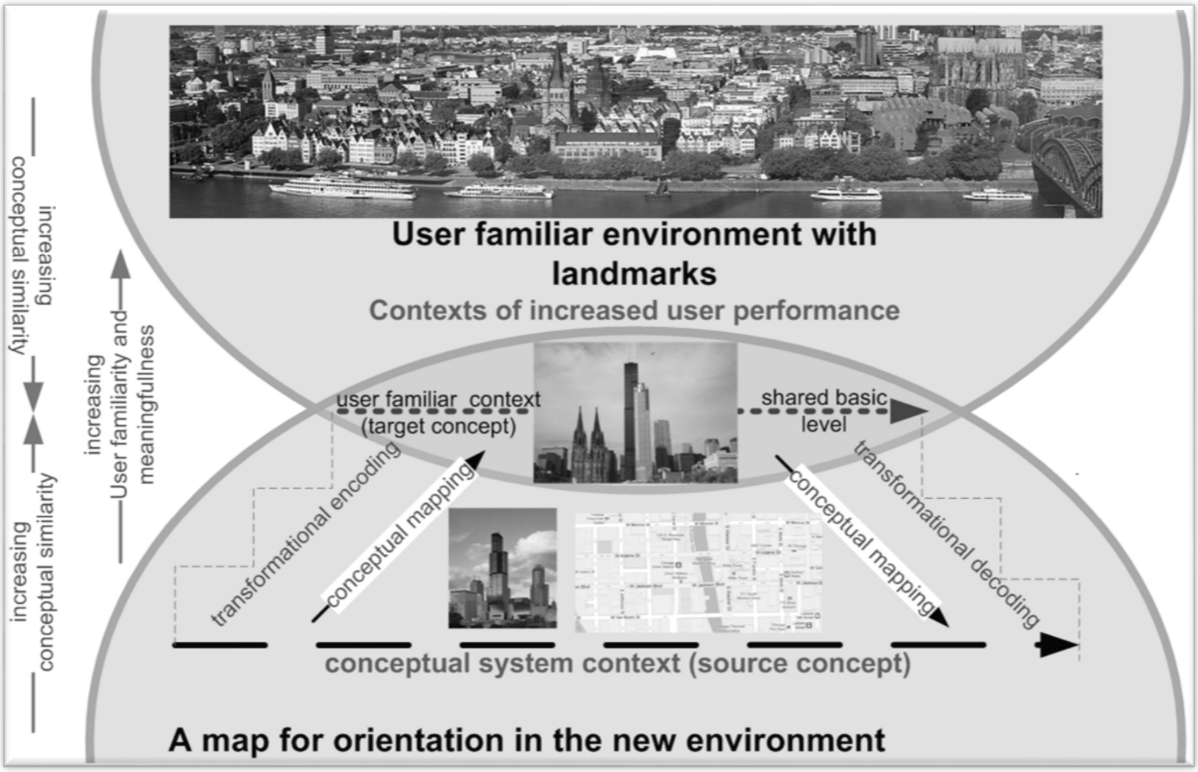


Figure 5-3. Digital Transformative concept for navigation through familiar landmarks.

Instead of providing a digitized map, the source context is transformed into a target context of increased user familiarity, providing landmark based references, embedded in their current environment. Landmarks need to be remarkable occurrences of a user familiar environment, such as known buildings of their home town. The new target context should be sufficiently similar, to allow for proper transformational encoding and decoding. Obviously the artificially embedded landmarks need to be mapped accurately with the right scaling.

In a similar way, car navigation systems may be modified. Standard navigation systems provide computer guidance by determining the shortest way to a certain destination. The suggested way is displayed on a two-dimensional top view showing a map, or a three-dimensional view from the drivers' perspective. The driving assistance is also often supplemented through audio driving directions such as, "*in 300 meters, turn right onto Whatever St*".

The Digital Transformative concept elaborated here, however, suggests addressing the environmental cognitive prototypes, as navigation references. Consequently, it would be more efficient to give directions like, "*turn right at the upcoming Shell gas* *station*". Obviously, the landmarks should be familiar to the driver, suggesting the use of commonly known objects, or references which are also present in the user's home environment.

5.1.1 Science Center To Go

Science Center To Go (SCTOGO) has been developed within the consecutive EU funded projects, CONNECT[19], EXPLOAR[20], SCeTGo[21], and PATHWAY[22]. In terms of complexity, time, and evaluation, it is the most comprehensively developed prototype, of those described here. The SCTOGO is a hands on learning environment, utilizing Tangible User Interfaces, Augmented Reality, and Transitional Objects. It addresses all three major categories of Bloom's taxonomy, the cognitive, affective, and psychomotor domain.

PROTOTYPE EVOLUTION

The main idea of modern science centers, or science museums, has already been expressed in the old Chinese saying: "*Tell me, and I will forget. Show me, and I will remember. Involve me, and I will understand*". Usually, museums visitors passively look at museum exhibits, often even advised not to touch any objects. In modern science centers, contrarily, visitors become an active part of each exhibit, and get involved into experimental learning sessions. Such an active involvement demands sophisticated exhibits, which work on many levels. Each exhibit is especially designed for providing a new perspective on a learning subject. This perspective is created interactively through the exhibit, and usually addresses multiple senses. In the EU-Project CONNECT such hands on science center exhibits have been extended through Augmented Reality (AR), to virtually show phenomena which are hard to implement in real models.

CONNECT aimed for an integration of science center visits into the school curriculum. Adaptability of learning content and remote participation were both key factors of CONNECT. Unfortunately, a science center visit is time consuming and relatively expensive for school classes. Moreover, it is very challenging to integrate given exhib-

[19] CONNECT was partially funded by the European Commission (FP6-2002-IST-1-507844)

[20] EXPLOAR was co-financed by the European Commission within the framework of the Life Long Learning Programme (135506-LLP-1-2007-1-GR-KA3-KA3MP)

[21] SCeTGo (505318-LLP-1-2009-1-FI-KA3-KA3MP) was co-financed by the European Commission within the framework of the Life Long Learning Programme (135506-LLP-1-2007-1-GR-KA3-KA3MP)

[22] The Pathway to Inquiry Based Science Teaching has received funding from the European Union Seventh Framework Programme (FP7/2007-2013) under grant agreement no. 266624.

its into the curriculum. The CONNECT platform provides a solution for both of these challenges. The virtual AR content can easily be adapted to the school curriculum via a web interface. Additional, audio-video streaming allows for distant participation of classmates. However, such a remote connection does not transport the full hands on experience of the real science center, reducing the learning experience of most participants back to a live video broadcast (Sotiriou et al., 2006; Wittkämper, Braun, Herbst, & Herling, 2007).

In the follow up project EXPLOAR, the CONNECT-platform has been evaluated in detail, and its AR component has been revised. As part of the revision, the problem of limited accessibility has been addressed by the development of a miniaturized version of one of the AR exhibits (Buchholz & Wetzel, 2009). The first prototypes were based on printed CAD/DAM models, using a Spectrum Z™510 3d printer. This procedure guarantees exact real physical representations of virtual AR models, which allows for accurate superimposition of additional computer generated information. The software was implemented based on the MORGAN AR/VR Framework, running on typical desktop PCs (Broll et al., 2005). For tracking purposes the marker based ARToolkitPlus computer vision library used (Wagner & Schmalstieg, 2007), a succeeding development of the ARToolkit library (Kato & Billinghurst, 1999). The three steps from the model to the first prototype are visualized in Figure 5-4.

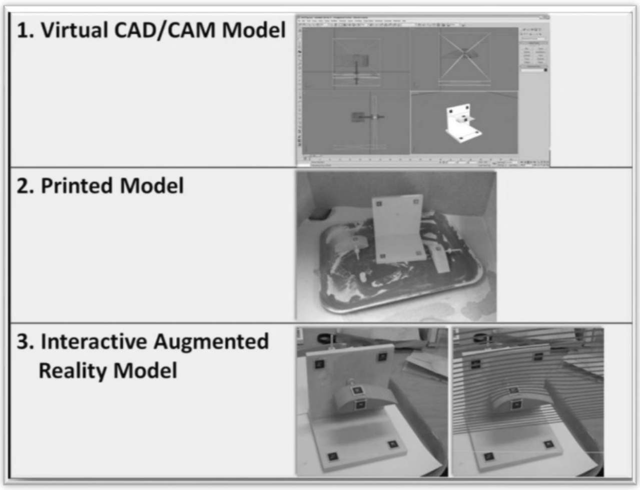


Figure 5-4. Three steps from a virtual to a tangible interactive Augmented Reality model

The SCTOGO was quickly refined towards more mobility. Therefore the second generation prototype incorporated a small box, holding the wing, and could be used on mobile devices, as shown in Figure 5-5.

Figure 5-5. 2nd generation airfoil wing in a box (middle). Augmentation on a mobile Device (bottom).

First tests seemed promising and led to a new project called SCeTGo, Within the SCeTGo project, a suitcase full of miniature exhibits has been developed and extensively evaluated. The project focused on a direct integration into the school curriculum, bridging the gap between science centers and learning in schools.

THE LATEST PROTOTYPE: THE SCIENCE CENTER TO GO SUITCASE

The suitcase stores all necessary elements for the existing five exhibits. Also included in the suitcase is a laptop with a touch screen, a webcam, and a little stand. The webcam is placed on the stand and connected to the computer. After booting, the computer directly opens the main screen, where each experiment is represented by an image. Each image displays the corresponding experiment in action, also serving as guidance for users to correctly setup and use the system. After setting up the desired exhibit in front of the webcam, a simple touch on the according image starts the software. The webcam stream is displayed on the computer and augmented with additional content. An exemplary setup, including the suitcase and laptop, is displayed in Figure 5-6.

Figure 5-6. The Science Center To Go setup for the Double Cone experiment.

In the following the five exhibits included in the suitcase will be described.

<u>THE MINI WING EXPERIMENT</u>

The MiniWing consists of a small box that stores the model of an airplane wing. The wing is about 5.5cm long, 3cm wide and 1.5cm high. It is connected to an axle that fits into a hole of the box (as seen in Figure 5-7). After the wing is brought into position, the user can easily rotate it, and try out all possible angles of attack. Only two markers are required for the Mini Wing: one is attached at the top of the box, while the other is placed directly on the wing. When the user rotates the wing, the software determines the current angle of attack by analyzing the tracking values of both markers.

The virtual representation instantly shows the air flow around the wing. Animated arrows visualize the different speeds of air, lift, and drag.

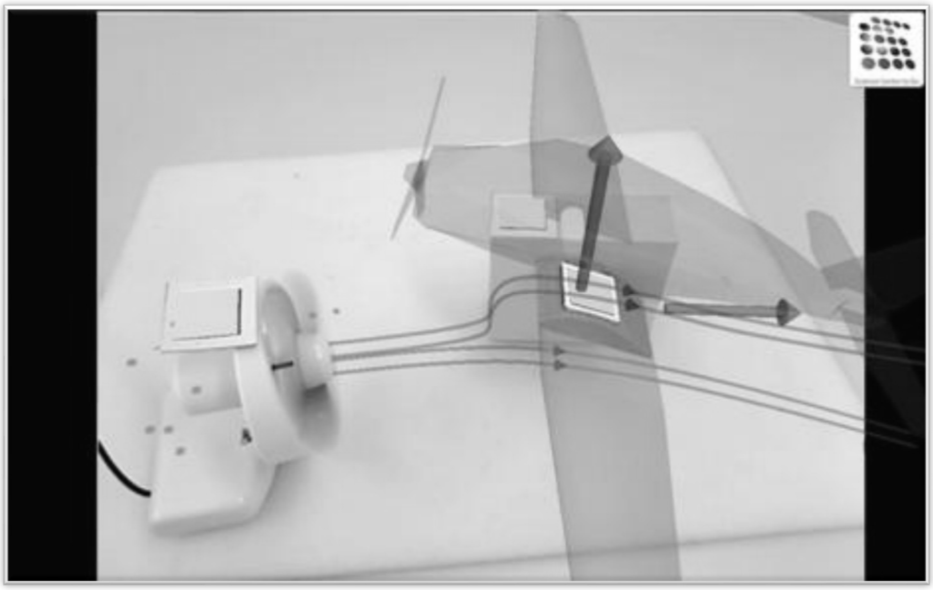


Figure 5-7. The Mini Wing exhibit augmented through the airflow. Two vectors are displaying the lift and drag.

This way the user is able to learn first-hand about the Bernoulli Effect. By experimenting and interpreting the results, they learn that the best angle, for optimal lift of the plane, lies between 15 and 20 degrees. Differently shaped wings, such as a flat door or a cylinder, may also be tried out, to compare different air flows.

THE DOPPLER EXPERIMENT

The Doppler Experiment consists of a fire truck and a virtual microphone representing a sound recording device, or listener. The fire truck holds a marker on its roof top. As soon as the truck is visible for the camera, sound waves are displayed and the sound of a fire truck horn goes off. The sound propagation is animated in a sequence of wave fronts that start off from the trucks siren, and expand concentrically away from the truck. The waves are emitted in a constant frequency. When users move the truck, they move the source of the sound waves, causing the wave fronts to be shifted closer together in moving direction. Obviously, the wave fronts are then shifted further apart in opposite direction (compare Figure 5-8). At the same time, the pitch of the siren audio is increased when the truck is moved towards the microphone, and decreased when the truck is moved away. The pitch of the siren also changes when the user moves the microphone.

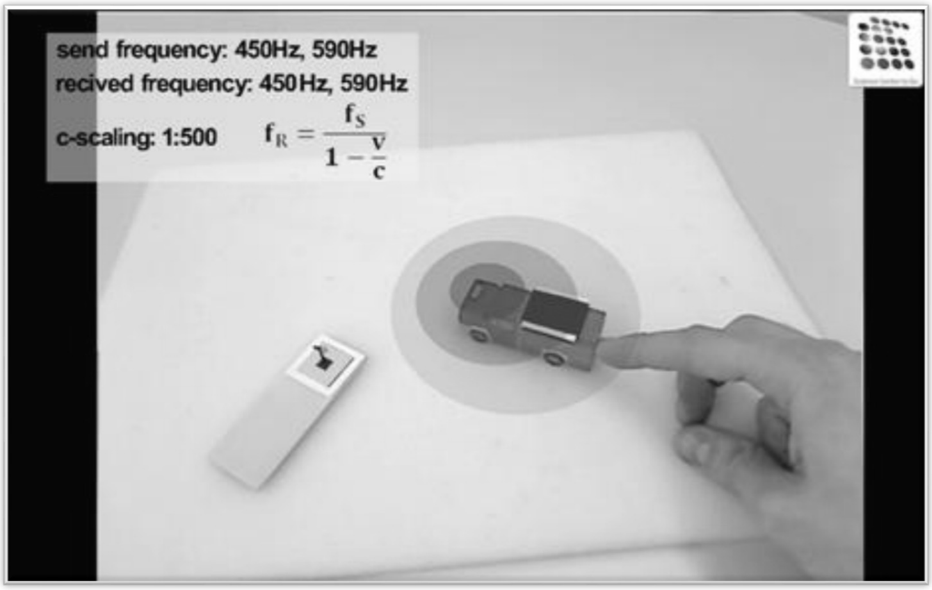


Figure 5-8. Sound waves of a fire truck displayed at the Doppler exhibit.

The audio-visual feedback, representing sound waves, allows users to learn about the Doppler Effect and the importance of relative difference in velocity between listener and the source. Observations made at this exhibit might easily be transferred to other physical phenomena related to the propagation of waves. Additionally the physical model is represented through a formula, which is updated in real time.

THE DOUBLE SLIT EXPERIMENT

The initial version of the double slit exhibit consists of the Mini Wing's box, a floor board, and a screen with either a single or a double slit (see Figure 5-9). The box serves as the end projection plane. The selected slit screen should be fixated on the floor board so that it faces the projection plane on the box. The slit screens and the box are registered via markers. The floor plane works as a fixation to ensure proper alignment and the right distances among all pieces.

Learners are able to change the double slit distances, the distance between the slit screen and the projection plane, and the emitting source. The emitting source can be a particle cannon, firing virtual little "cannon balls" towards the screen. Some of the balls are deflected while others pass the slit. The box projection plane is not deflective. Balls hitting this plane stick to it, and create a pattern, matching the slit screen.

Users might be surprised when they test the same setup with waves. In this mode the cannon is replaced by a source sequentially emitting waves with a certain frequency. The wave fronts are spreading concentrically. When a wave hits a slit on one side a new concentric wave goes off on the other side of the slit. The projection plane at the box finally shows the resulting interference pattern. A single slit creates a bright band in the middle of the box. For a double slit the projection turns into several bright bands of light, as shown in Figure 5-9.

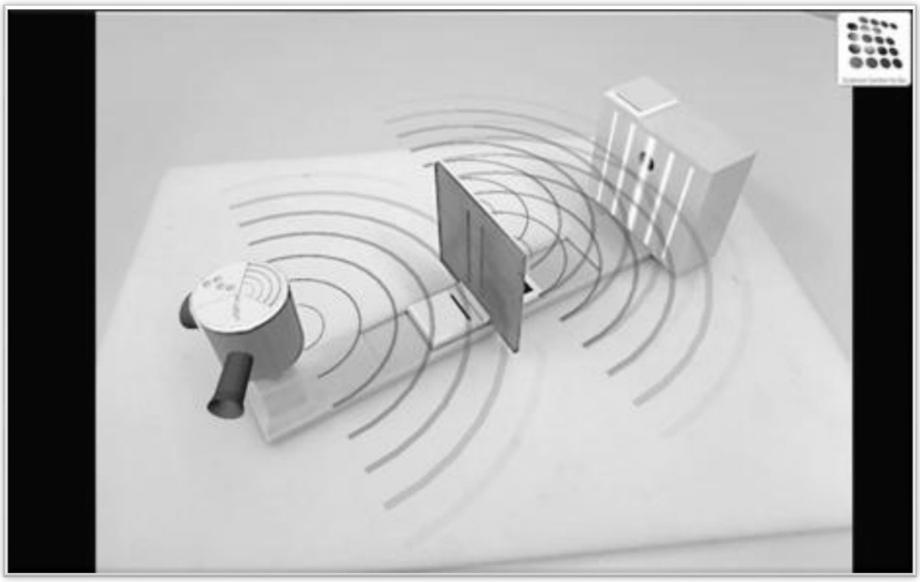


Figure 5-9. The double slit experiment simulating a wave field.

The waves are visualized and animated, allowing ambitioned observers to interactively search for areas of constructive and destructive interference. By experimenting with the double slit exhibit users learn about wave propagation, interference, the particle wave duality of light, and quantum particles opposed to normal particles.

THE DOUBLE CONE EXPERIMENT

The double cone miniature consists of two rails, each 12cm long. The rails are jointly connected on one side; on the other side each rail rests on a ramp. The ramps provide an inclination of 1.5 cm by 3 cm. Additionally, four rolling objects are available to be put on the rails. Three of the rollers are double cones and one is a cylinder. The opening angle, measured alongside the double cones, differs between 15, 30 and 45 degrees.

As shown in Figure 5-10 the rails are resting on ramps on one side. If the cylinder is put on the construction it will roll down the slope. However, when a double cone is set on the rails, it might as well roll the opposite way, up the hill.

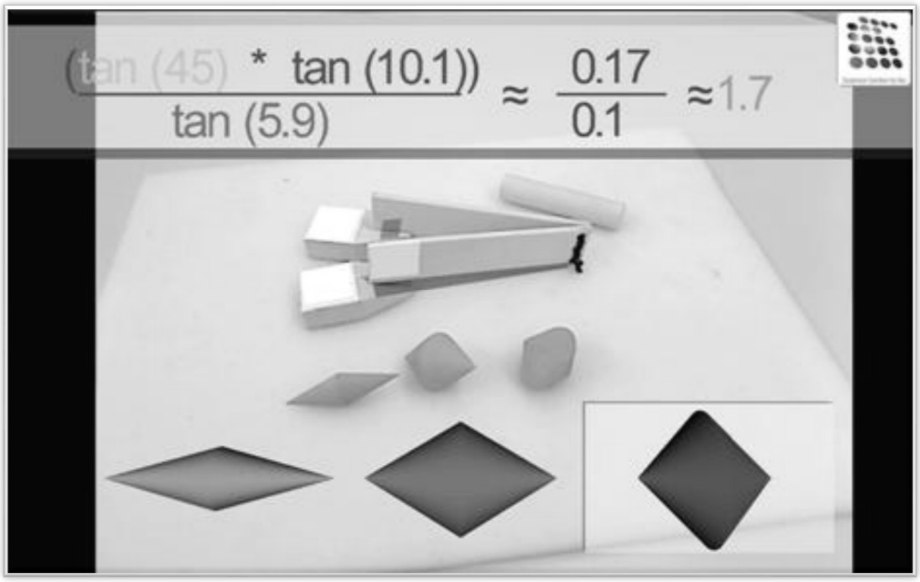


Figure 5-10. The double cone exhibit - consisting of two rails, two ramps, and four roller objects. In the double cone selection interface buttons are shown in the bottom, the resulting formula is displayed at the top of the screen.

Three angles are important to understand and predict the behaviour of the double cones. One is the opening angle of the double cone roller (α), the other is the opening angle of the rails (β), and the third one is their inclination (χ). The angle referring to the double cones' shape is selected via the user interface. Three markers are used to precisely capture the remaining angles. The opening angle is calculated from the two markers on the ramps and the resulting distance between both ends of the rails. The slope could have been determined directly from the orientation of the marker attached to the rail. Though, for increased precision, the marker on the adjacent ramp is used to determine the position of the rail alongside on the ramp, which gives us the lift of the rails' ends. From here we are able to precisely deduce the inclination. The relation of all three relevant angles, for an up-rolling element, may be described by the following expression:

$$\frac{\;}{\;} > 1$$

Formula 4. Determine the rolling direction of the double cone in the double cone experiment.

If the result of this expression is greater than 1, the selected double cone should roll towards the ramps, otherwise it rolls in opposite direction. The setup allows learners to easily change all relevant angles. The opening angle of the ramps is changed by moving the ramps apart. The incline is changed by pushing the rails up, or pulling them down the ramps. The roller object has to be selected using the buttons on the screen. The formula for predicting the behaviour of the experiment is shown, and instantly updated, at the top area of the AR screen. Additionally the angles are color-coded, directly at the physical objects, visible through the AR view.

This exhibit directly estimates the prediction model, described by a mathematical expression, and a real experiment. Hereby learners should be able to learn about the physical underlying logic of the double cone experiment. Those experiments might also reveal typical misconceptions related to gravity.

THE BOLTZMANN EXPERIMENT

The Boltzmann experiment contains three objects: A freezer, a heating surface, and a thermometer. Since touch is important for hands on learning, the refrigerator actually gets cold and the heating surface heats up. The exhibit also includes a functioning infrared thermometer, which displays the real temperatures of all objects.

Each object is registered through a marker. The markers of the freezer and the heating surface are used to determine the areas of high and low energy. The energy level between those two extremes is smoothly interpolated to provide a realistic transition.

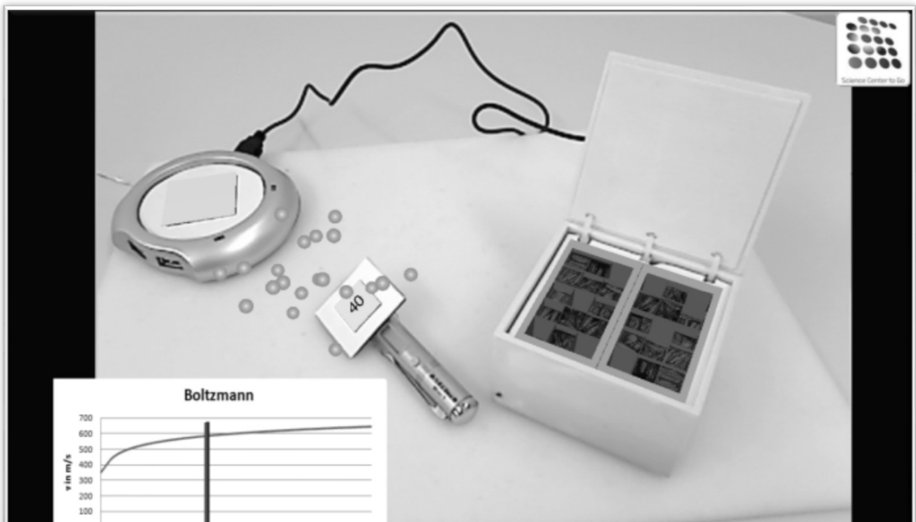


Figure 5-11. The Boltzmann exhibit: Heating plate in the front refrigerator in the back. The user holds the thermometer into an area of low energy, the molecules here are moving slower. The Boltzmann distribution is shown in in the bottom left.

After setting up the experiment users are able to measure the temperature with the thermometer at different areas of their setup. Additionally, molecule movement is visualized at the top of the thermometer (see Figure 5-11). In AR users might observe that molecules in areas of a high energy, near the heating surface, move faster than molecules around areas of low energy, e.g. inside the freezer.

With this experiment learners should get a deeper understanding and insight into the relation between energy, temperature, and molecule movement. This understanding is supplemented through a graph showing the Boltzmann distribution.

SCIENCE CENTER TO GO: A DIGITAL TRANSFORMATIVE

Obviously the SCTOGO includes elements which foster familiarity, such as Tangible User Interfaces or the use of Transitional Entities (Buchholz & Brosda, 2012). Tangible User Interfaces foster usability especially through perceived affordances, as detailed in *Cognitive and Practical Background Constructivistic Learning with Transitional* Objects and Digital Manipulatives (pp. 89) and *Affordances from UI Concepts* (pp. 99). Those affordances are based on familiarity of physical object handling. Transitional Objects increase motivation and learning interest through object affection (Buchholz & Brosda, 2012). There is evidence that affection is also related to familiarity, which has also been manifested in Feature 3 of Digital Transformatives (Langlois et al., 2000; Rhodes, 2006). The Science Center To Go builds on physical objects in our everyday environment, such as wings, freezer, thermometer, a fire truck. Those objects are generally familiar to learners in the western world, and should therefore come with increased affection. Abstract drawings, or even formulas from conventional learning situations, do not frequently appear in our everyday environment; hence, they are more unfamiliar and less affective.

The most fundamental difference, however, relates to familiarity of how we learn. The Science Center To Go aims at improving access to theoretical abstract models and formulas known from conventional learning situations in schools. Learning with formulas and models presented in text books differs dramatically from natural ways of learning. First of all, most formulas and abstract representations occur comparatively seldom in everyday live. Hence, a majority of individuals, especially children are highly unfamiliar with reading, using, and interpreting them. According to cognitive prototype theory (see *Concepts, Prototypes, and Categories* pp. 23), those unfamiliar areas correlate to decreased cognitive performance. Dealing with formulas and abstract models is arduous for many learners.

Based on the idea of Digital Transformatives, the new system should offer a performance context, with maximum user familiarity, which corresponds to a cognitive prototype, on a shared basic level of sufficient target similarity (compare Feature 10).

This requirement can easily be achieved for physical theories, since corresponding formulas usually describe real phenomena. We are used to interpret physical object behaviour. We are familiar with haptic interaction in situations where actions show direct impact. We are used to learn through exploration in a responsive environment, as manifested in constructivistic ideas (J. Bruner, 1986; JS Bruner et al., 1956; Montessori, 1946; Piaget, 1954, 1962; Vygotsky, 1964). Obviously most of our abilities, such as walking, behaving, communicating, or socializing are being learned outside of a school environment. This kind of learning happens self-directed in a natural inquiry based way. A natural way of understanding systems in our environment, such a system could be a ball rolling down a hill, a hot air balloon raising towards the sky, or simply communication with other individuals, may look as follows: the system is observed, hypotheses are raised, the system is being manipulated in order to test such hypotheses, and the result is evaluated (Kerres, 2001; Schulmeister, 2007; Weidenmann, 2001).

Consequently, the Science Center To Go provides improved access to conventional learning materials, as illustrated in Figure 5-12.

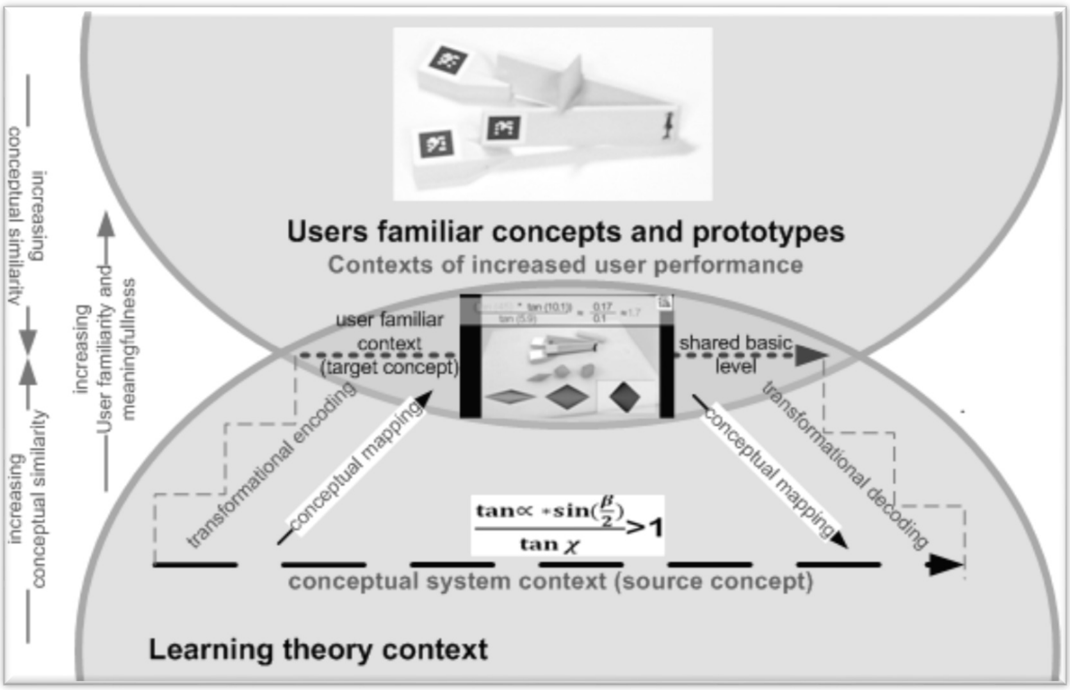


Figure 5-12. The Mixed Reality context of the Science Center To Go, aiming for a natural tangible interface to formulas.

The Science Center To Go offers an inquiry based learning platform for natural learning in a school environment (Rocard, Csermely, Jorde, Lenzen, & Wallberg-

Henriksson, 2007; SCeTGo Consortium, 2012). The Science Center To Go uses the familiar natural learning environment as an interface for controlling the formulas. It hereby generates a new perspective on formulas and models, and enhances them with familiar interaction possibilities known from physical objects.

Science Center To Go Evaluation

The Science Center To Go was developed to offer an alternative way for experimental learning, which ideally improves learning results of curriculum relevant information. Due to the various interfering influences of learning with the system, general learning success can hardly referred to just a single characteristic of the system. Hence, the evaluation results only provide an indicator of the impact of the considerations given above. As elaborated above, the implemented Digital Transformatives (DT) directly seek to increase affection in order to induce better learning effects. This may also correspond to increased usability and motivation during the use of the system. Thus, affection motivation and usability are the most direct measures for the impact of DT characteristics. Based on the idea of Transitional Objects for learning, those characteristics should indirectly also induce better learning results; both has been the case during the evaluation, as detailed in the following.

The Science Center To Go was evaluated quantitatively and qualitatively with teachers, students, and pedagogical research partners. Overall 890 users tested the system, resulting in 627 Student questionnaires, 43 Students interviews, 146 Teachers' questionnaires, 66 Teachers' interviews, and 8 research partner interviews. The validation and quality assurance was conducted by the Department of Biology Didactics at the University of Bayreuth. Multiple suitcases have been conceptualized, designed, and created. They were traveling all over Europe, and were evaluated at schools in various countries, such as Finland, Greece, Romania, Spain, Sweden, or the UK. The exhibits were presented and discussed at training session with teachers. The suitcase was also evaluated with teachers using the suitcase in the real learning environment of the classroom. The target group ranges from fourth to twelfth graders.

The results indicate that technology' acceptance is high in general, and that the usability of the system is rated very positive by pupils and teachers. In summary, all students enjoyed working and learning with the miniatures and most of the teachers assessed pedagogical effectiveness and technological aspects throughout positive.

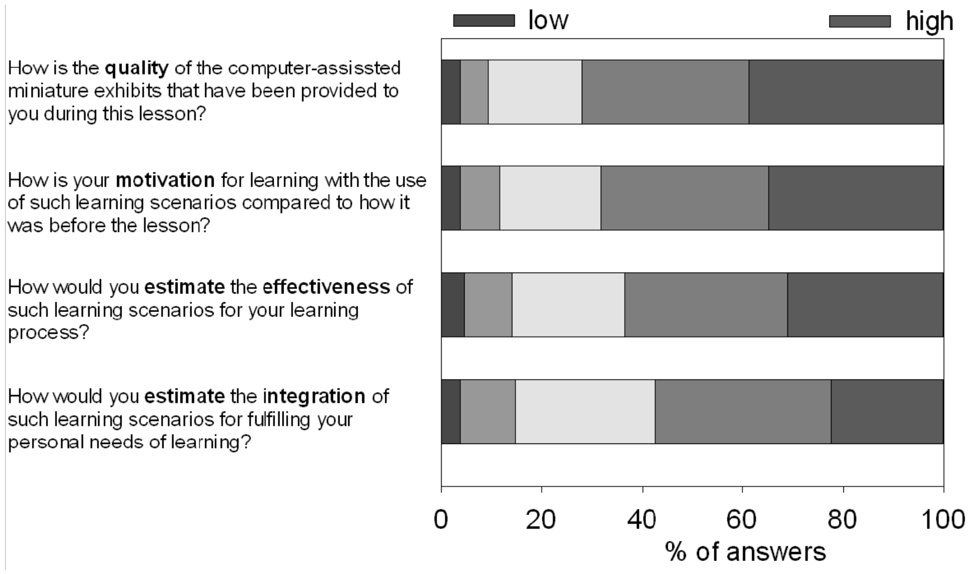


Figure 5-13. Evaluation results of students rating the motivation and usability of the Mini Wing exhibit (Larsen & Bogner, 2012; Larsen et al., 2012).

Usability features, along with software application were rated well (Larsen & Bogner, 2012; Larsen et al., 2012). The majority of teachers did not have problems, concerning the usage of the SCeTGo exhibits (64.9%). However, a minority felt uncomfortable to use the software (13.5%). The same ratio of interviewees found the physical phenomena difficult to understand and need more information (10.8%).

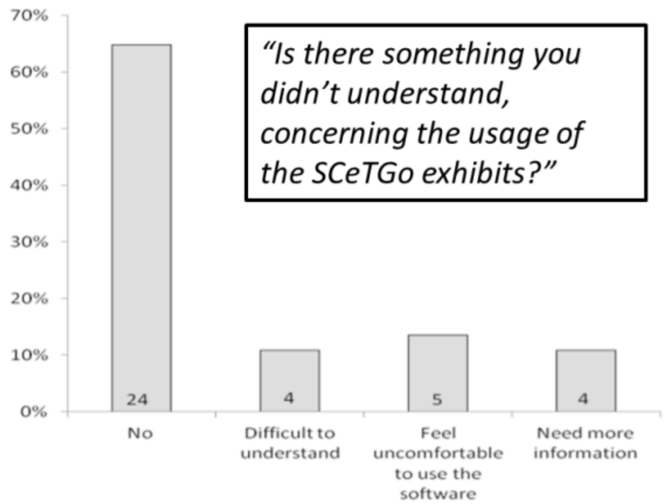


Figure 5-14. Teachers' rating of the general usage of the system (Larsen & Bogner, 2012; Larsen et al., 2012).

Further evaluation of the AR Science Center system also show that students reached better understanding through the system, compared to conventional learning environments (Salmi, Sotiriou, & Bogner, 2009).

5.2 Generation 2 – First Approaches based on Familiarity

Digital Transformatives try to address a context with maximum user familiarity, including cognitive declarative, as well as procedural knowledge. From an action based perspective, Digital Transformatives aim at finding automatized processes in order to shift the usage context into areas of high performance. This way Digital Transformatives offer a highly adapted user interface, enabling efficient usage by addressing well practiced behavior. Ideally this behavior is habituated.

In the recent years another research field gained increasing popularity, also trying to take advantage of habituation. In the field of *Captology* it is researched how technologies can be used to change user behavior.

5.2.1 Captology and Persuasive Technologies

Ever since we interact with technology it is affecting and changing our behavior. What would our daily life be without clocks, automobiles, television, computers or the internet? While many technologies are usually designed with a functional focus they may also be designed with a behavioral focus. While many changes happen casually, *Persuasive Technologies* aim at intentionally affecting human behavior.

Captology is a term describing the study of computers as persuasive technologies. BJ Fogg derived the term captology in 1996 from an acronym: Computers As Persuasive Technologies = CAPT (Persuasive Technology Lab, 2010). Captology approaches the topic of intentionally changing behavior from a psychological site. In the definition of persuasive technology a major aspect lies in the "*noncoercive attempt to change attitudes or behaviors*" (B. J. Fogg, Cuellar, & Danielson, 2007, p. 134)

Most work in this area is of empirical matter, and has been summarized in heuristics, such as *Foggs Behaviour Model (FBM)* (B. J. Fogg, 2011). According to this model, behavior is mainly influenced by the users' motivation, abilities, and external triggers. Whether a behavior is being established or degraded, depends on those three factors. Their relations in this constellation are depict Foggs Behaviour Model, shown in (B. J. Fogg, 2008)[23].

According to the FBM, behaviors occur if three factors are coming together in the right portion: *motivation*, *ability*, and an *effective trigger*. Three steps are explicit to initiate a behavior. First attention is drawn to a trigger, for example via a signal. Second the trigger is associated with a target behavior. Third the target behavior is performed when we are motivated and able to conduct this behavior. The chance for a successful trigger increases with higher motivation or simplicity. Thus, if one is highly motivated for a certain behavior, then it is not so important whether this behavior is hard to do. On the other hand, if a behavior may be easily conducted, not much motivation is required for a trigger to initiate a behavior. The core factors of Foggs Behavior Model are motivation, ability, and triggers (B. J. Fogg, 2011). The relations of those factors might be expressed in the following formula:

$$\text{Threshold}(\text{Behavior}) < Motivation * \text{Ability} * \text{Triggers}$$

Formula 5. A Behavior is triggered when the three factors of motivation, ability, and external triggers are exceeding the threshold of this behavior.

The Fogg Behavior Model includes three key factors for motivation: sensation through pleasure or pain, anticipation through hope or fear, and social cohesion through social acceptance or social rejection (B. Fogg, 2009a).

According to (B. Fogg, 2009a) a behavior will not be performed when at least one of six simplicity factors fails. A behavior is not performed if it takes too much time, money, physical effort, brain cycles, social deviance, or non-routine; thresholds of those factors are not specified further. According to this model, social deviance occurs when a behavior leads to social resistance. Fogg also believes in the common laziness of human beings. If a behavior takes too much effort to be acquired, then our non-routine might hinder us from performing it (B. Fogg, 2009a).

Finally a behavior only appears on a trigger. Whether a trigger is sufficient depends on motivation and abilities. Therefore, different types of triggers are considered. If we are highly motivated, but have low abilities, a trigger is needed, which makes behavior easier, a so called facilitator. Usually facilitators convince users that something is easy to do. When a person lacks motivation then the trigger should build on motivation. For example, triggers could include additional information for fear or potential hope. Such triggers are called sparks. Least demanding are behaviors where skills and

[23] Upon request the author of this text unfortunately did not get permission to publish this figure.

motivation sufficiently exist. In such cases a notification can be enough to trigger a behavior.

Fogg and his fellows offer different guiding materials for the design and implementation of persuasive technologies. For example the *Behavior Grid* proposes a matrix of 15 ways of behavior changes (B. Fogg, 2009b).

Based on the former heuristics of the Behavior Grid, and the FBM, a web based Behavior Wizard has been developed, which is meant to give interactive guidance on designing behavior change (The Behavior Wizard, 2012). Moreover, Fogg expresses the following five guidelines for the design of persuasive technologies (B. Fogg, 2009c):

- Choose a simple behavior to target.
- Learn what is preventing the target behavior.
- Choose the right tech channel.
- Start small and fast.
- Build on small successes.

An example for persuasive technology is “alarm clocking”. Alarm clocking is based on a common technique known from traditional sales stores. From time to time stores offer a certain deal to attract visitors. Alarm clocking can be used analogous in the internet. By building on the bad feeling of regret, when a user misses a great deal, a website could have a great deal presented every day. So users supposedly were willed to visit the page, just not to miss the deal.

CONCLUSION AND CRITICAL REFLECTION

The field of persuasive technologies seems to be highly empirical, as further emphasized by a statement of Fogg on the reason why he chose his behavior model among others. “Many people in psychology, marketing, and related fields have proposed different ways to view motivation (for references, see www.BehaviorModel.org). But for the purposes of persuasive design, I find my three-element approach to be the most useful.” (B. Fogg, 2009a, p. 44). Accordingly, most of the theoretical background on Persuasive Technologies has been developed iteratively. Models like the FBM have been refined over time providing an interesting basis for further investigations and evaluation. However, on the one hand they do not seem mature enough to provide sound evidential findings, in order to reliably refine the conceptual basis of Digital Transformatives. On the other hand, they mark a reference point for an adjacent research field of Digital Transformatives (DTs). In contrast to Persuasive Technologies, Digital Transformatives focus on the design of technology rather than behavior design.

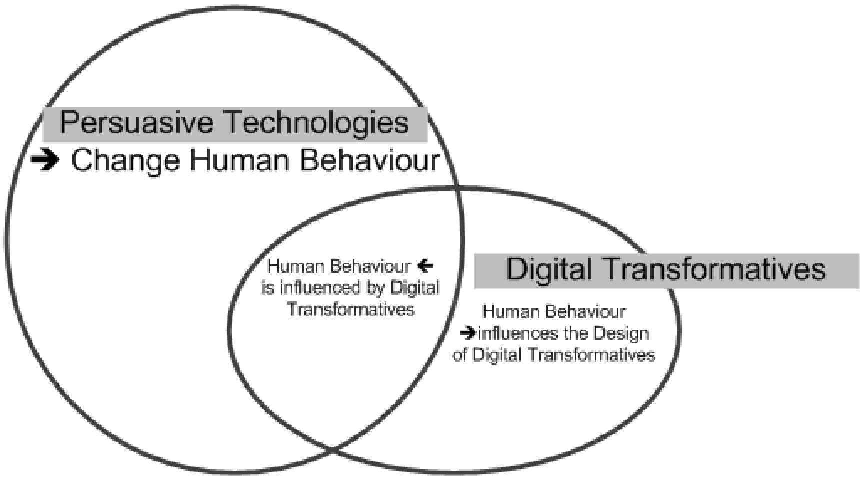


Figure 5-15. Persuasive Technologies aim at influencing human behavior while Digital Transformatives are trying to adapt computer interfaces to human behavior.

Persuasive Technologies seek for an intentional change of human behavior while Digital Transformatives are trying to adapt computer interfaces to human behavior, which in a long term may change human behavior, as Transitional Objects show. This is quite in line with Foggs Model of behavior. By building on familiarity Digital Transformatives are addressing familiar behaviors, which deliver the easiest access to a behavioral change (compare B. J. Fogg, 2011).

From a psychological point of view, Digital Transformatives predominantly address the simplicity factors of Persuasive Technologies. In this context, a remarkable statement in (B. Fogg, 2009a, p. 44) is as follows:

“People are generally resistant to teaching and training because it requires effort. This clashes with the natural wiring of human adults: We are fundamentally lazy. As a result, products that require people to learn new things routinely fail. Instead, to increase a user’s ability, designers of persuasive experiences must make the behavior easier to do. In other words, persuasive design relies heavily on the power of simplicity. A common example is the 1-click shopping at Amazon. Because it’s easy to buy things, people buy more. Simplicity changes behaviors.”

The statement that People are generally resistant to teaching and training might be disputatious. Besides that (B. Fogg, 2009a) expresses the heuristic importance of simplicity.

A subset of Persuasive Technologies and Digital Transformatives seems to be given by Transitional Object. This leads us to the following guideline:

Guideline 13. Transitional Objects offer valuable anchors to induce behavior change with Digital Transformatives.

5.2.2 Drinking Garden – a Digital Beverage Coaster

The human body is constituted between 50% and 70% of water. Body water is an essential element of the human body. Hence, drinking is essential for our health and well-being. In general it can be assumed that a human demands approximately 2400 milliliter of water a day. Liquid is ingested not only through drinking water but also through high moisture food. Deficient drinking might have little effects on our well-being, have stronger effects on our productivity and concentration, or even lead to dehydration and critical life threatening states; similar impacts can be observed for water intoxication (Becker-Carus & Herbring, 2004; Rossaint, Werner, & Zwissler, 2008; Schmitz, Lehrl, Schröder, & Wagner, 2003). Therefore, it is important for every one of us to raise awareness for water ingestion. The system described in the following aims at raising awareness and inducing a behavioral change if necessary.

CONCEPT DESIGN

The first design concept was based on the hypothesis that simple feedback, displaying drinking amounts is sufficient to improve the sense for drinking behavior. It is assumed that a better feeling for water consumption induces self regulation for a more balanced drinking behavior. This hypothesis is heavily based on other work in the field of Smart Metering (Fraunhofer FIT, 2011a, 2011b). (Darby, 2006) investigated the impact of feedback about energy consumption of smart metering systems. In most cases it was sufficient to display energy consumption, hence creating awareness and a different perspective on this matter, in order to cause measurable change. It is argued that the new perspective helps consumers to train themselves; feedback as a learning tool for reflective learning. In this context studies on smart metering systems on fuel consumption emphasize the importance of an accurate and responsive feedback system (Darby, 2006). Such insights suggest a simply interaction cycle for the first concept shown in Figure 5-16.

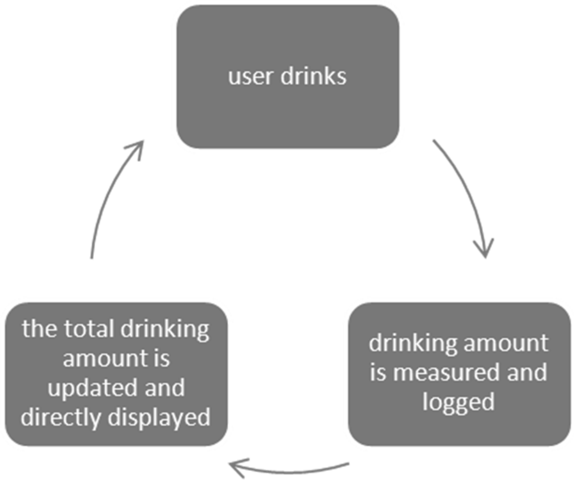


Figure 5-16. Concept of a system for direct reflection and logging of drinking behavior.

Whenever the user drinks a sensor records the drinking amount which is directly feedback to the user. Direct feedback has the advantage of providing information when the user is in the context of drinking. In this situation it is likely that his awareness and interest for drinking is increased. Contrariwise, drinking itself is a casual action, which happens almost unconscious continuously throughout the day. Analogous the system needs an ambient design, otherwise users might easily get annoyed simply by the frequency they are disturbed through the display. Consequently the system is designed as an ambient user interface, which does not demand additional user attention at any time (T. Gross, 2003; Ishii & Ullmer, 1997; Prinz, 99). Keeping a proper balance of body fluids depends on many factors, such as general physiognomy, body temperature, or stress (Rossaint et al., 2008). Due to this complexity it is hard for any system to precisely track the perfect need of body water. Therefore, the initial system design simply aims at a behavioral change rather than being an advisor for optimal drinking behavior.

IMPLEMENTATION

The first prototype has been tested in an office environment equipped with personal computer on every desk. Drinking behavior was easiest measured through a digital scale. The scale is connected to a personal computer running the software for recording and displaying the drinking amounts. A scale was assumed to be an accurate, unintrusive, and very flexible sensor. It was introduced to users as a digital beverage coaster. This way it did not introduce any change into the normal behavior or processes. If test persons were used to drinking out of a bottle they could still use the bottle, as long as they put it back onto the digital coaster. The scale also allows users to easily change drinking containers and drinks. Beverage containers may change and vary as long as they are not exceeding weight of 2000 gram. A concept drawing is shown in Figure 5-17

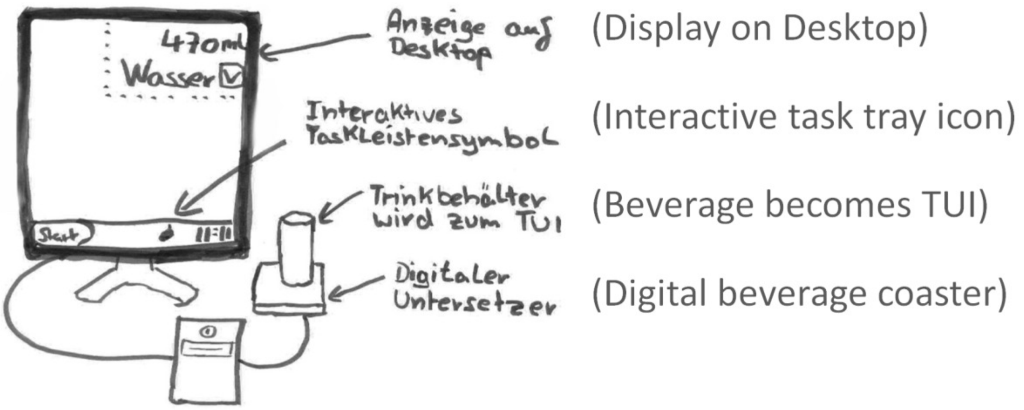


Figure 5-17. First concept drawing showing the setup of the digital beverage coaster.

As described above, the usage of the system is almost as simple as using a normal beverage coaster. Users are only required to set their beverage container onto the coaster every time after they had something to drink. As soon as users lift their beverage from the coaster the system assumes two possible actions. First the user might change the fluid level inside the container. A reduced weight indicates drinking while the user might as well refill the drink, which causes an increase in weight. Secondly, the user might change the drink. Hence, as soon as users lift of their beverages a selection list pops up on screen, showing the current drink. The list also gives users the chance to select a new one if desired. The pop-up automatically disappears as soon as the beverage is put back onto the scale. In this case the drinking display is reduced back to its standard peripheral view as shown in Figure 5-18.

When the beverage is on the coaster only peripheral information about the drinking amount is displayed. The accumulated drinking amount of the current day is shown on the desktop. Additionally a task bar indicates the current status of the system and the type of drink. If the beverage is on the coaster the icon shows a iconic glass including a drop of liquid. The drops color indicates the current type of drink, for example, it is white-blue for water and brown if the user is drinking coffee or tea. When the glass is lifted of the coaster the iconic glass disappears showing that the system is in drinking mode. The scale turns the beverage container into a tangible user interface to the software running on the computer desktop. This allows the system to directly respond to any user interaction with the drink, which raises awareness for the system.

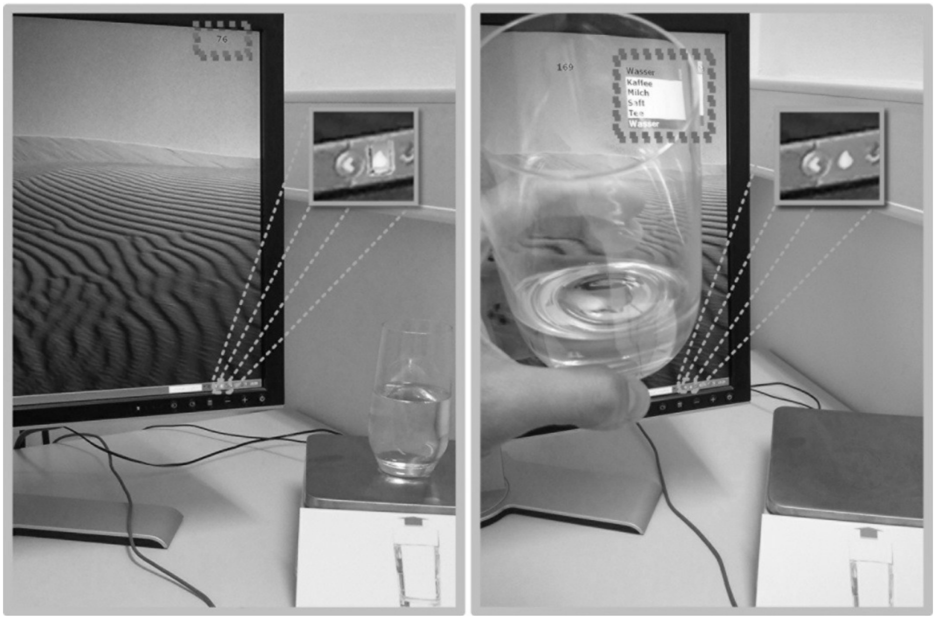


Figure 5-18. The graphical user interface of the digital beverage coaster. Only peripheral information is shown on the desktop if the beverage is on the coaster (left). An beverage change interface is shown as soon as the glass is lifted off the coaster (right).

The system in its current design is able to capture almost all cases of relevant drinking behavior. The only inaccuracies occur if a user exchanges a container with a lighter container, without changing the type of drink. Luckily this behavior has not happened during the evaluation phase of the system, since test persons usually did not use more than two containers, a cup for tea or coffee and a glass for water, juice, or milk.

FIRST TESTS

The system was tested with four co-workers in an office working environment in Germany. The youngest test person was 23 the oldest 32 years old. Half of the subjects were female. Every person tested the system in their own office in a period of consecutive working days. The test period ranged from October 2010 to May 2011. The office temperature was fairly constant while the outside temperature most of the time was below 22 degrees Celsius. The relatively constant temperature inside the office and the moderate or cold outside temperatures suggest only little influence of temperature on the measuring results.

After system installation subjects were shortly introduced to the usage of the system. They were informed about the basic idea of the tests, and that their drinking will be recorded. Further information about the actual goal of the study has not been communicated. The tests were separated in two phases. In the first phase the system did not give any feedback about the drinking amount. In this phase system simply recorded control values. This was also used to get users used to the new hardware setup. The control value phase lasted approximately one week. After that time the drinking amount display was activated, initiating the actual test phase. By collecting data in the first phase the same users could serve as their own control group. In total the system was tested for 116 days. 29 of those days were used to collect control values, on 87 days the users were in the actual testing phase.

<u>TEST RESULTS</u>

The general drinking behavior of the test persons, recorded in the control value phase, varied quite drastically. While two test persons, who drank about 200-400 milliliters per hour, highly exceeded the commonly recommended drinking amount per day, one person almost exactly met the recommended amount of 100 milliliters per hour, and another one drank 10% to little.

The display of the drinking in the test phase caused a general drinking increase of 121%. While the drinking amount number lead to a change of approximately 3% for two persons, one person drank 213,7% more water, while another reduced their drinking amount by nearly 67,7%.

<u>RESULTS IN DETAIL</u>

The drinking behavior of the first test person is shown in detail in

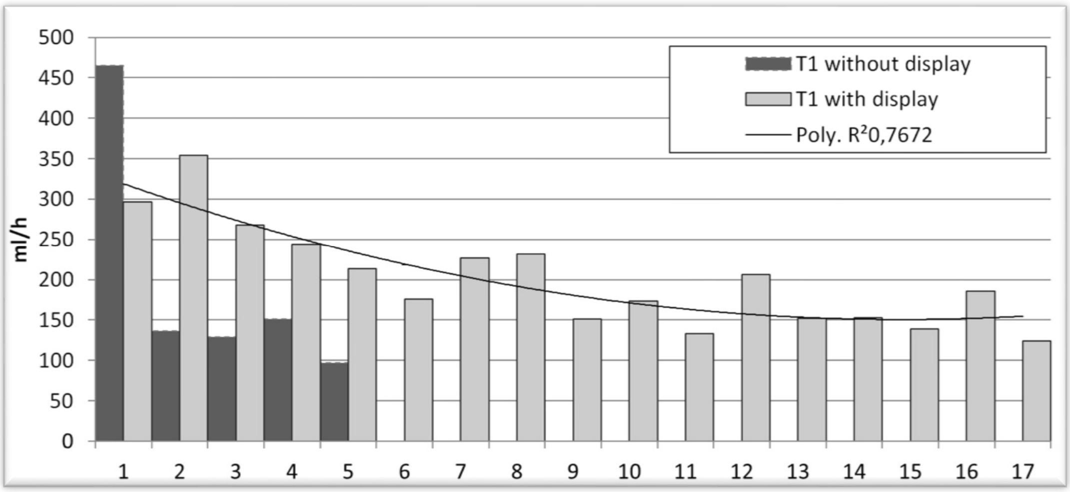


Figure 5-19. Drinking behavior of test person T1 with drinking amount display during the test phase (bright bars). Control values were gathered without giving any feedback on drinking amount (dark bars). The polynomial trend line of the test phase with stability index R^2=0,7672.

Although no drinking values are shown, the sheer installation of the system seems to have an effect on the drinking behavior. The initial maximum of the control values indicates this clear effect for the first test person. As the further control values show, the test person got used to the system already from day two on, with comparatively little changes among the following control values. The test phase is displayed through bright bars. The transition between control value phase and test phase happened from one to the other day. The introduction of the display directly had an impact. The drinking amount increased remarkably. In comparison to the average drinking amount of 195,6 ml recorded in the first five control days the subject drank 275,4 ml per hour in the first five days of the test phase. On a long term, the drinking amount steadily declined. After three weeks the drinking amount was barely above control value level.

In order to get more information about the long term effect of the digital beverage the second person was used the system over a longer period of in total 54 working days. The drinking behavior is shown in Figure 5-20.

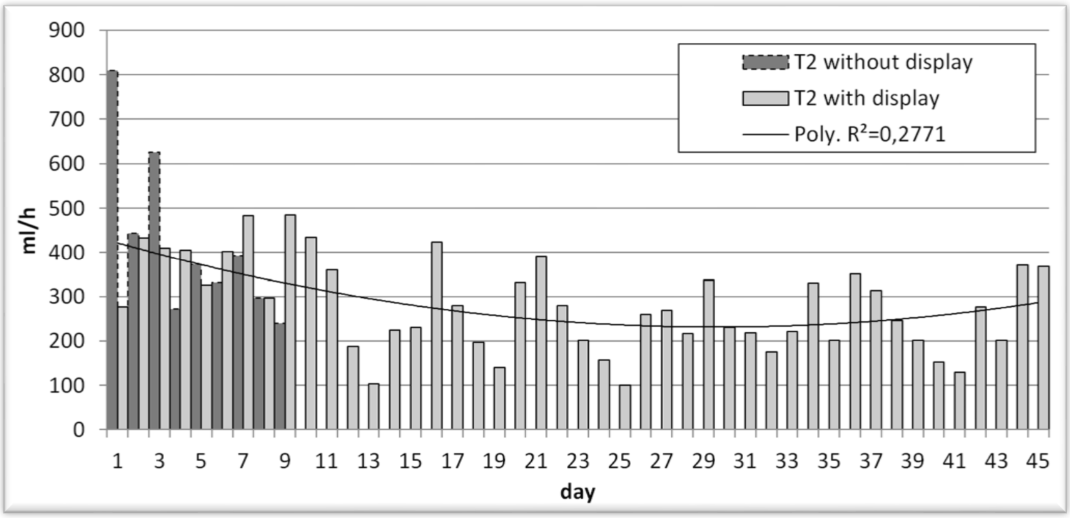


Figure 5-20 Drinking behavior of test person T2 with drinking amount display during the test phase (bright bars). Control values were gathered without giving any feedback on drinking amount (dark bars). The polynomial trend line of the test phase with stability index R^2=0,2771.

The curves are similar to the one of the first test person. The impact of the hardware installation of the system is indicated through a clear peak on the first day. The control value phase was extended here due to the peak on day three, however, from day 3 to day 9 the variation of the values settled, and the user seemed to be used to the system. The display of the drinking amount during the test phase caused only a very moderate increase in drinking behavior, in the beginning. Similar to the first test person this effect vanishes after approximately 11 days of usage. In the following 34 days the values do not differ remarkably beyond the general variation. A very similar behavior was observed by the fourth test person. Here the initial effect in the test phase lasted only for the first five days.

The behavior of the third test person, however, was comparatively exceptional as shown in Figure 5-21.

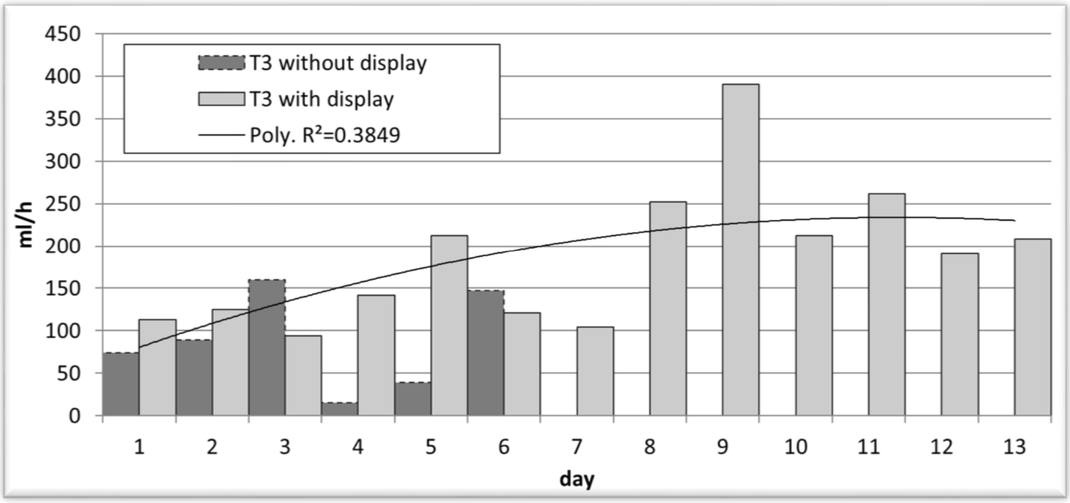


Figure 5-21. Drinking behavior of test person T3 with drinking amount display during the test phase (bright bars). Control values were gathered without giving any feedback on drinking amount (dark bars). The polynomial trend line of the test phase with stability index R^2=0,3849.

It can be said that the subject showed a clear behavioral change even after a longer period of time. In an succeeding interview the test person reported that the system made her aware of a drinking deficit. Moreover, the test person had the feeling that the drinking behavior was changed even after the system was removed. In this case, the general drinking amount of 87,3 milliliter was remarkably low, and definitely below values usually recommended in the health sector (Becker-Carus & Herbring, 2004; Rossaint et al., 2008; Schmitz et al., 2003).

CONCLUSIONS

A persuasive system for changing drinking behavior has been described, eventually trying to prevent critical states such as dehydration or water intoxication. A prototype has been developed and tested. A first evaluation indicates a clear impact of the installation with a receding effect after several days of usage.

At first glance the system seems to induce behavioral change. By showing a body parameter which is usually not perceived that accurately through our system, test persons obviously changed their drinking behavior. According to the idea of Persuasive Technologies this change was caused self regulated. Half of the test persons showed significant changes in their drinking behavior, in average the drinking amount increased by a factor of 1,21 among all participants.

On the other hand, the behavioral change vanished over a longer period of time for three of four persons. For those subjects the impact of the system was not sustainable over a longer period of time.

It seems that the biggest impact comes from the physical presence of the system, which causes its users to reconsider their common drinking behaviors. However, the tests also show that the first prototype is not able to keep up awareness over a longer period of time, as it was meant to be.

Therefore, the next iteration prototype demands an interface which raises, and keeps up, awareness for drinking. This led to the development of a virtual garden. The garden is connected to the digital beverage coaster. Whenever users drink they also water the plants of their virtual garden. When plants are not watered they wither. Flowers are positioned on the computer desktop as shown in

Figure 5-22. The Drinking Garden user interface. For raising long time awareness, users of the system need to care for a virtual garden. They water virtual plants by drinking.

Since drinking behavior is highly different, users are able to adapt the plants in their virtual garden to their drinking needs. While a sportive person with a generally higher drinking volume may keep up a big tropical garden, others might only be able to keep a small garden of various cacti. Moreover, shared virtual gardens are planned, allowing several users to share a garden.

5.2.3 Assess the Assessable

Computers in all kind of variants are saturating daily life progressively more. Human machine interaction (HMI) is inevitable in our modern world. Usually human machine interfaces are designed iteratively. Three major steps are performed per iteration during a HMI design process: a conceptualization phase is followed by an implementation phase which is then evaluated. The results of the evaluation phase kick start the next iteration demanding a re-conceptualization, re-implementation, and so on (Dix, 2004; Jakob Nielsen, 1989; Rogers et al., 2011). The importance of thorough evaluation is fundamental for any further development. Budget and time often do not allow for a quantitative evaluation at every iteration. This leads to a wide use of small scale qualitative assessments during a software design process.

In psychology it is commonly agreed on the relativity of human perception (Norwich, 1983). We judge, think, and rate relatively. Interestingly, a majority of evaluations in human computer interaction are conducted in a non-comparative way (Brooke, 1996; Shneiderman & Plaisant, 2010). De Bruin et al. state that a fundamental problem of psychology lies in mapping subjective feelings onto quantitative measures (de Bruin, Fischhoff, Millstein, & Halpern-Felsher, 2000). The use of non-comparative scales might reinforce this problem.

On the other hand, there is a comprehensive set of analytical tools available for transferring non-comparative assessments into expressive mathematical representations. Moreover, it only demands pen and paper to mark boxes in a non-comparative assessment, while it is much harder to provide a comparative setup that allows for arranging and capturing ratings relatively to each other.

However, pen and paper are not standard tools for capturing information anymore. Computers, especially in a mobile version, are rising in their availability. Compared to pen and paper they offer more ways for capturing user data.

A look at the evolution of human computer interfaces shows a trend from abstract to relative input mechanisms. While the command shell was a predominant input mechanism for early computers, the desktop metaphor and corresponding input mechanisms became typical for later personal computers. Especially the upcoming generation of multi touch devices is building on input methods which necessarily do not make use of a keyboard anymore.

Looking at the development of HCI and the commonness of computers, one might wonder why questionnaires did not evolve in a similar way. Today's digital questionnaires merely differ from their paper ancestors of 20, or even 50 years ago – but why? Shouldn't we make use of latest developments in HCI and change digital questionnaires to offer more natural mechanisms for user assessment input? Should comparative scales be used more often, and would this really change lead to significantly better results?

As a starting point, for designing a comparative tool as a Digital Transformative, the following hypothesis is formulated: Users perform better if they judge in comparative ways. This should especially be relevant for small scale qualitative tests, where quantitative mathematical operations may not be applied with sufficient significance.

BACKGROUND

Not just since constructivistic views were becoming increasingly popular in learning, it is commonly agreed that reality is perceived in different ways by different viewers. Everyone has his or her own perspective. Perceptions, interpretations and other cognitive processes are related to previous knowledge and experiences (Ginsburg & Opper, 1988; Solso, 2005; Vygotsky, 1964), consequently human decisions or judgments are hardly being objective (Baron, 2008; Kahneman & Tversky, 1984; Tversky & Kahneman, 1981). *The Prospect Theory*, for example, described by Tversky and Kahneman includes multiple experiments indicating that decisions are not always rational but often affected through cognitive biases, such as framing (Kahneman & Tversky, 1979).

Cognitive bias and illusions are strong indicators for the relativity of perception and other cognitive processes (Solso, 2005). One should be aware of such effects when evaluations are conducted to minimize distortions.

The *Adaption Level Theory* by Helson says that the judgment of an individual is based on a frame of reference consisting of previous exposure to certain stimuli as well as recollection of past judgments of similar stimuli (Helson, 1964).

Many comparative and non-comparative techniques have been developed and described in literature. As stated above, the Likert scale is probably the most common scale for non-comparative assessments (Shneiderman & Plaisant, 2010).

Usually probands are asked to rate items on a scale consisting of a certain amount of classification levels. Such scales usually aim on capturing the degree of agreement of interviewees with certain statements. A Likert scale is subject centered with the purpose to assess respondents, while a comparative scale should be used to capture stimuli (McIver & Carmines, 1981). Other non-comparative scales are continuous rating scales, line marking scales, itemized rating scales or semantic rating scales.

Pairwise comparison is usually used to compare tangible and non-tangible subjective impressions, e.g. is item A nicer than item B. Usually every item is compared with every other item. Thus, pairwise comparisons are considered to be relatively time-consuming for each test person since n test items result into n*(n-1)/2 comparisons. A preference matrix might be used to finally put items in order.

Thurstone describes a probabilistic basis for pairwise comparisons (Thurstone, 1927). He argues that the comparison regarding a certain feature of an item is usually based on a latent scale, inherent in every individual. The subjective judgment would differ among individuals and also for the same individual in different situations. He proposes that the rating for all items on such one dimensional latent scales underlie normal distributions, where all distributions are identical, only differing in their mean value. This leads to the assumption that the difference of two compared ratings is also normally distributed. Later investigations, especially those by Bradley and Terry (Bradley & Terry, 1952) and Luce (R. D. Luce, 1959), lead to further refinements of this model, also finding some flaws such that the normal distribution is oftentimes not a proper and accurate description of the distribution of rated items.

Thurstone later found that his model could be extended to comparisons with more than two choices (Bock & Jones, 1968; R. D. Luce, 1994; Thurstone, 1945). Such refinements are coming close to the *Rasch Model* which is a probabilistic model to determine how well a subject dealt with an item of certain difficulty (Rasch, 1960). Other refinements of the law of comparative judgment focused on group measurements through rank order data and paired comparisons (Li, Cheng, Wang, Hiltz, & Turoff, 2001).

While paired comparisons can be transformed into ordinal scales, rank order data might also be directly captured. Bogardus (Bogardus, 1925) used ranking to measure social distances. He asked persons to allocate several ethnic groups into 7 categories ranging from (7) “close kinship by marriage”, over (4) “employment in my occupation in my country” to (1) “Would exclude from my country“. Hereby he determined the levels of social distance for each of the seven categories. Sometimes the personal allocation of ranks is not always universal. In such cases users are asked to order items or statements themselves. In this case they typically first need to name items marking the extreme values and then rank all other items in between those extremes.

<u>System Concept for Releasing User Intrinsic Performance</u>

The hypothesis motivated in the beginning will be taken as the starting point for a first experimental concept. It is assumed that many non-comparative ratings, known from common questionnaires for evaluating human computer interface design, can at least adequately be replaced by comparative ratings, which might even result in more accurate assessments. This hypothesis is based on the observation that usual questionnaires, often used in human computer interface evaluations, offer an absolute rating scale in an indefinite expressive reference frame. Such a schema is controversial to the relativity of human behavior, thinking, and judging.

When relative subjective impressions have to be mapped on an absolute scale sedulous test persons might in this case try to relativize such a question. They could concretely start thinking of other examples, and rate requested items compared to ones they know. On the other hand users might not take the extra effort to create themselves a reference frame. Those users simply try to rate the gut feeling they perceive according the given question, and put it somewhere on the provided scale. Thus, ideally users transform the absolute rating system into a mindset that allows them for a relative rating. This transformation demands extra effort. Consequently, interviewees need to burden extra load with every non-comparative rating as visualized in Figure 5-23.

According to this concept users need to perform three steps before they feed back their assessment results to the interviewer. First they should get aware of the reference domain, then they are required to structure the domain according to the demanded question, and including the item to be rated, Finally, they have to map the structured domain onto the absolute scale providing an assessment for the requested item. Steps one and three are communication efforts for the assessor to input their results into the questionnaire. The actual assessment happens in step two. Instead of providing the interviewees with an interface that speaks their language the interviewees are requested to map their output onto a given scale. They need to learn the interfaces language. This translation might cost extra capacities for the assessor.

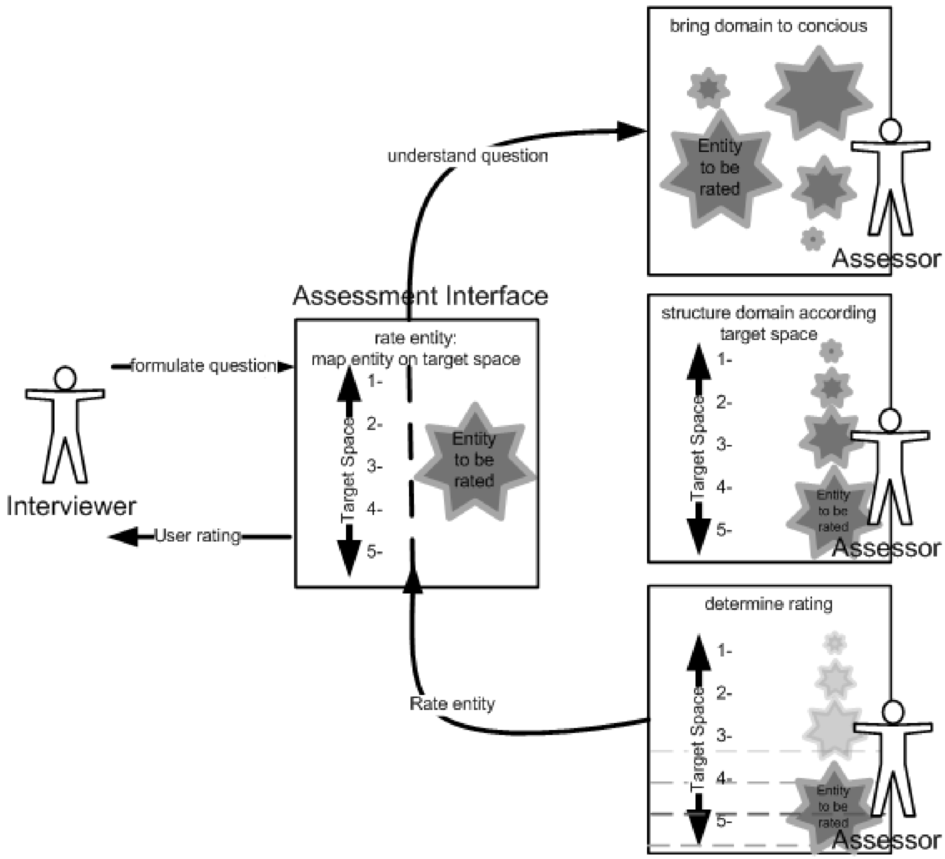


Figure 5-23. Workload model of a non-comparative assessment.

The assessment interface should be more intuitive for assessors, freeing capacities for assessment. Such an interface should reduce the interviewees' workload. In this case the interface should consist of step one and three. The users task is reduced to structuring concrete items in a provided domain space (see Figure 5-24).

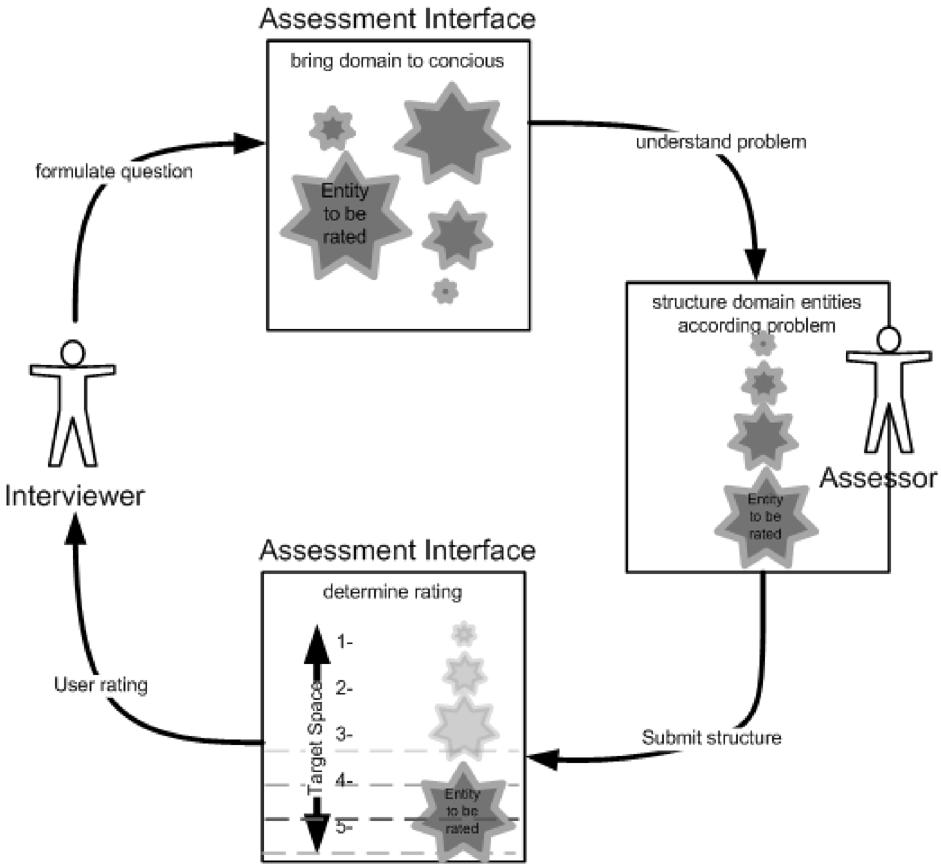


Figure 5-24. Workload model of a comparative assessment.

The basic idea is to take away some assessment load from the probands, and offer them a tool to rate in a natural way – a highly familiar way of relative judgments. The assessed items are brought into reference by the evaluation system and consistently mapped onto a scale, allowing for further analysis.

Assess the Assessable a Digital Transformative

As argued earlier, currently most assessment procedures and tools request absolute judgments from users. However, there is no cognitive prototype known where humans conduct an absolute assessment without using additional tools. If we measure distances or time, we utilize tools. Assessments simply based on human judgment are never absolute, but always comparable. In order to design an improved assessment tool, one needs to find a shared basic level which allows for mapping on absolute human judgment values. So what is the most salient prototype for relative judgments? One of the first, coming to mind are spatial comparisons. Since we are living in a spatial environment, spatial comparative assessments are likely to be a general salient cognitive prototype.

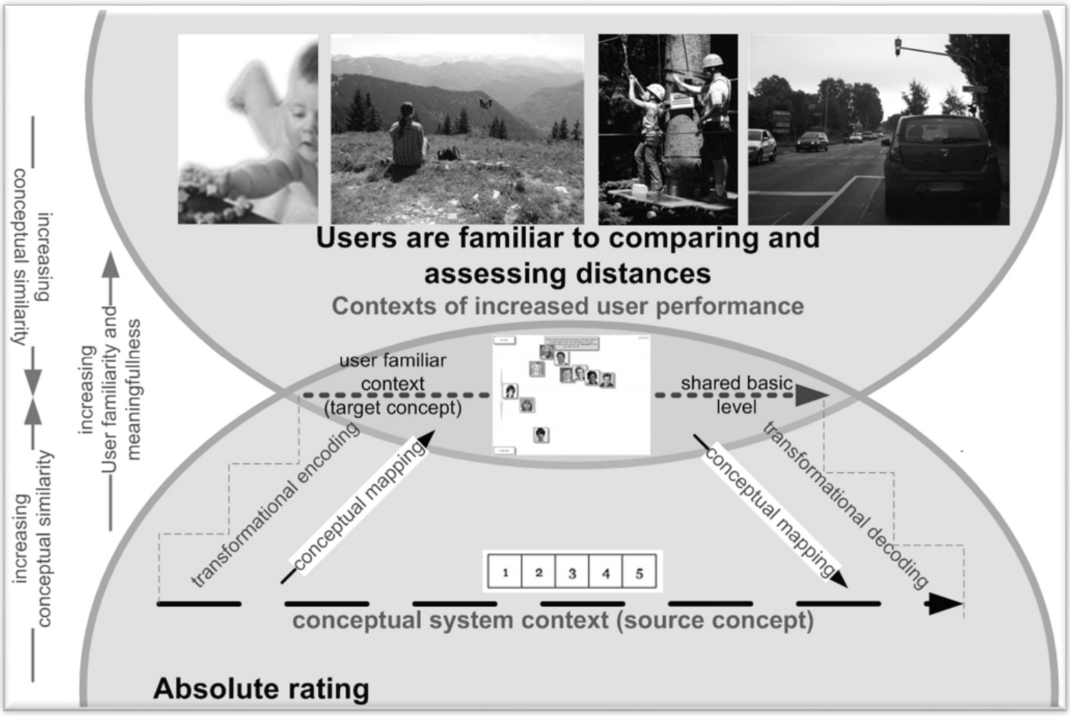


Figure 5-25. DT principles involved in deriving absolute user assessments from comparative inputs.

Various situations in daily life demand spatial judgments, for example, when we reach out for something, or when we orient ourselves and navigate through our environment. Spatial judgments also highlight the importance of the frame of reference. If we are sitting in a car, a distance of 1 kilometer is considered to be nearby, while the same distance is far away if we are walking.

Hence the rating interface needs to be based on spatial input elements, allowing users to express comparative ratings between entities. The concept described above led to testing spatial based comparative assessments in comparisons to non-comparative ratings. The experimental methodology is described in the following.

EXPERIMENTAL METHODOLOGY

Two basic requirements were formulated for the test case. First, assessments on physically measureable, ratio scaled occurrences were demanded. This way, human judges with objective definite control values can be evaluated. Secondly, judgments should be minimally influenced by cognitive biases. In order to minimize disturbances of influencing factors it has been decided to build on basic human perception without semantic connotations. Examples of occurrences that meet both requirements are the intensity of light and sound. The frequency of both would not be sufficient, since it could only be mapped on an interval scale. Since interval scales are affine spaces assessments of user judgments were also only relative.

Hence, the first tests built on different levels of brightness. Those stimuli are very low level, meaning that their processing does not interfere much with other cognitive concepts and experiences. Moreover, the can easily be shown on a computer display, and test persons are able to perceive the full intensity spectrum from dark to bright.

Additionally, a comparative rating interface was needed. Most computers provide us with audio visual in- and output devices. Since pointing devices are most common computer input devices, it has been decided to build on a spatial mapping. Additionally, judging and comparing positions is a natural human skill, not demanding much cognitive load.

The initial pre-tests have been conducted with pen and paper, which led to some refinements of the later tests.

PRE-TESTS

The initial pre-test was separated in two phases: In the first phase users were shown five cards of the same color but with different levels of brightness, one after the other. For every card they were asked to rate and write down the estimated brightness on a scale from 1 (dark) to 100 (bright). In the second phase all five cards have been put on a sheet of paper. Users should now vertically arrange the cards according to their brightness. In the end the brightest card should be placed at the top, the darkest at the bottom of the assessment area, and all other cards accordingly in-between.

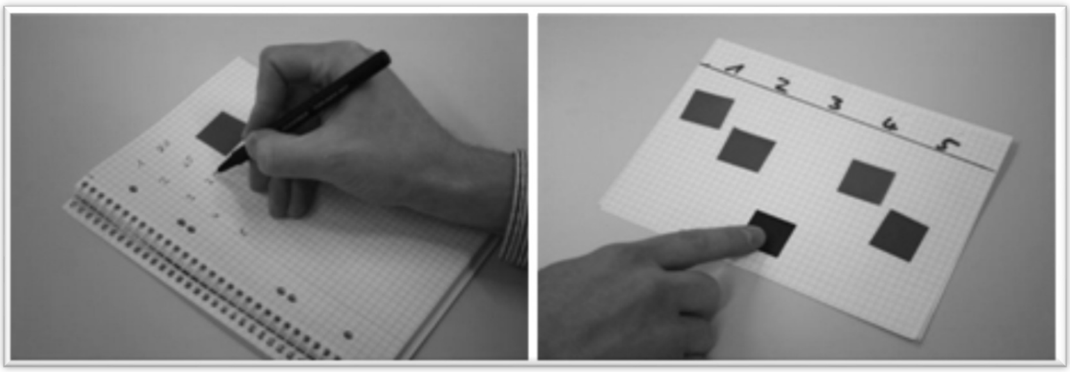


Figure 5-26. Pre-Test: Comparative and non-comparative assessment of brightness using a pen and paper prototype.

The pre-tests were conducted with four test-persons, who were asked to think aloud during the test. Every test person had to rate five green and five red cards. Each test ended with an interview. The pre-tests revealed clearly better results for the comparative assessments. The tests also showed that test persons memorized cards set in phase one, and compared them with cards from phase two. Although this strategy gives further hint that the initial hypothesis points into the right direction, such disturbances should be minimized.

Users should only give one absolute rating, since every further rating might be compared to a former one. Thus, the test was adapted. Instead of starting with a set of absolute inputs it was decided to reduce this part to only one single non-comparative assessment, and then continue iteratively with another four comparative assessments. Thus, the first of the five assessments was made without any visual reference.

TESTS

The tests were conducted via a public web interface on a voluntary basis. Participants were free to decide where and when they perform the test. The users were not informed about the real aim of the test.

The test interface is shown in Figure 5-27.

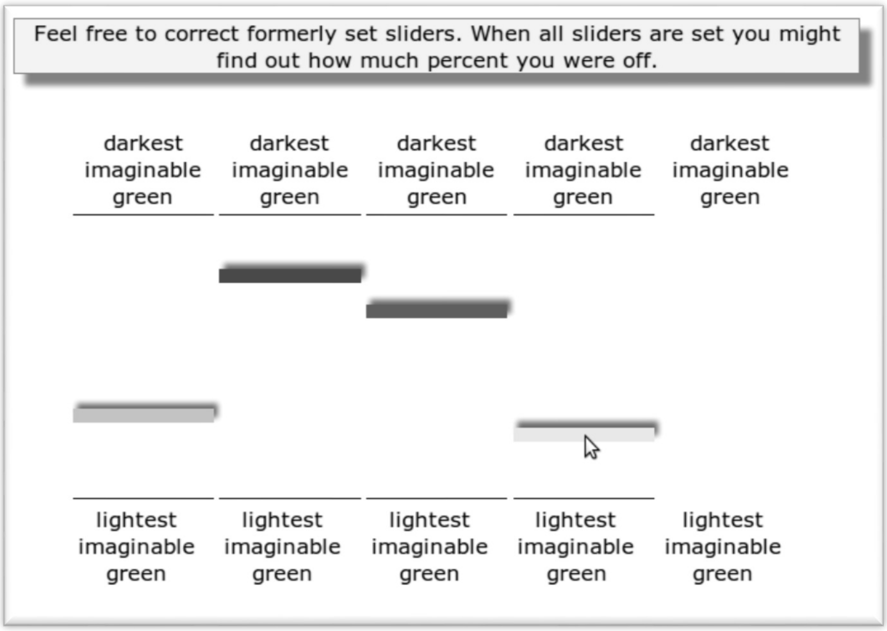


Figure 5-27. Test User interface: Five sliders had to be arranged from top (dark) to bottom (bright). The sliders appeared sequentially from left to right. The two consecutive states between setting slider 4 (left) and showing slider 5 (right) are shown above. Users were engaged to adapt former sliders whenever desired.

At each run users had to arrange a set of 5 sliders vertically according to their brightness via drag and drop. At random, all sliders of a set were either red, green, or blue – similar to the color perception mechanism of the human cone cells. The brightness of every slider was randomly generated on a equally distributed basis, varying from dark to bright. At each set the sliders were shown in consecutive order from left to right. Each run started with only the first slider being visible, ready to be rated. After it had vertical positioned the next slider appeared to be assessed, and so on.

Every new slider provides a new reference for comparison. Users were engaged to adapt formerly set sliders whenever desired. Time was not announced to be a success criterion, although it was internally measured.

The brightness range was mapped linearly onto an interval from 0 to 1 with a step size of 0.01. From 0 to 0.5 the saturation was set from 0 to 100%, and from 0.5 to 1 additionally the brightness level was linearly increased to 100%. Giving the example of a red slider, the darkest value corresponds to the red green blue (r, g, b) tuple of (0, 0, 0). Brightness values between 0 and 0.5 were linearly mapped on the corresponding red values between 0 and 255 while green and blue stay zero; (0, 0, 0) to (255, 0, 0). And the range between 0.5 and 1 was mapped on green and blue values between 0 and 255 while red stayed 255; (255, 0, 0) to (255, 255, 255).

At the end of each run users were provided with the average deviance from the real reference values. Further statistics were provided to help putting performance into context, and to give little additional inducement as shown in Figure 5-28

Besides the average slider offset, no feedback was given on slider level, making it harder for users to learn reference values. Additionally, users were provided with the recent top scorers, an all times high score, and high scores of their average error distinguished by color. The high scores were complemented by an overall distribution, providing additional references on assessing results in comparison to others. A personal history complemented the competitive scores in form of a personal learning curve. Besides statistical incentives, most users were also provided with candy after finishing five runs.

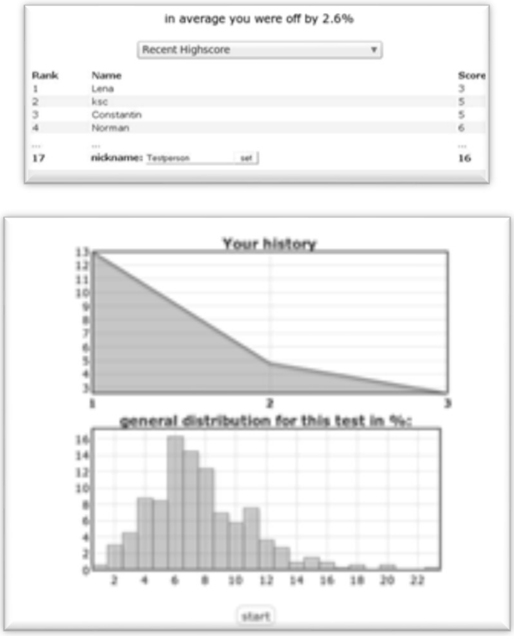


Figure 5-28. High score, history and general distribution statistics were provided after each assessment run.

<u>EVALUATION RESULTS</u>

540 runs have been conducted. Due to a submission problem only 396 of those runs were valuable. This added up to 1980 valuable samples by 39 subjects. The test subjects were between 9 and 65 years old, with approximately 41% female users. First we will have a look at the absolute deviance, which is determined as the difference of the user rating and the real reference value. From the given samples the absolute offset of each slider is shown in Figure 5-29

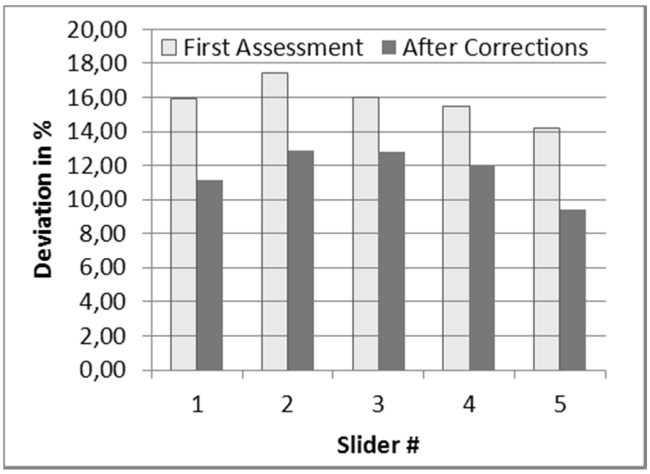


Figure 5-29. Absolute average errors for each slider at first assessment and after corrections.

With an average error of 15.81% the first slider inputs were clearly more error-prone than the inputs after correction, with an average error of 11.65%. The first input of the first slider was considered to be non-comparative since no clear references were given at this point in time. The error of such non-comparative inputs was 15.96% in average. Interestingly the error of slider two (17.45%) was even higher while the third slider (16.00%) reached almost the same level as slider one. Slider four (15.45%) and five (14.20%) in average had better results at first guess than the non-comparative setting of the first slider.

The differences in the error rates of slider one to five are unlikely to be caused by optical illusions, since the reference brightness levels of each slider were randomly assigned on an equally distributed basis. Such characteristics may rather reflect the users' uncertainty during assessment. In this case the second slider seemed to even increase uncertainty, indicated through an increased error rate. The error rate decreased with the additional references of slider three and four, which might also have raised certainty. Slider five shows the lowest deviance. The first assessment also shows a continuous error decline starting from slider two, getting below the initial error after slider three. The error rate might correspond to test persons' uncertainty. Another indicator for this hypothesis might be given by the time taken to set each new slider in context of the others, as shown in Figure 5-30.

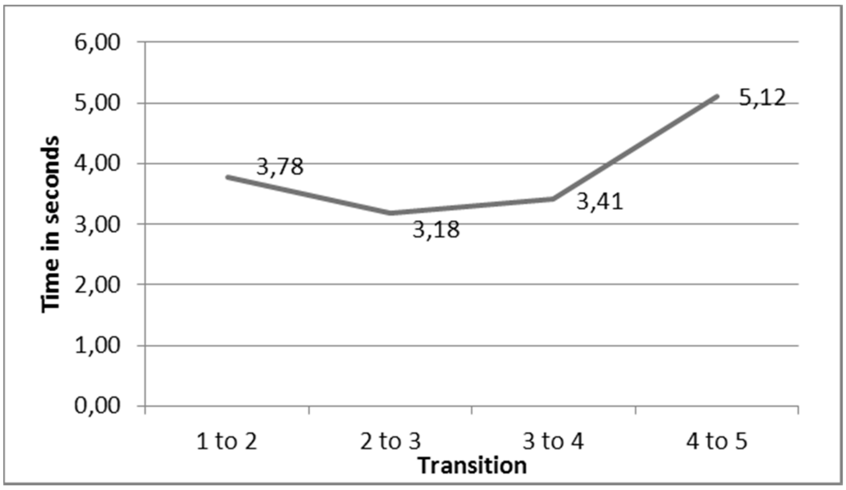


Figure 5-30. Time to continue with the next slider (including a 1 second fade in effect of the user interface).

Including a one second fade in effect, in average it took subjects 3.78 seconds to set slider two after it was visible, 3.18 seconds at the transition from slider two to three, and another 3.41 seconds for assessing slider four, the first time after it appeared.

While the differences of the first three assessment time spans vary by only 0.6 seconds users in average demanded 1.71 seconds more to set slider five. This difference might give hint on the time test persons took for reconsideration, after slider four was shown. Another evidence for this reconsideration break is visualized in Figure 5-31. The diagram shows how often users decided to set further sliders before they started correcting previous ones.

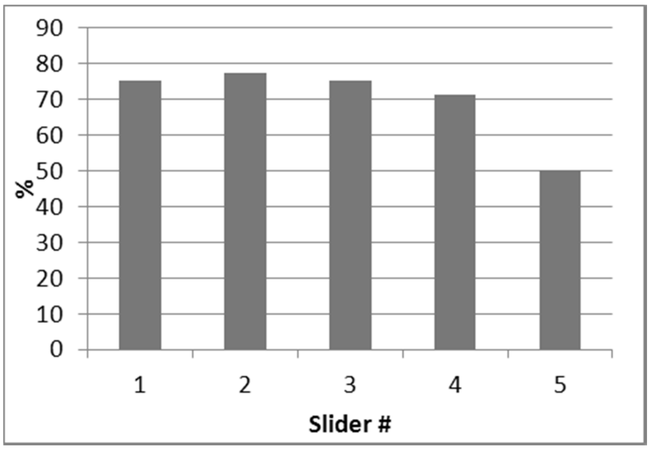


Figure 5-31. How many users continued their input without correcting previous sliders, determined relatively to those who continued.

The diagram shows that between 22.72% and 28.90% of users decided to start reconsidering formerly set sliders for slider one to four, while 50%, decided to reconsider former sliders before they set slider number five (compare). The time spans shown in Figure 5-30 together with the re-consideration rate from Figure 5-31 gives evidence that in average four references were necessary to reach enough certainty for the reconsideration of formerly set sliders.

The sequence of Figure 5-31 is also similar to the one shown in Figure 5-29, indicating a correlation between users not correcting an input right away, and an increased error rate. Figure 5-29 also indicates that users do not tend to make big relative changes after the slider position is set the first time. Although the corrected absolute error in average is lower the relations stay the same (compare outlined and bold bars in Figure 5-29).

Although the initial absolute error of slider one was just above average (compare Figure 5-29 outlined bars), it has been corrected by the users significantly more often, as shown in Figure 5-32.

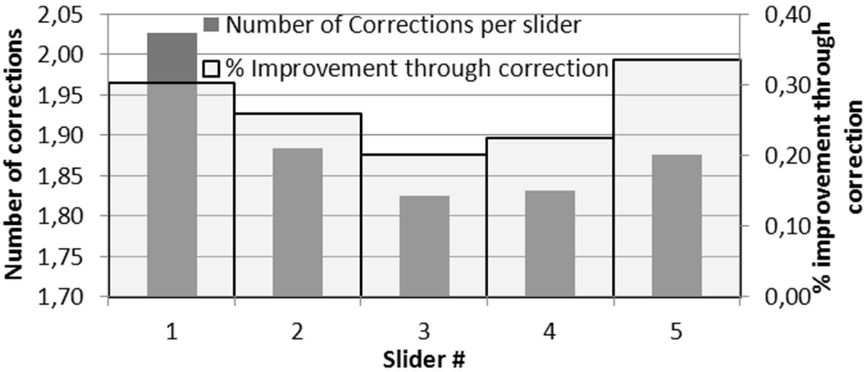


Figure 5-32. Number of corrections per slider, and improvement per correction in percent.

While sliders two to five have been changed between 1.83 and 1.88 times, slider one has been modified 2.03 times. This difference may be seen as an indicator for uncertainty of setting slider one, it may also only be a matter of fact that slider one served as first reference.

First Conclusions

Comparatively corrected results were clearly better than first guesses. The non-comparative input of the first slider was 1.37 times more error-prone than the average error after comparing five sliders.

Results and users' certainty seem to improve when at least four sliders are provided; on the other hand a second slider seemed to raise uncertainty. Consequently, it may be interpreted that in average users needed three sliders to create their own frame of reference and four to gain certainty. Slider four and five showed the best results from start as well as after corrections. It should be further investigated how much this differs for other tests.

Figure 5-29 also shows that corrections helped to improve the absolute result in average by 26.3%, however it did not level the absolute error among sliders, as one might have expected. For more detailed information the distribution of the absolute error among the input values will be studied further.

Detailed Look at the Distribution of the Absolute Error

The basis for this test were the brightness reference values, which were generated randomly and equally distributed between 0 and 1 using the Lehmer random number generator algorithm (Park & Miller, 1988). The distribution of the absolute error of the users input for each control value is shown in Figure 5-33.

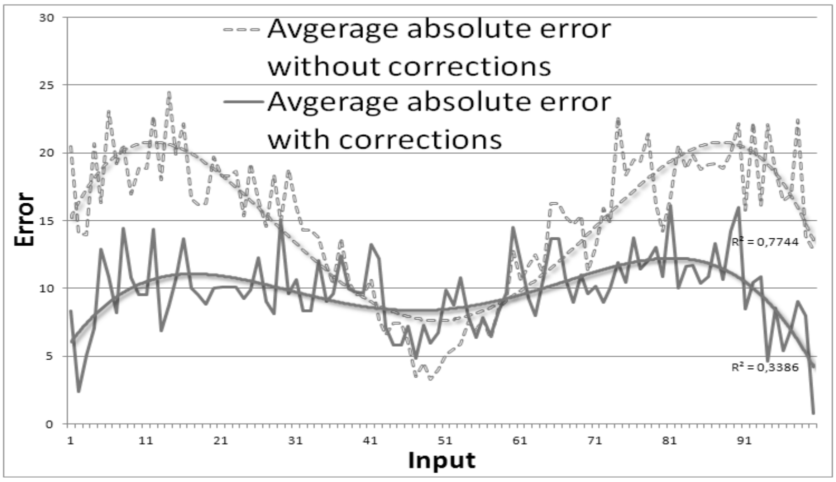


Figure 5-33. User input error for each control value; with and without corrections. Trend lines are of 4th degree.

Both distributions show a minimum error around 50, and an area of maximum error at the intervals of 20 to 10 and 80 to 90. The minimum error around 50 is similar for both inputs, with correction and without correction, while the maxima of the curve without correction are approximately 10% above the maxima with correction. The curve with correction is more flat than the curve without corrections. Both curves show diminished error values at the borders towards 0 and 100, while the corrected error curve in these areas is even below the minimum around 50.

The minimum and the maxima observed in Figure 5-33 might be related to the epistemic uncertainty described by Fischhoff and De Bruin (Fischhoff & Bruine De Bruin, 1999). They found an unexpected 50% "blip" in several surveys, where small probabilities were highly overestimated. This blip may be caused by users when they are not sure about the right answer. They believe that subjects may be choosing the 50% answer in a way that one is using the phrase "fifty-fifty". Such a 50% blip can also be found at this test as shown in Figure 5-34.

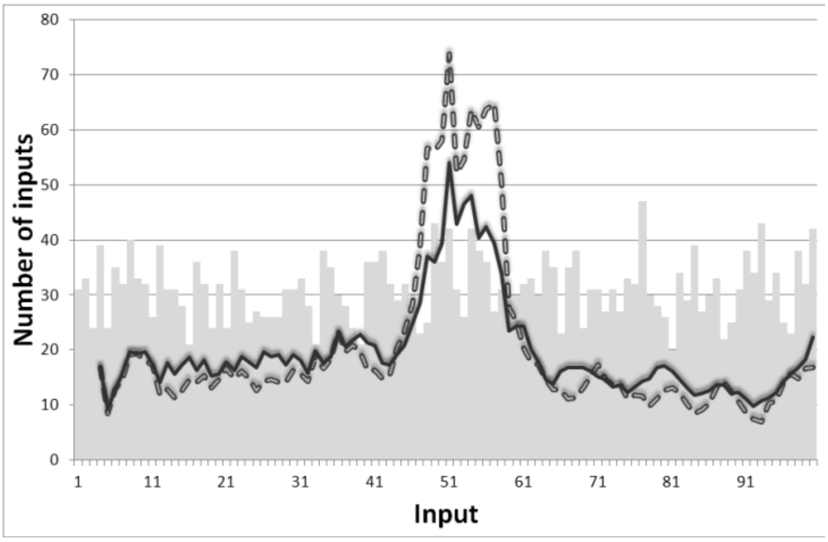


Figure 5-34. Accumulated distribution of user inputs before (stitched line), after corrections (bold line), and control values (gray bars).

The input values clearly accumulate around 50s. The smoothed curves also show that the mean without corrections shows a stronger 50s blip than the curve with corrections. It also shows that a higher 50s blip leads to a reduced input at the outer areas (the bold curve is slightly below the stitched curve in Figure 5-34).

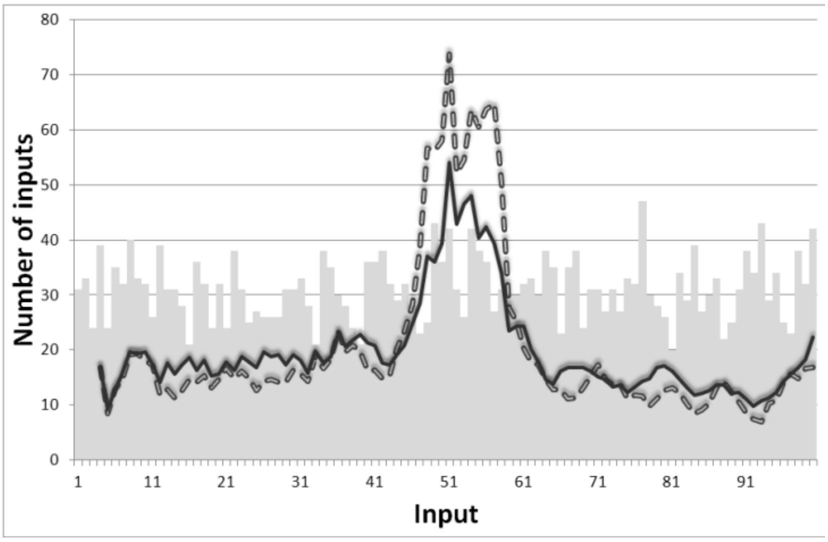

80
70
60
50
40
30
20
10
0
Number of inputs
1
11
21
31
41
51
61
71
81
91
Input

Figure 5-34. Accumulated distribution of user inputs before (stitched line), after corrections (bold line), and control values (gray bars).

Choosing the middle when one is uncertain also has a strategic value, since the probability of being completely wrong is minimized. Such a tendency of selecting an average answer could be the reason for the curves shown in Figure 5-33.

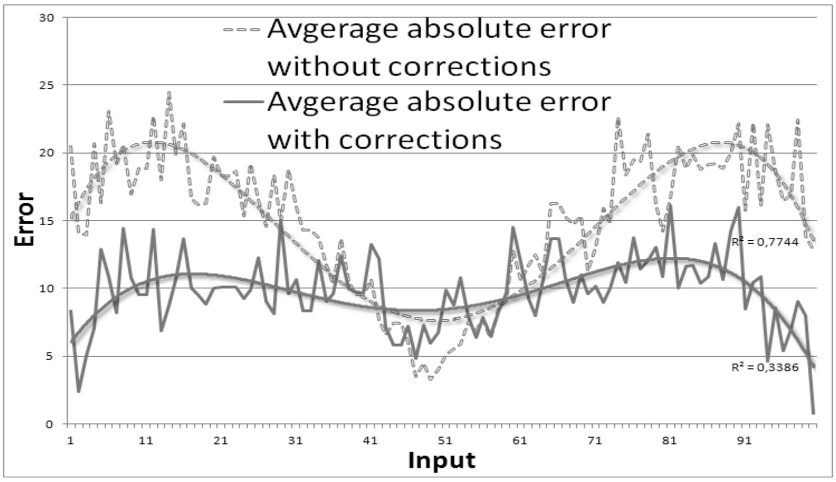


Figure 5-33. User input error for each control value; with and without corrections. Trend lines are of 4th degree.

When users tend to prefer a 50s answer the chance to actually hit a 50s control value is higher. On the other hand the probability of an error for the values further away raises. Based on this hypothesis the uncertainty of the user is clearly shown in the curve without correction. The minimum being below the minimum of the corrected curve underlines this hypothesis.

Further confirmation that the outer error maxima are based on the 50s minimum in the middle is given in Figure 5-35.

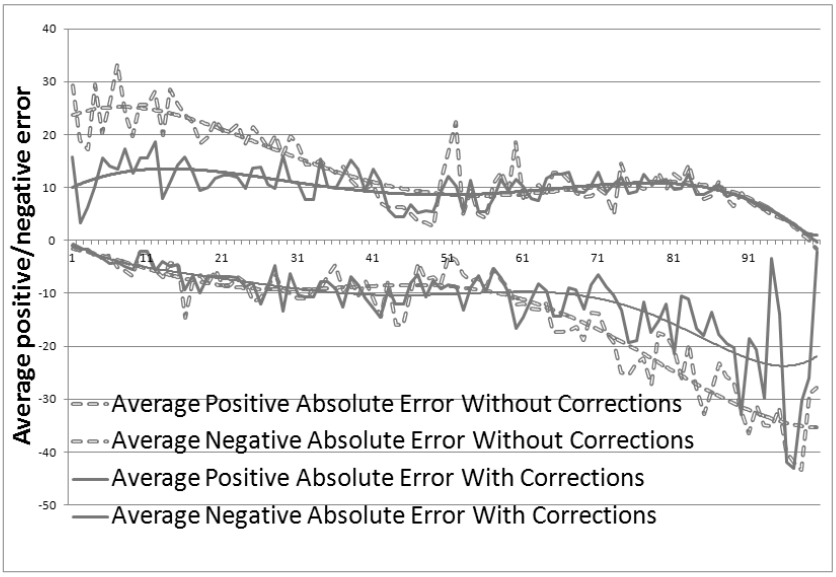


Figure 5-35. average positive and negative absolute error per slider number with 4th order polynomial trend lines.

This diagram shows the direction of the error, visualizing that the errors at the edges are directed towards the middle. The question is on whether such findings are based on a nonlinear psychological distance of measurement (Edwards, 1968). On the other hand the flattening of the curve through correction clearly indicates that it is not a psychological or perceptual distortion.

The decreased error rates towards the edges at 0 and 100 (compare Figure 5-33) may also be based on the fact that a ratio scale for this test was used, which was not open-ended. The darkest and brightest slider was imaginable as black and white, although not visualized. Such imaginable references seem to increase certainty, as it can be observed especially at the corrected input data, where those error values lie below the minimum. Thus, on an open ended scale an exponential growth towards the edges might be expected. On the other hand one might ask the question on how often an open ended scale actually appears in a real evaluation case?

Gained certainty through corrected inputs might also be indicated through a decrease in the standard deviation of both sets. The standard deviation of initial ratings (SD = 12.27) was significantly higher than the standard deviation of corrected ratings. (SD = 7.90); obviously on the same number of samples.

Conclusions

Besides the standard deviation, the minima and maxima in the absolute error curves give further hint on an increased uncertainty with little references. They also visualize how epistemic uncertainty results into a 50% blip complemented by error maxima towards the edges. Giving references seems to flatten this curve and minimize the distributed error.

Detailed Look at the Distribution of the Relative Error

So far the user input referred to the absolute slider errors was analyzed –the user inputs were compared to the brightness of the control values. We will now have a look at the comparative error among pairs of sliders. Therefore, the relative difference of two reference sliders was compared to the relative difference of the corresponding rating sliders. If two reference values are close together, than the corresponding user ratings should be close together as well. Every slider has been compared to every other slider of a set, using the described pair wise relative comparison

The relative error as described above resulted in an overall mean of 11.76%. At first glance it seems surprising that it is slightly higher than the mean of the absolute error with corrections. An explanation could be the lack of the black or white imagined references at each end of the scale. The relative error described above only considers the relative error between the sliders, and does not consider the edges, which give imaginable reference as stated above. Consequently one could also argue that the difference between the absolute error value with correction of 11.65% and the average of the comparative slider error of 11.76% is providing the advantage of the imaginable static references provided through the edges. For a future test it is interesting to see whether concrete references of a black and white square at each end lead to better results.

The distribution of the comparative error in respect to the relative difference between the reference values is shown in Figure 5-36.

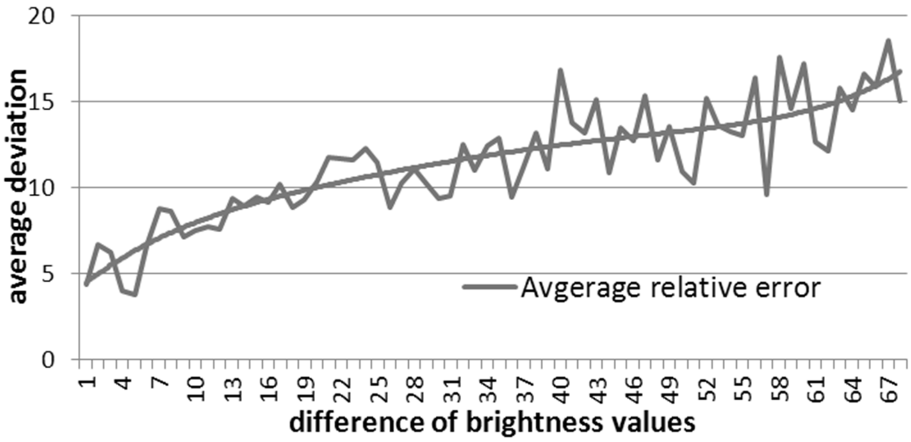


Figure 5-36. Comparative distribution of the error in respect to the relative difference of the control values.

As expected the relative input error is minimal if the reference value difference is small, and it rises for reference values which are further apart. The average relative error also shows the imprecision of the system itself. Even when two sliders had the same color in average users created a relative error of 4.35%. This offset might be due to the user interface or caused by other flaws of our perception. Consequently the slider input accuracy could be adapted to the precision reached by its users, from a slider granularity of 100 to 23.

The first derivative of the trend line in Figure 5-36 also shows a strong increase of the error in the first 20% (increase of 0.34) followed by a little regression between 25% and 50% (increase of 0.024) and another rise afterwards.

Conclusions

Figure 5-36 shows that an increasing amount of references leads to increased test result accuracy. The accuracy gain was greater when the relative reference value difference was below 25% of the whole spectrum. Assessments of reference values with a relative brightness difference between 25% and 50% were characterized by a less steeper error curve. Thus, the relative gain achieved through three references, equally distributed in 25% steps over the whole spectrum, would be comparatively small compared to the gain of a single reference in the middle. A bigger gain can be expected if four or more references are provided, equally distributed, while more than 23 references do not give any winning effect due to the inaccuracy of the setup. These findings correspond with the results stated above, where more than three sliders were demanded for a user to create a valuable reference system.

DISCUSSION AND FUTURE WORK

In the previous text, a very common methodology was reconsidered, from the perspective of today's needs and possibilities. Especially the iterative design process of interactive systems is characterized by multiple small scale user assessments, where quantitative mathematical operations may not be applicable with sufficient significance. Therefore, a very special test case for determining the possible potential of using comparative assessments in such cases was created, giving an indicator for the need of further investigations on that topic. Concerning this, the above tests underline the potential of comparative assessments. Usually non-comparative tests are based on a scale without concrete references, as it is for example known from the Likert scale. Such a scale seems to leave users with a particular amount of uncertainty. In case of the above tests non-comparative assessments have been represented by the first slider which was set without concrete references. On the other end of the spectrum five sliders have been set comparatively, finalizing each run. The error of the first non-comparative slider was 27% higher than the average error of the comparative set of five sliders. Following the investigations above, this difference may majorly be based on increased uncertainty of the assessors.

Uncertainty was observed through various factors. The number of corrections of the first slider was 9.2% higher than the adaption of the other sliders. The tendency to choose the middle is a strong indicator for uncertainty. It implicitly led to a minimized error in the area around the fifties alongside with a maximal error in the outer area.

A second slider increased the error rate, and therefore seemed to increase uncertainty. Four or more comparative sliders lead to an error rate which was below the non-comparative rate of one slider. Certainty seemed to be gained with more than three references. This observation was supported by the fact that users in average took most time to reconsider former sliders after the fourth slider has been set, while consideration spans of the first three did not differ much.

Although results were not expected to be very general, some findings turned out to be more remarkable than expected. The gain of certainty through a raising number of references, as described above is very likely not to be limited to this special test case. Another remarkable finding, not necessarily reduced to this special case, was the fact that users improved their results through adjustments. Based on the error rates users seemed to gain certainty by seeing multiple different sliders. Corrections and adaption of such sliders among each showed to improve results. Also the tendency to choose the middle was reduced after correction, which describes another indicator for gained certainty. Since these tests were based on imaginable references, assessments right at the edges showed better results compared to the rest. This might differ with an open ended scale. However, it seems that such effects may also occur in more complex cases.

For future work it would be interesting to investigate user ratings on an interval scale, such as the frequency of sound. Moreover, in future tests one could let the users decide on their demanded amount of references. This way one gets a deeper insight into the users' personal scaling.

The initial hypothesis, that users perform better when they judge in comparative ways, is supported by these outcomes. Although some results were unexpectedly already found, giving hint for general guidance, most of such questions need to be verified through more complex test cases. In order to start this process the most remarkable findings will be transformed into first guidelines.

Comparative Assessment guidelines

Use comparative measures to get comparative results!

Finding: Abstraction demands extra work load otherwise leading to high uncertainty. Do not provide an abstract description of a target domain. Let the user work on concrete items. It should be the work of the interviewer, and not the work of the interviewees, to abstract their assessment.

Guidance: Instead of "How much did you like the system on a scale from 1 to 10". Rather ask: "Compare the following five systems and rate them relatively to each other"

Encourage users to correct their input!

Finding: When users correct their inputs, the error curve is flattened and the 50s blip is reduced.

Guidance: Design the evaluation in a way that users are able to correct formerly set values.

Provide a certain amount of references! Finding: References are leading to more precise ratings. Guidance: Provide enough examples to sufficiently represent the target domain.
Let subjects determine the scales' scaling! Finding: An imaginable minimum and maximum helped improve the result. Guidance: Ask subjects for their personal min and max values of the domain – corresponding to the imaginable black and white of this test. This gives further indication about the users' frame of reference. The personal min and max value helps with scaling and comparing different inputs among users.

Table 5-1. Initial comparative assessment guidelines.

5.3 Generation 3 – Cooperative Cognitive Prototyping Tools

The third generation of tools aims at extending the methodological support of the Digital Transformatives framework (DT framework). The DT framework builds on online tools, allowing for identifying and gathering cognitive structures of possible end-users. Therefore, all tools are implemented as web services, using the Seaside web development framework and Pharo programming environment (Ducasse et al., 2010; pharo-project.org, 2012; Seaside.st, 2012). Additionally, they make use of the Bootstrap Framework, for offering a common look and feel, and database back ends based on MySQL or MongoDb (mongodb.org, 2013; Oracle Corporation, 2013; Twitter, 2013). All tools further implement user management. Users are able to log in, be invited, and related to each other.

In the following a brief description of web based implementations of the two developed methodologies of Sub-Action Modeling, and Salient Super Modeling will be given, also including a set of guidelines based on first experiences with such implementations. Hereby, the designs and implementations of the two tool are different. The initially described Sub-Action Modeling module was conventionally designed and implemented, through an iterative design cycle addressing the original problem context. In contrast, the design and implementation of the Salient Super Modeling module, exemplifies DT specific design and development features by focusing on cognitive prototypes.

Finally, another tool is described in form of a speed reading game. The tool is supposed to provide a module for generally extracting cognitive structures of users – not related to a specific use case.

5.3.1 DT Web Module Sub-Action Modelling

The following Web Module provides a cooperative interface, which implements the sub-action modeling approach, as described in detail in chapter *Sub-Action Modeling Approach* (pp. 261), demands users to cooperatively create, and refine action chains. The development of the according web module followed a typical design cycle. Initial paper prototypes were used to perform first tests. While a first implementation offered a cascade based interface for modeling actions in a directed graph, this procedure turned out to be too complex. Hence, in further refinements, the user interface was adapted, letting users model actions hierarchically through bullet point lists, as shown in Figure 5-37. Bullet point lists are commonly known, but typically not sufficient for modeling action chains, especially not if such chains include alternative ways. Since the modeled action chains are split up in later phases of the Sub-Action Modeling procedure, users are not required to model complete lists. Hence, users are asked to model simply a part of choice, rather than requesting a complete action model of a certain task.

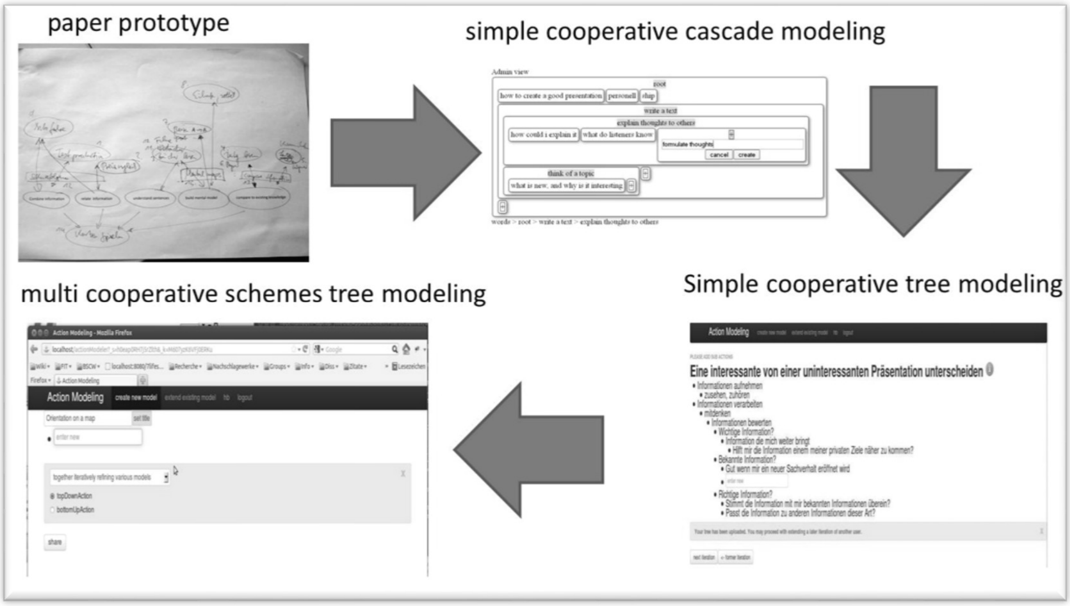


Figure 5-37. Iterative development of the Sub-Action modeling web module, starting off with paper prototypes (top-left), an experimental online interface showing cascaded action lines (top-right), and the final hierarchical bullet point based interface for modeling incomplete action chains (bottom).

The web interface allows users to move, modify, extend, or delete branches of the action chain bullet point tree. Modeling happens in two basic modes, top-down or bottom-up. Top-down modeling starts with an action, and requires users to determine sub-actions. The Bottom-up modeling mode implements the second part of Sub-Action Modeling. Therefore users need to find concrete tasks that integrate a given set of sub-actions. A major requirement of the determined approach is distinct user groups, for the two different modes; bottom-up and top-down.

Since it is not clear how users cooperatively create such action chain models in the best way, several forms of cooperation are provided.

1. Simple cooperative modeling, where all users model on a single tree. This approach demands a lot of communication, in order to agree on a final solution.
2. Cooperative cascade modeling, where various versions of action chains in the same context are created. Users start with creating root models of an action chain, and later extend the models of others. Those models can be modified, and refined, resulting in numerous different alternatives, which need to be concluded manually.
3. Disjunctive iterative modeling, where different users refine the next iteration of each other. Leading to several final iterations.

First experiences of the web module are summarized in the following table.

challenge	experience
Finding super actions is harder than finding sub actions.	Ask users for concrete examples for each sub action. Optionally others might start from those examples and determine sub actions again (Starting a new iteration cycle)
Similarity of new performance context actions depends on depth of level of modeled sub actions, or the number of consecutive iteration cycles.	The more basic an action, the more general it is. Using very fundamental sub actions as seeds might result in completely different concrete examples

Table 5-2. Experiences of the action modeling web module

5.3.2 DT Web Module Salient Super Prototype Modeling

The Salient Super Prototype Modeling web module implements the procedure described in *Salient Super Prototype Identification Approach* (pp. 175). According to those explanations the tool seeks to span a salience-similarity space of end-user concepts, related to source actions. Hereby, concept salience correlates to user response immediacy, performance, or typicality, and to similarity of concepts, which gradually decreases with rising generalization. Those prerequisites set a coarse frame for the tool:

I. Users need to input concepts.
II. Inputs should be made immediately, possibly under time pressure.
III. Inputs should be of high typicality.
IV. Similarity is reached by implementing the procedure as described in chapter 4.3.3 (starting from pp. 178).

<u>COGNITIVE PROTOTYPE ORIENTED SYSTEM DESIGN</u>
To add some DT elements, the design of this web module will start off with the consideration of salient concepts addressing the aspects given above. Moreover, the web tool should be fun to use. Therefore, the author spontaneously determined typical cognitive prototypes for each aspect, as outlined in the following:

Prototypes addressing previously listed aspects:

I. Typical fun actions involving user input are
   - communication via instant messaging,
   - social networks, or
   - micro-blogging.
II. Typical fun actions requiring time critical performances are
   - competitions,
   - races,
   - games of skill.
III. Typical fun actions requiring good examples of a kind are
   - guessing games, such as the board game Taboo (Wikipedia.org, 2013d), or
   - the design of icons or signs.

Aspect IV describes a prerequisite that needs to be fulfilled on a procedural level, not a user interface level; hence it is not considered here.

The above list provides ingredients for a cognitive prototype driven design of the web module. The ingredients could be combined in several ways, of which the author chose to address the three demanded aspects of, instant messaging, competition, and guessing games. In the next section, the ingredients chosen by the author will be combined in a system concept.

The previous procedure showcases the basic principle of cognitive prototype based digital system design. In this example, the author determined possible prototypes simply through immediacy. The identified prototypes may be extensively varying among persons; hence, more elaborate designs should capture concepts directly from possible end-users. This highlights the importance of methodologies and tools, in order to get a more precise shared cognitive prototype structure.

<u>Salient Super Concept Modeling Game Concept</u>

The above ingredients suggest a system implementation in a game context. According to this, the user input element should follow the design of a social network activity stream. Contrarily to a common activity stream for direct communication, its elements should be optimized towards communication in a guessing game, where users implicitly communicate with each other. The according user interface is displayed in Figure 5-38.

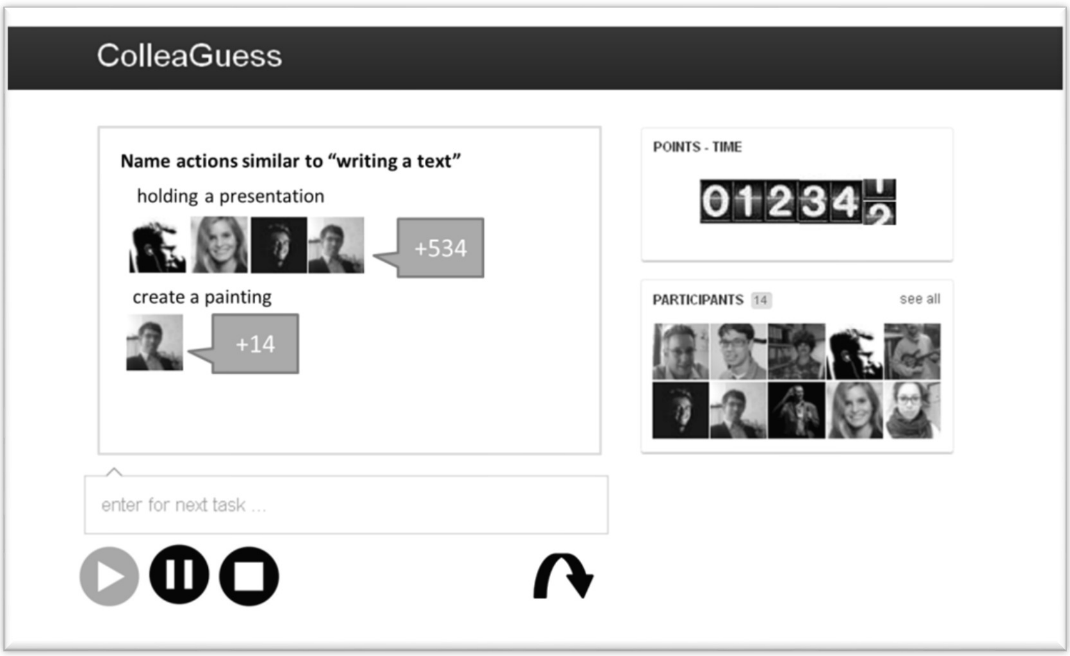


Figure 5-38. ColleaGuess user interface consisting of an activity stream, and competitive game elements,

In order to start the procedure, respectively play the game, users need to click on the play button. As soon as the button is clicked, a task is displayed in the activity stream, such as "name actions similar to writing a text", and players are required to provide answers. Whenever such answers correspond to answers of other participants' all players with matching answers get points, Herby points correlate to the amount of matching answers. An answer is more valuable if it is given by a higher amount of co-players. This game element fosters typicality, since users try to find answers possibly general for all other participants. Typicality is an important feature of salient prototypes. Another feature promoting salient answers is immediacy. Therefore, the score also works as a count-down timer. This way, players are urged to answer quick and immediately, in order to increase the score. Players are also able to pause the game, which also pauses the count-down of the score, and empties the task pane. They may also skip the current task, if they desire. Moreover, like in social network communities, it is possible to create new groups or join different existing one. For the underlying purpose of capturing a salient cognitive concepts structure, it is important to setup possible end-user groups. While the user interface elements ensure the basic aspects required for salient super concept modeling, the underlying procedure, schematically shown in Figure 4-24 (p. 181), also needs to be implemented. With some additions, this procedure can be realized through the game logic illustrated in Figure 5-39.

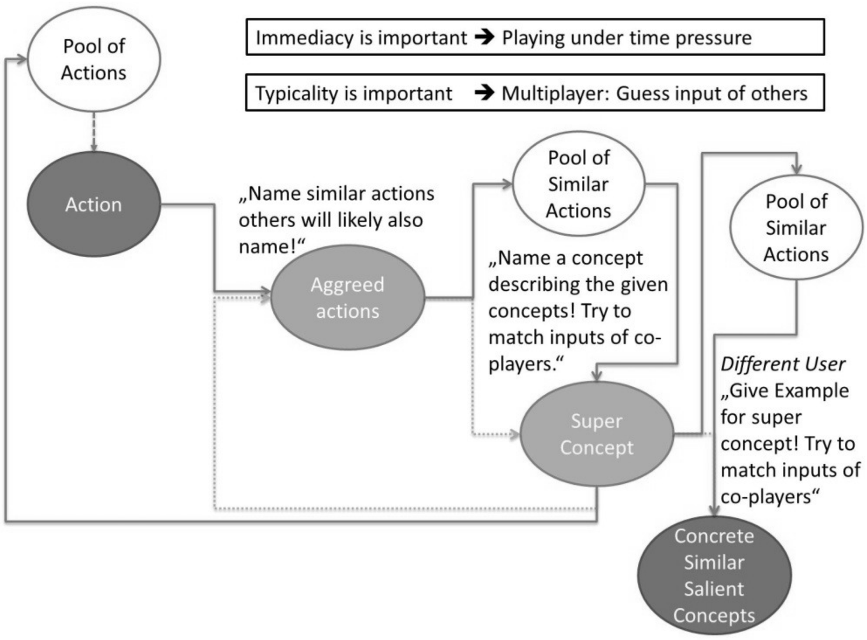


Figure 5-39. The game logic adapted from super modeling approach procedure shown in Figure 4-24. Gray dashed lines indicate transitions of the original procedure.

The game logic describes the different types of tasks, and how those fit into the original procedural concept. The tasks are provided from several pools, which are filled during the game. The initial pool of actions may be filled by system designers, or end-users using the sub-action modeling module, described earlier.

The game logic in combination with the user interface addresses all four aspects underlying the salient super concept modeling system. The implementation of fun elements should furthermore foster the usage of the system. Besides a high-score list, the module also provides feedback on cognitive affinity among colleagues, which might be another interesting feature raising motivation for using the system.

More importantly the module delivers a merged conceptual model of all user-inputs. Enriched with general salience information from search engine query amounts this model should provide a valuable basis for further developments.

CONCLUSIONS – DT MECHANISMS AND GUIDELINES

Since this digital system followed cognitive prototype driven design principles of Digital Transformatives, the transformative mechanisms can be summarized in the following schema.

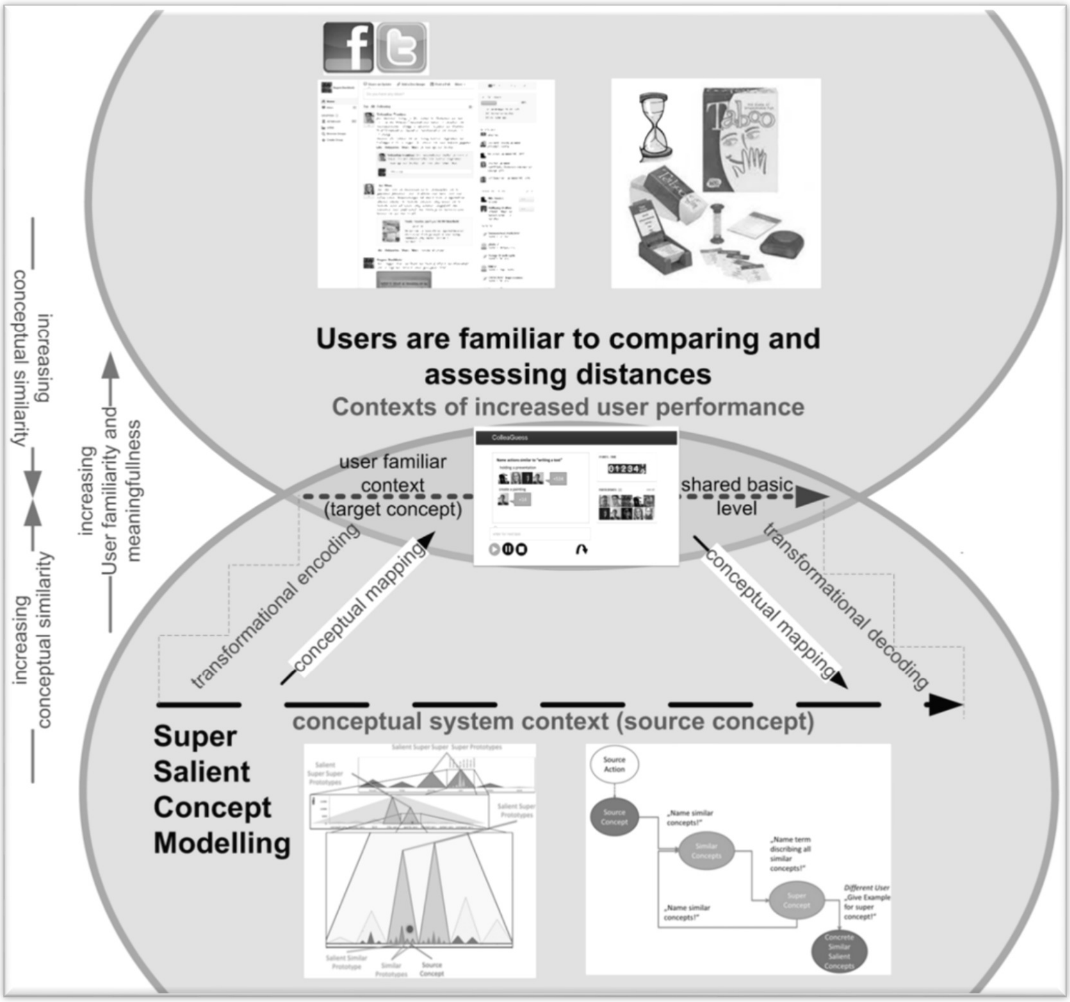


Figure 5-40. Transformative mechanisms of the super salient concept modeling game.

The user interface provides a shared basic level of cognitive prototype concepts. The source system context demanded user input of concepts. The input context was addressed by a social communication inspired interface. Moreover, the source context demanded salient and typical concept inputs. Therefore, the prototype of a guessing game was implemented to acquire typical user answers under time pressure. All was combined through social competitive game rules, where points are gained when participants of a community agree on inputs. Competition and team play are prototypes for fun, both included here.

First tests during the development of the system also led to the following guideline:

Challenge	Experience
End-Users are usually not designers. Their cognitive concepts for design are not elaborate.	Instead of providing tasks in a design context, it is advantageous to let users operate in an action context. Cognitive action concepts are more salient and prototypical in end-users minds. Taking, for example, the design of the craftsman's cup (introduced in chapter 4.3.2, pp. 172). Instead of starting with the source concept of a handle, which allows for hammering, it is better to activate the cognitive concept of the action of hammering. Hence, the first users need to determine actions similar to hammering, instead handles similar to a hammer handle.

5.3.3 Reading Speed Test

While the previous tools were mainly created for capturing the conceptual model of possible end-users in a specific context, the following tool aims at a casual method for cooperatively modelling concept structures more generally.

The tool is based on the comprehensive investigations and discussions of the influences of language during the cognitive development of concepts. For example, the linguistic relativity hypothesis led to multiple studies and discussions on the importance of language for our understanding of the world (Berlin & Kay, 1969; Davidoff, 2001; Derose, 2005; D. Gentner & Goldin-Meadow, 2003; P. Kay et al., 2009; Saunders & Van Brakel, 1997; Whorf & Carroll, 1956; Woodbury, 1991). While language surely influences knowledge construction, in turn, it also provides a good representation of conceptual shapes in our minds. In this context, studies on reading provided a profound reflection on learning processes, automatization, and chunking (G. D. Bower, 2008; LaBerge & Samuels, 1974; LaBerge, 1975, 1976; Samuels et al., 1978). Moreover, several studies show an improvement of reading speed with rising age in the phases of cognitive development, as visualized in the diagram in Figure 5-41 (Artelt, Naumann, & Schneider, 2010; Buswell, 1922; Gilbert, 1953; Grissemann, 1981; Hunziker, 2006; Lefavrais, 1967; Linder & Grissemann, 2003; Mayringer & Wimmer, 2003, 2008; Stanford Earl Taylor, Frackenpohl, & Pettee, 1960).

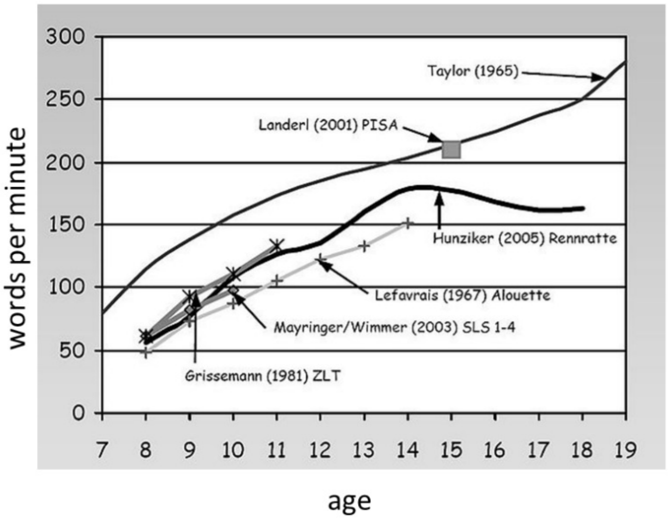


Figure 5-41. Diagram visualizing results of multiple studies on the relation of reading speed and age, from (Hunziker, 2006).

Test persons with better reading performance also achieved better text comprehension, and showed fewer regressions in their eye-movements (times words had to be re-read). Other studies on reading speed and comprehension underline further correlations with expertise (Artelt et al., 2010; Hunziker, 2006; McNamara & Kintsch, 1996; Stanford E. Taylor, 2006). Since expertise is context dependent this also indicates a relation between the topic context and its reading speed.

To get further insights a small scale online test with 8 subjects has been conducted on reading speed. Four of the subjects were female and the age ranged from 25 to 58. In the test readers were asked to read similar texts in their mother tongue, German, and in the foreign language English. The reading test was succeeded with comprehension questions, to validate proper comprehensive reading. The results indicate an approximately 30% higher letter processing speed of the test subjects in their mother tongue.

For further indications, the first prototype of a simple reading speed test has been improved to be more appealing for users. Therefore, it was transformed into a speed reading game, as shown in Figure 5-42.

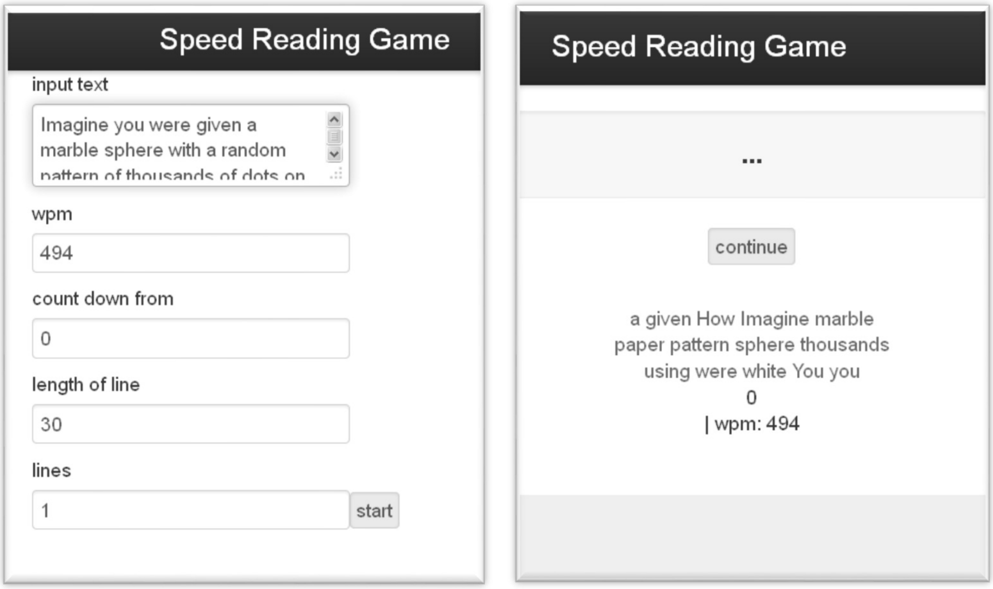


Figure 5-42. Screenshots of a cooperative online speed reading game for capturing cognitive concept structures. Session setup interface (left) and in game interface (right).

The game can be played in training mode, to practice speed reading, or in a competitive mode against friends. During the game a selected text is presented sequentially subdivided into small chunks. Each text chunk flashes up for a certain time period, disappears again, and a word cloud of the text shows up. From this word cloud players select the words they recognized in the previous text chunk. The display time of the text chunks is adapted instantly during the game, according to the error rate. Prior to each session users are able to select the input text, and set the starting values of the words per minute related display time, the length of the text chunks, and the number of rows shown at once.

The multiplayer mode allows for competition between multiple friends. Therefore, one player needs to initiate a session by selecting a text, and sending the session URL to possible participants. Each invited participant may join the session by reading the given text, while the final score is dependent from the words per minute (WPM) value and the error rate.

Besides the motivation of practicing speed reading, a high-score, and the competition should raise interest in using the system. First tests showed that reading speed differed clearly for single readers among various subject domains. Reading speed and error rates performances were much better on user familiar domains. Therefore, it is strategically advisable in multiplayer games, to setup sessions with texts of familiar domains, which are ideally unfamiliar for the competing readers. Hence the multiplayer mode fosters the tool to be used in a broad variety of subject domains. As a result a broad domain based model of cognitive concepts of each participant is generated.

For future work, the speed reading game should be tested and evaluated in large scale. However, first indicators, combined with evidences on reading speed from other studies, show the potential of this casual game for generally capturing cognitive prototypes.

5.3.4 Conclusion – How to Create Tools for Capturing Cognitive Concept Structures

This section briefly gives a general overview on how to create tools for capturing cognitive concept structures. It will be a summarizing guidance, concluding experiences based on the practical design of cognitive tools, theoretical knowledge about cognition, and the methodological developments for Digital Transformatives. The following summary should help with the creation of tools for capturing cognitive conceptual structures in a salience similarity space.

The required salience-similarity space demands

- representations of concepts,
- their relative salience,
- and relative similarity distance to each other.

CHOOSE PROPER REPRESENTATIONS OF CONCEPTS

Today's digital interfaces allow for addressing cognitive concepts of various types. Therefore, one should consider the closest common interface mechanism to user actions involved in the targeted task. Hereby, linguistic terms provide valuable representations of concepts, and are easily captured and analysed. Moreover, aural and visual information are also common digital in- and output mechanisms. Even haptic interfaces in form of force-feedback devices or 3D printers are available. However, of the named representations only linguistic terms can provide anchors to all kind of concepts, while the other representations are often mainly addressing concrete concepts.

After the concept representation, and an appropriate interface is chosen, it is necessary to determine salience of user concepts.

DETERMINE RELATIVE SALIENCE

In order to determine relative salience of captured concepts, it is advantageous to measure user performance of involved cognitive concepts, since salience corresponds to efficiency. Usually time is a good measure for performance. In practice, it is helpful to identify tasks involving the targeted concepts, such as recognizing colors, sounds, haptics, or terms, and measure performance times for those. Alternatively, one could capture comparative user ratings for familiarity, to get further refinements.

DETERMINE RELATIVE SIMILARITY DISTANCE

As the work on super salient prototype modelling shows, similarity distances of concepts can well be approximated through cognitive categorization. Hereby it is important to keep in mind that cognitive categories are not as well defined as the commonly understood categories. They change dynamically over time and context. Similarity distances can be approximated by integrating similarity and generalisation requests into the tool.

CONCLUDING GUIDELINES

The above explanations are concluded in the following guidelines:

Challenge	Experience
Choose proper concept representations	Linguistic terms provide valuable representations of cognitive concepts. Additionally, one should consider the closest common interface mechanism, to the requested concepts, may it be aural, visual or haptic.
Determine Relative Salience	Implement a task involving the addressed cognitive concepts and measure user performance. Usually time is a good measure for efficiency. The quicker the performance the more salient the referring concept. Also, let users comparatively rate for familiarity or typicality for further refinements of the structure
Determine Relative Similarity Distance	Implement request for similar concepts, or generalisation tasks to approximate similarity distances

Table 5-3. Guidelines on how to create tools for capturing cognitive prototypes.

6 CONCLUSION, CRITICAL REFLECTION AND OUTLOOK 253

6.1 DIGITAL TRANSFORMATIVES SUMMARY 253
6.1.1 Basic Schema 253
6.1.2 Background and Working Principle 255
6.1.3 Designing Digital Transformatives 258
6.1.4 Use Case Prototypes 267
6.1.5 Function Characteristics and Design Guidelines 268

6.2 FUTURE WORK – COGNITIVE PROTOTYPES BEYOND INTERACTIVE SYSTEM DESIGN 273
6.2.1 CPs at Conventions, Standards, Norms, and the Force of Habit 273
6.2.2 CPs in Perception and Recognition 275
6.2.3 CPs Help Promoting an Image 277
6.2.4 CPs on Pleasantness and Affection 279
6.2.5 CPs in Communicating and Innovation 281
6.2.6 CPs in social structures 282

6.3 NEW CHANCES TO ANALYZE AND UTILIZE COGNITIVE PROTOTYPES 283

6.4 CRITICAL REVISION 284
6.4.1 DTs might Hinder Personal Development 284
6.4.2 Failures through the Force of Habit 284

6.5 WRAP-UP AND FUTURE WORK 285
6.5.1 Research Questions 285
6.5.2 Outcomes Beyond Research Questions 285
6.5.3 Outlook 286

6 Conclusion, Critical Reflection and Outlook

The first part of this chapter provides a summary of the previous text. Afterwards, most significant findings are reflected in the context of interactive digital system design. The final outlook starts with a perspective on the potentials of cognitive prototypes beyond software design, and how new digital methods may help with the identification and exploitation of such potentials.

6.1 Digital Transformatives Summary

This work was initiated by the observation that mnemonic devices release hidden user intrinsic potentials. Hereby, users transform a given task in a different context of increased efficiency, where they achieve remarkably better performances despite the extra transformational effort needed for encoding and decoding information. Motivated by such cognitive tools of enhanced efficiency, this thesis aims at establishing a new class of interactive digital systems, which utilize the same basic mechanisms for improving usage efficiency. Since the transformation is conducted digitally, those systems are called Digital Transformatives (DT).

6.1.1 Basic Schema

The basic schema outlines the class and helps with the identification of DT systems[24].

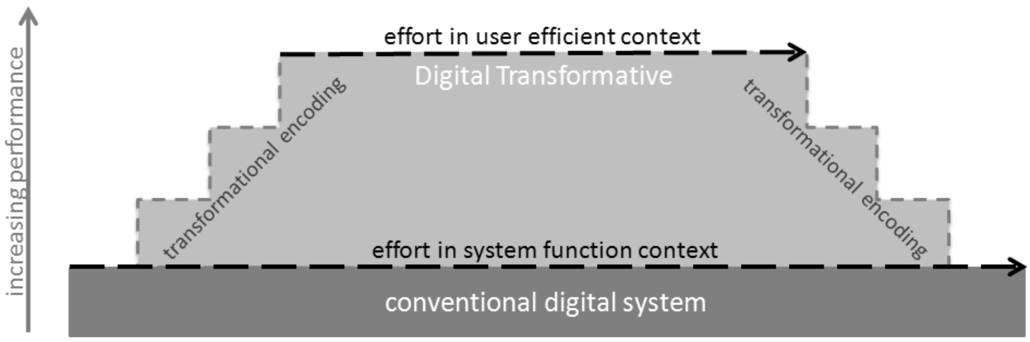


Figure 6-1. Digital Transformative basic schema for reducing user effort by offering an interface in a user efficient context. Information encoding and decoding becomes part of the design, and needs to be implemented digitally by the system.

[24] further elaborated in chapter 1.5.2 *Basic Schema of Digital Transformatives* (pp. 8)

While conventional digital systems often offer interfaces in a functional context, DTs transform user interaction to areas of high cognitive performance, as visualized in Figure 6-1. Hereby, the DT user interface may be situated in a completely different usage context.

This working principle corresponds to the one of mnemonic devices. However, in contrast to those, DTs additionally relieve the user from any transformational cognitive load, since context shifts are implemented digitally by the system. Therefore, the transformation should already be considered in the design phase, to promote information encoding and decoding between the new user interface context and the original function context.

Prominent examples meeting this schema are passphrase pattern inputs or the LOGO Turtle, shown in Figure 6-2 (Logo Foundation, 2000; Meacham, 2013).

Figure 6-2. Examples for DTs: Android Pattern Lock (left) and Papert's LOGO Turtle (right) (Logo Foundation, 2000; Meacham, 2013).

The pattern number lock provides users with a visual input interface, which internally maps patterns onto number codes. There is no need to memorize or even think abou abstract number codes. Users instead work with shapes or paths, which appear to be more natural.

The same applies for the idea of learning with the LOGO Turtle. The LOGO Turtle internally consists of a programmable computer with the appearance of a movable object. Learners are able to "talk" to this turtle in a programming language. This way they can teach it behavior, such as how to move on a certain path. While the learners get the impression to talk to a turtle, functionally they are programming a computer. Further examples are described in section 2.2.

6.1.2 Background and Working Principle

The basic schema helps with identifying DTs, however, more detailed knowledge about the characteristics and mechanisms of high performance contexts is necessary for fully understanding, and finally designing DTs.

<u>BACKGROUND ON COGNITIVE EFFICIENCY</u>

A review of evidence based research on mnemonic devices and associative cognitive processes, such as conceptual metaphors, cognitive categorization, and semiotics, highlights the importance of cognitive prototypes (CP) for this work[25]. CPs define areas of high cognitive efficiency. The formation of CPs is related to chunking and automatization. In order to understand cognitive efficiency in communication, it is necessary to investigate shared cognitive models. Therefore, the term of *shared basic levels* is being introduced. A shared basic level describes a salient prototype concept shared among different communicating parties[26].

<u>RETROSPECTIVE STUDY OF EXISTING SYSTEMS</u>

An evaluation of the cognitive findings against existing systems further showed the importance of the two properties of meaningfulness and increased familiarity. Hereby, a retrospective property extraction revealed that the two properties were common to all context shifts conducted by the investigated systems[27]. Consequently, they can be considered as fundamental characteristics of Digital Transformatives. Other features did not occur in all context shifts. The relative frequencies are visualized in a tag cloud shown in Figure 6-3.

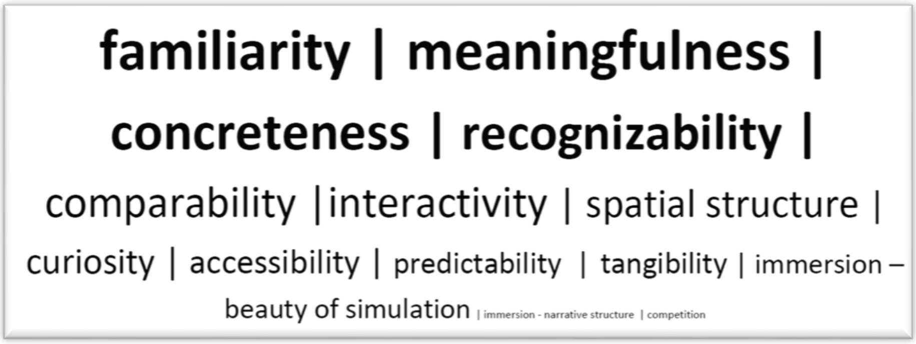


Figure 6-3. Context shift property tag cloud sorted by frequency. Familiarity and meaningfulness appeared in all context shifts of the investigated DT systems.

[25] 2.1 *Relevant Cognitive Mechanisms* (pp. 14)

[26] 2.1.7 *Cognitive* Efficiency Catalysts in Communication (pp. 48)

[27] 2.2, *Retrospective Property Extraction of Existing Digital Transformatives* (pp. 57)

Those results are in line with findings on the cognitive background[28]. The colloquial term "familiarity" meets a commonly understood concept, which corresponds to cognitive prototype categories and processes in practice, as evaluated in this thesis[29].

Model for Cognitive Efficiency in Communication

The results are summarized in a concept for cognitive efficiency in communication, shown in Figure 6-4[30].

According to this model, efficiency in the communication of new knowledge from speakers to listeners is dependent on the conceptual mapping between base and target concepts. If speakers want to communicate new target concepts to listeners they may utilize a base concept, which is known and interpreted by the listeners.

Hereby, efficiency is dependent on the similarity of target and base concepts, and on the cognitive salience of the utilized shared base concept. This suggests the use of shared basic levels with certain similarity to the target concept, for most efficient communication.

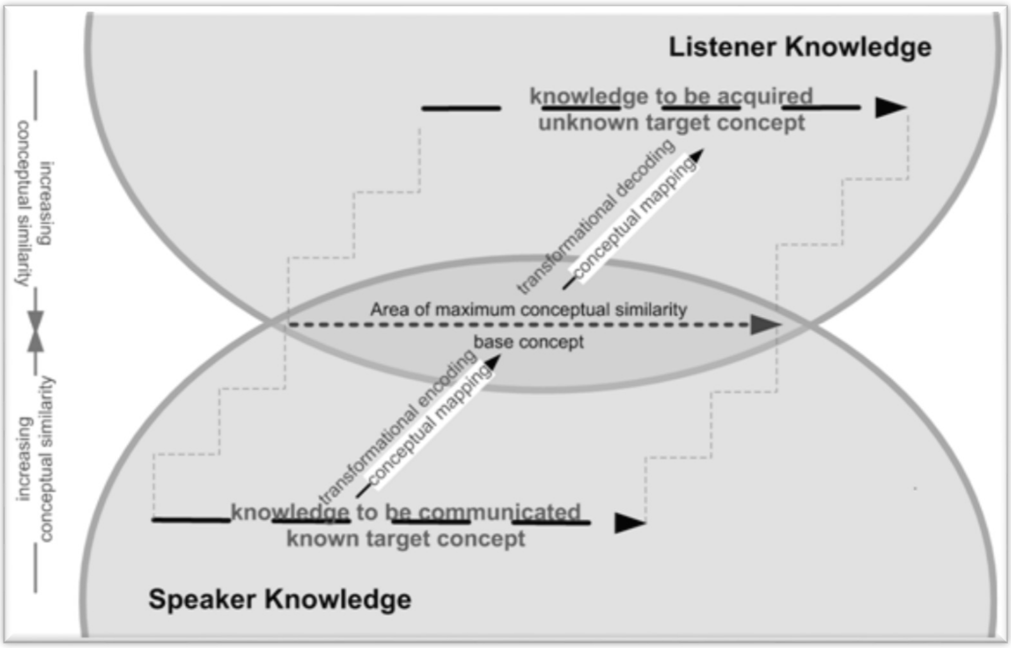


Figure 6-4. Efficiency model of human communication. Deduced from evidential cognitive research.

[28] 3.3. *Digital Transformative Main Characteristics* (pp. 116)
[29] 3.4. *Basic Concept Validation* (pp. 119)
[30] 3.1. *Concept of Cognitive Efficiency Drivers in Human Communication* (pp. 110)

WORKING PRINCIPLE OF DIGITAL TRANSFORMATIVES

This efficiency model in human communication can be further adapted towards a DT concept model, shown in Figure 6-5.

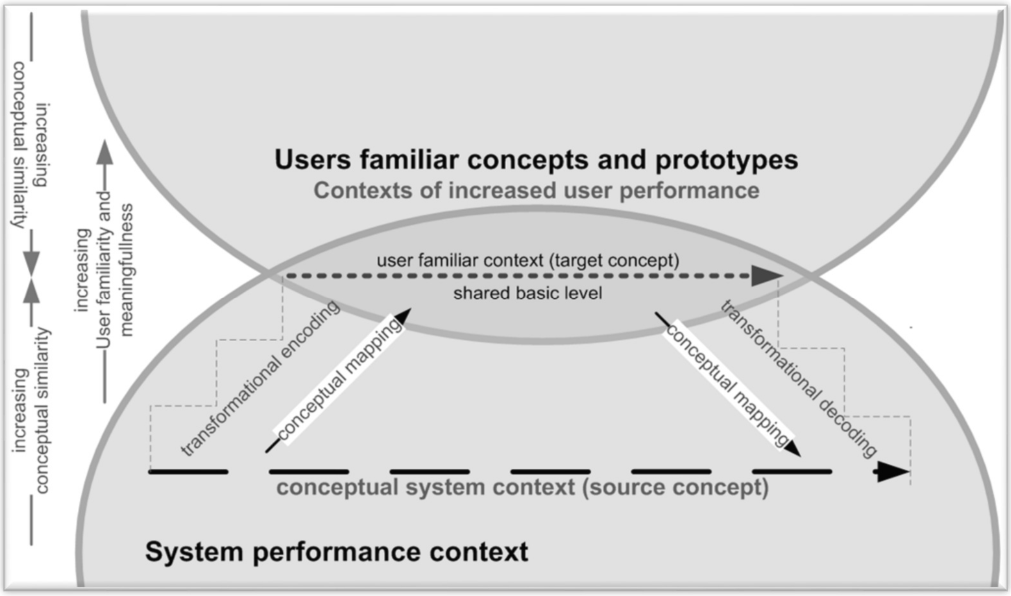


Figure 6-5. Schematic representation of the Digital Transformatives concept and its working principle.

According to the given concept, context shifts are implemented through the user interface, which offer bidirectional conceptual mappings from original function contexts to highly efficient usage contexts, through transitional encoding and decoding. The user interface resides at a cognitive shared basic target concept, achieving maximum user familiarity and sufficient similarity to the original source concept.

In principle, Digital Transformatives cognitively work analogous to cognitive categorization or metaphors in communication. An essential requirement for a successful implementation of conceptual mappings is similarity between base and target concept. Hereby, increased similarity of two concepts seems to correlate to a greater amount of shared salient features.

The details of this working principle become clearer if we consider them on a level of cognitive concepts. Therefore it is necessary to approximate our cognitive structures through a model of concepts arranged in a salience-similarity space[31]. The salience-similarity space represents our cognitive structure, visualizing salience values, and the relative similarity of concepts. Cognitive prototypes are recognizable as peaks, the

[31] 4.3.2 *DT Design Challenge* (pp. 180)

more similar two concepts, the closer they are on the similarity plane, as depict in Figure 6-6.

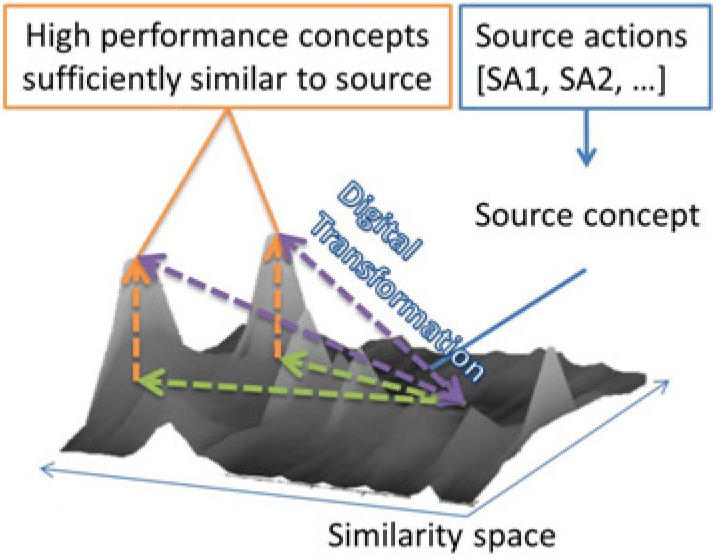


Figure 6-6. The working principle of Digital Transformatives referred to the visualization of a salience-similarity space of cognitive concepts.

In this visualization, conventional systems address concepts in close similarity to the original, usually function related source concept[32], while DT design aims for nearby super salient cognitive prototypes, which typically mark areas of high cognitive efficiency.

Consequently, the design of Digital Transformatives initially requires the identification of relevant prototypes with sufficient similarity to the source concept[33]. Similarity comparisons are also active in the innate cognitive processes of categorization[34]. This suggested the development and implementation of methods for extracting cognitive prototype categories, in order to span the salient-similarity space, as described in the next section.

6.1.3 Designing Digital Transformatives

A major challenge, during the design of interactive digital systems, lies in the provision of a proper user interface. Therefore, system designers typically first analyse

[32] Source concepts are cognitive complements to source actions, which may be determined through initial task analyses.

[33] Section *Similarity Measures for Prototype Categorization* (pp. 33)

[34] 2.1.6 *Pervasiveness of Similarity Comparisons* and 2.1.7 *Cognitive Efficiency Catalysts in Communication* (pp. 45)

relevant user actions to provide interface solutions directly addressing demanded actions. During such a design procedure, the given problem context predominantly activates salient concepts within the designer, in close similarity to the source concepts. Hereby, based on previous experiences, system designers almost automatically tend to think of common interface solutions[35]. Those direct associations possibly hinder them from finding further cognitive prototypes, with close similarity and superior salience. Creativity techniques, such as Brainstorming, may get our thoughts beyond those obvious concepts. Unfortunately, such techniques also produce some sort of noise, resulting in strong random concepts, which are often completely unrelated to the source concept. Moreover, those concepts may outshine good candidates.

Therefore, based on research on categorization and cognitive prototypes[36], two approaches have been developed to systematically span up a salience-similarity-space (Figure 6-7 (left)), and identify DT candidate prototypes.

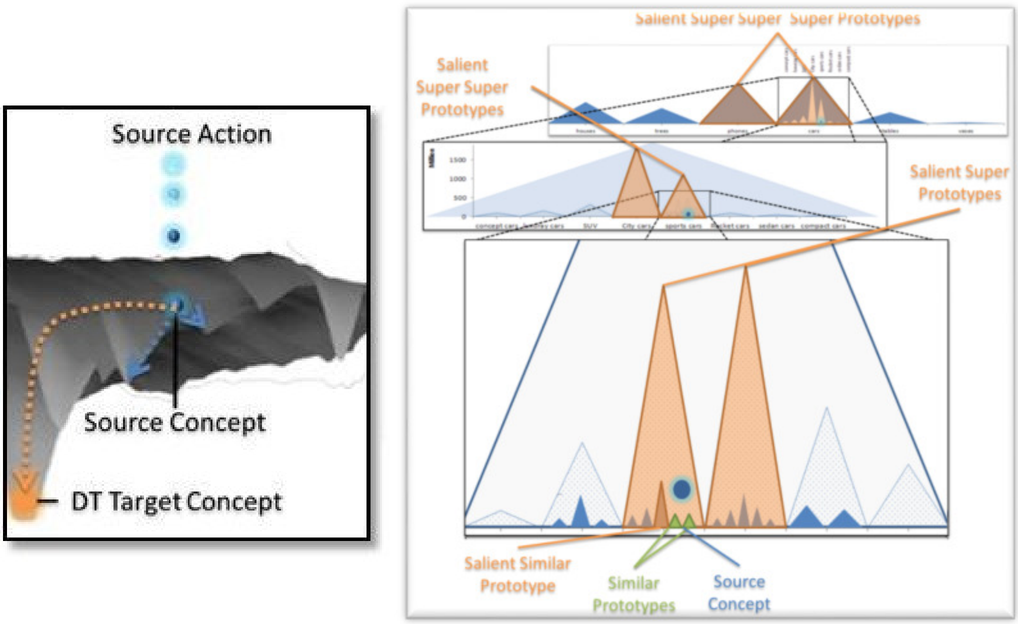


Figure 6-7. The DT design challenge (left). Salient Super Prototypes Approach for systematically identifying target contexts of high salience and sufficient similarity to the source context (right).

The processes initiate from source concepts and gradually expanded the salience-similarity-space towards prototype concepts of adjacent similarity. One procedure identifies relevant concepts through generalization, as detailed in *Salient Super Proto-*

35 4.3.2 *DT Design Challenge* (pp. 180)

36 2.1.4 *Concepts, Prototypes, and Categories* (pp. 24)

type Identification Approach (pp. 175). The other procedure builds on sub concepts of the source concept, elaborated in section *Sub-Action Modeling Approach* (pp. 183).

Both approaches capture concept structures directly from possible end-users. The single steps involved will be briefly described in the following section. Ideally, each step should be performed without knowledge about the previous step. Consequently, they may be performed by distinct user groups. In any case, results should be sorted by frequency or immediacy, as research on cognitive prototypes suggest.

SUPER SALIENT PROTOTYPE IDENTIFICATION

The cognitive background of the super salient prototype identification approach is visualized in Figure 6-7 (right). Similar prototypes are requested from participants to iteratively determine super prototypes. Every super iteration reveals more general concepts, going along with prototypes of increased salience and a broader context. Hence, the iterations progressively expand the conceptual context and increase dissimilarity. Examples of those generalized concepts, offer further salient cognitive prototypes. This way the procedure concentrically spans the super-salience-space, The iterative identification procedure of salient super prototypes is described in Figure 6-8.

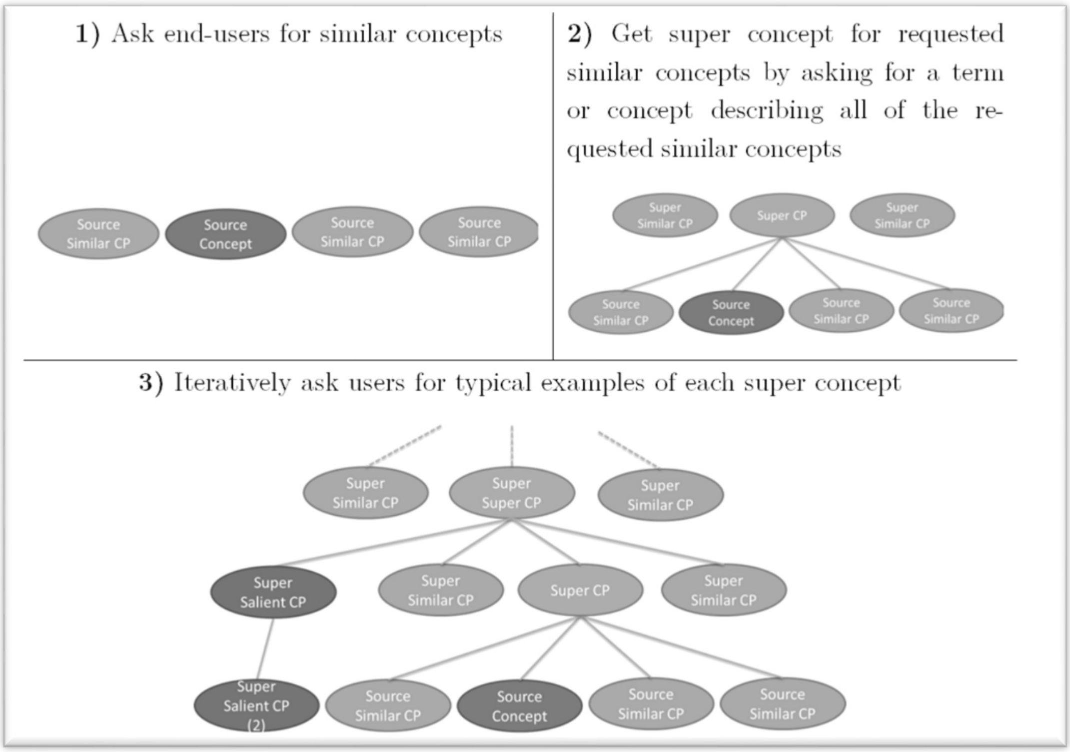


Figure 6-8. Iterative identification procedure of salient super prototypes.

It initially starts from a source concept, iteratively revealing super concepts. Additionally, each super concept provides access to most salient concepts of this class, offering super salient concepts similar to the source concept. Similarity decreases with every super iteration.

<u>SUB-ACTION MODELING APPROACH</u>

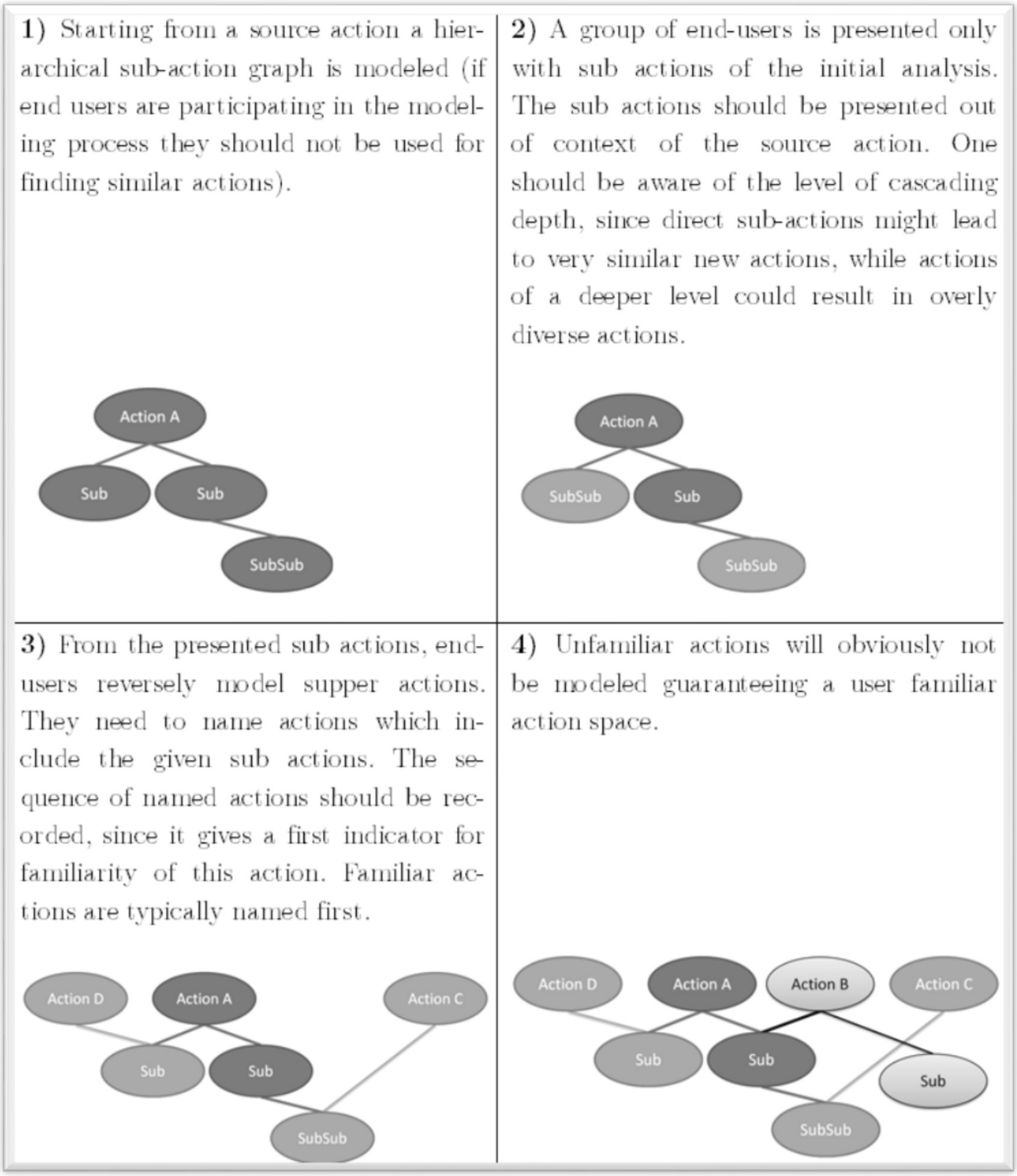


Figure 6-9. Iterative identification of similar salient prototypes through sub-action modeling.

While the super salient prototype identification approach was derived only from evidential cognitive findings, the sub-action modeling approach was developed from a technical perspective, originating in the Hierarchical Task Analyses. The procedure is summarized in the following illustration.

Both methods should provide numerous relevant cognitive prototypes of differing salience and similarity. In a next step, the results may be sorted by end-users, providing additional familiarity and efficiency ratings.

METHODS FOR FAMILIARITY ASSESSMENT

The primary requirement for any Digital Transformative (DT) should be maximum user familiarity, which correlates to cognitive prototypes[37]. Hence, an initial assessment of design ideas according to user familiarity increases the chances for determining proper DT concepts. Four possible methods for familiarity assessment are proposed[38]:

Heuristic expert estimation: Experts assess the level of familiarity for certain tasks or actions. This is considered to be the most effortless, but also most inaccurate method for familiarity assessment.
Probabilistic environmental observation: Cognitive prototype categories are being developed based on environmental or behavioral occurrences; hence, an adequate measurement of such occurrences may offer a fairly accurate probabilistic pattern, corresponding to cognitive patterns of high performance within users.
Learning curve analysis: A learning curve analysis is based on the idea that user efficiency of processes and procedures increases with practice and occurrence frequency. Performance analyses allow for determining the learning curve for certain actions and the performance state on that curve. This way they give indirect implication on the current level of familiarity.
User rated familiarity: User rated familiarity offers an accurate method for assessment. The method requires potential end-users to assess action contexts by familiarity. Tests showed that user rated action familiarity significantly correlates to performances.

Table 6-1. Methods for assessing familiarity.

[37] *User Rated Familiarity* (pp. 171)

[38] 4.2.1 Methods for Assessing Familiarity (pp. 162)

While expert estimations or learning curve analysis are common, well elaborated techniques in user interface design, the two methods of probabilistic environmental observation and user rated familiarity need further explanation and evaluation.

Probabilistic Environmental Observation

A color perception test, conducted in this work, highlights the relation between environmental input stimuli and the development of conceptual structures[39]. The test application also provides an example implementation for the measurement of environmental stimuli. Its interface and results are shown in Figure 6-10.

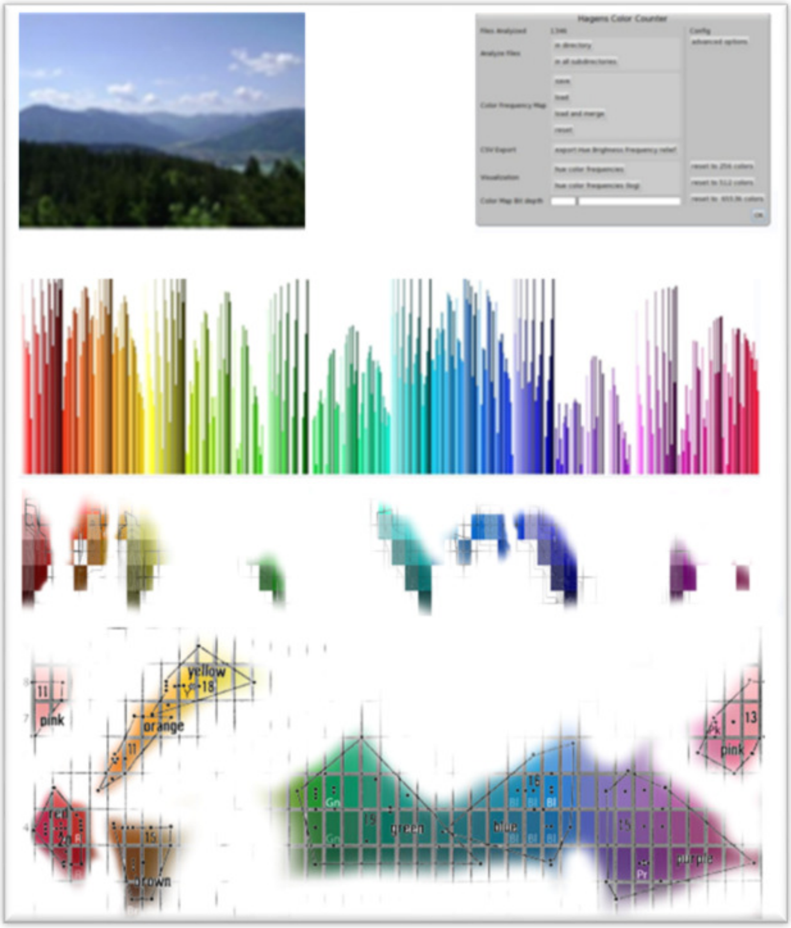


Figure 6-10. User interface of an application developed for analyzing the probabilistic environmental color stimuli of test-persons (top). Test results (middle) compared to the a distribution found through the Basic Color Terms study (Berlin & Kay, 1969).

[39] *Probabilistic Environmental Observation* (pp. 162)

In this test, the quantized color perception of multiple persons from the same area was analyzed, accumulated, and averaged, to get a probabilistic environmental analysis. The results show clear similarities to some outcomes of the World Color Survey and the Basic Color Terms study. Those studies captured cognitive concepts on colors and color terms of various cultures around the world, and provide an extensive ressource for assessing cognitive concepts (Berlin & Kay, 1969; P. Kay, Berlin, Maffi, & Merrifield, 2007; Richard Cook et al., 2012). Thus, the color test clearly indicates that the occurance of environmental stimuli probabilistically maps onto the formation of cognitive conecpts. This gives hint for the validaty of this method for deducing cognitive concept structures from environmental user stimuli.

User Rated Familiarity

Additionally, a validation of user rated familiarity showed that users are able to rate familiarity with sufficient accuracy for this application[40]. The test provides evidence on a correlation between user familiarity assessments and user performances, indicating the efficiency potential of familiarity, as visualized in Figure 6-11.

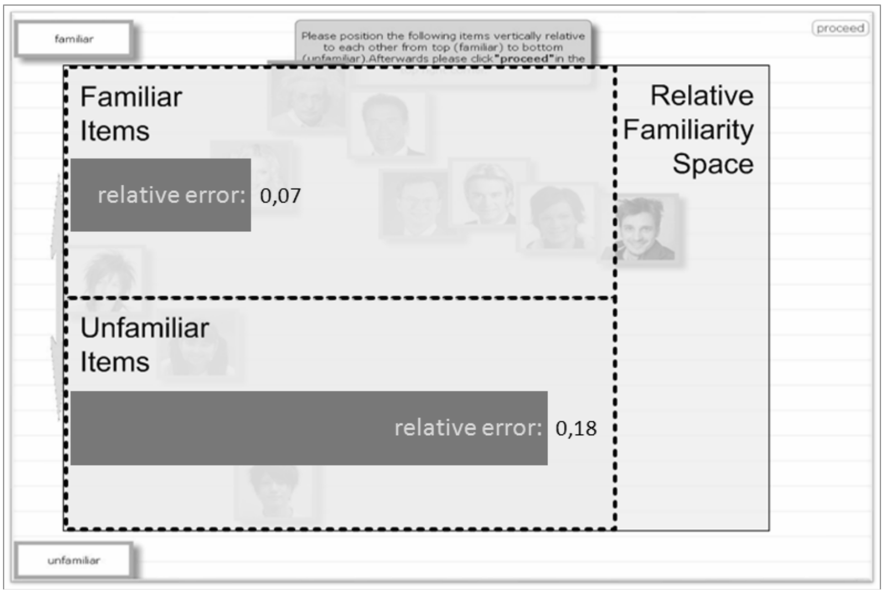


Figure 6-11. Extensive test on accuracy of user rated familiarity and its relation to performance.

Further analysis, of more than 29000 matches of the German professional football league showed a clear tendency for performance improvements if teams had fewer turnovers during a season[41]. Hereby it was assumed that fewer fluctuation of team

[40] *User Rated Familiarity* (pp. 171)

[41] 3.4.3 *Correlation between Familiarity and Performance in Team Sports* (pp. 126)

members correlates with higher team familiarity. The average performance improvement of the top teams, which had 4 times less turnovers, was determined with 22 percent.

RELATIONS BETWEEN FAMILIARITY, AUTOMATIZATION, AND COGNITIVE PROTOTYPES

The investigations on familiarity and efficiency also emphasize a correlation between practice in automatization processes and familiarity. Based on a stability index for the logarithmic growth of un-familiarity of approximately R^2=0,98, the relation of familiarity and performance can be described with a logarithmic progression, with a stability index of R^2>0,9. Hence, analogies to the formula for the law of practice, describing automatization processes[42], underline the assumption that familiarity corresponds to practice in automatization processes (illustrated in Figure 6-12 (left)).

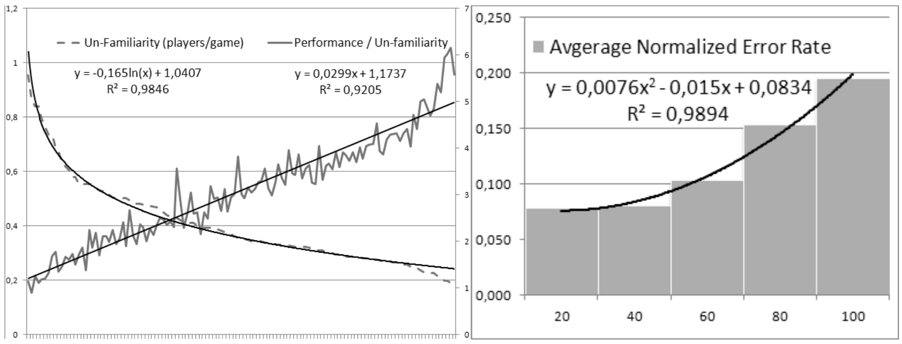


Figure 6-12. Indicators for familiarity corresponding to practice. Familiarity in team sports in relation to performance (left). Familiarity and performance in the face memory test (right).

Those findings are further supported by a user rated familiarity test, as shown in Figure 6-12 (right). Consequently, familiarity seems to be a valuable measure for finding prototypes, and the studies give further insight on how those cognitive prototypes are formed.

INTEGRATED SYSTEM DESIGN METHODOLOGY

The findings and systematic design elements of Digital Transformatives integrate well into typical interactive system design procedures. Accordingly, Figure 6-13 shows an adaption of the common design approach towards a cognitive prototypes oriented interactive system design procedure[43]. This adapted procedure suggests the consideration of DT heuristics during the conceptual idea development phase. Methods for

[42] *From controlled to automated processes to habituation* (pp. 38)

[43] 4.2, *Interactive System Design Methodology for Digital Transformatives* (pp. 160)

systematically identifying salient cognitive prototypes may be used to foster the finding of DT concepts. In a next step, the common requirements and guidelines based evaluation of the design ideas should be complemented by familiarity assessments. This leads to a prototype oriented initial system concept, which is further refined in an iterative development phase.

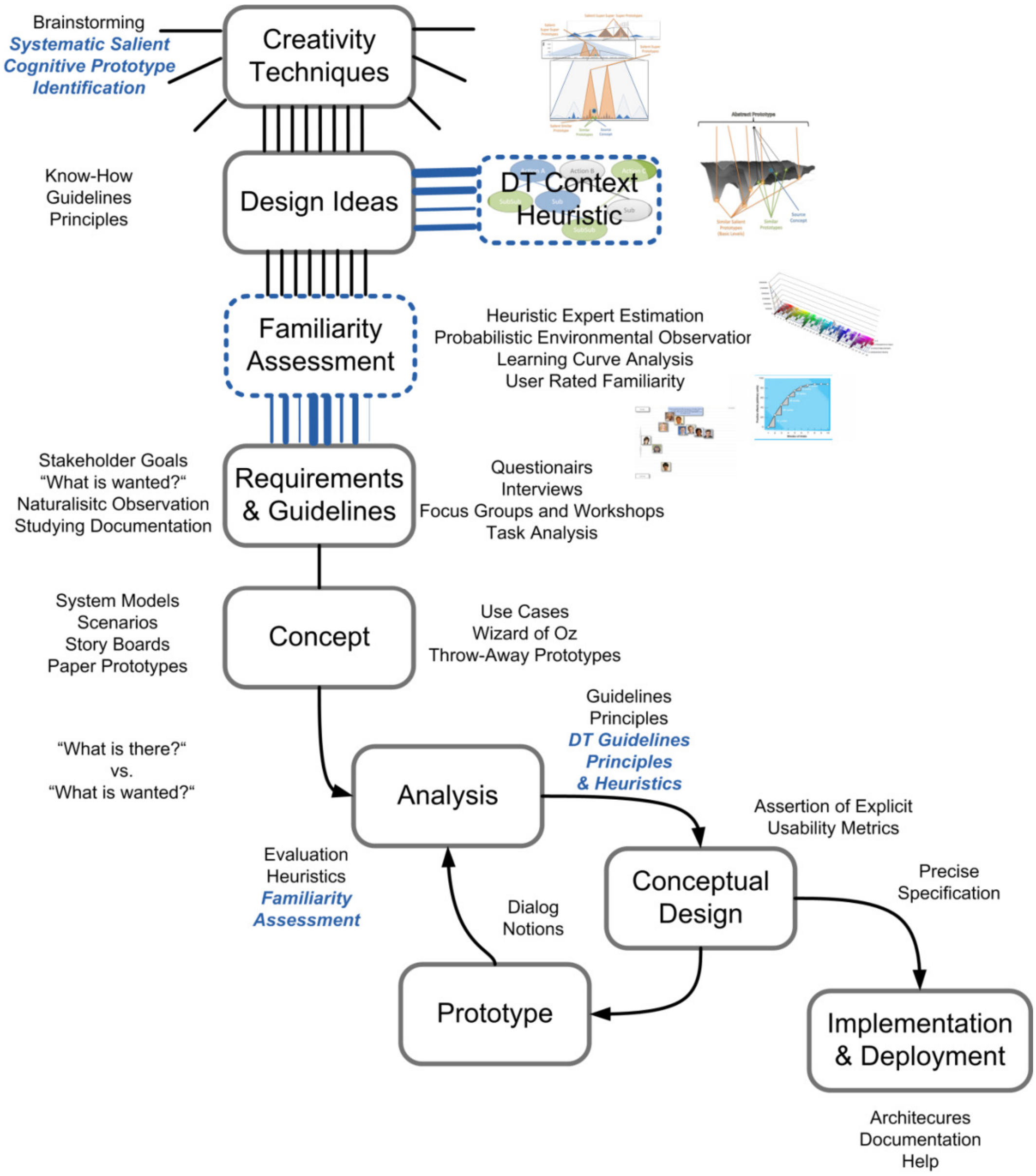


Figure 6-13. An adaption of the common interactive system design approach fostering systematic Digital Transformatives design. New aspects are highlighted in blue.

Within the iterative development phase, DT guidelines and principles should be considered at every conceptual refinement. Additionally, familiarity assessments should be implemented in the evaluation phases.

6.1.4 Use Case Prototypes

Besides the previously described test cases, multiple use case systems have been implemented, throughout this research, to study the theoretical concepts practically[44]. The use cases started off with the description of initial concept design ideas, highlighting the integration of landmarks into navigation systems[45]. Moreover, the influences of this work on the Science Center To Go have been detailed.

The second generation[46] describes first approaches which were developed to investigate the importance of familiarity, and the correspondence to performance and cognitive prototypes. The importance of a comparative assessment tool has been evaluated[47]. This Digital Transformative was also used to underline that users were able to accurately rate their own familiarity, and hereby implicitly assess areas of high performance. Moreover, based on the concept of Transitional Objects in learning, it has been detailed how Digital Transformatives may be used for inducing behavior change. In this context, an application, called digital beverage coaster, was developed to elaborate on the relation between Persuasive Technologies and Digital Transformatives[48].

In the third generation[49], methodological tools were introduced and developed for supporting cooperative web based Digital Transformative design. At first, a conventional iterative design cycle has been used for the implementation of a cooperative web module for Sub-Action Modeling. In contrast, the second development of a module for Salient Super Prototype Modeling emphasizes on DT specific design elements. Its user interface is shown in Figure 6-14 (left).

[44] Detailed in chapter 5 *Iterative Use Case Prototypes (pp. 197)*

[45] *Generation 1 – Concept Designs and Science Center To Go*, (pp. 198)

[46] *Generation 2 – First Approaches based on Familiarity* (pp. 214)

[47] *Assess the Assessable* (pp. 227)

[48] *Drinking Garden – a Digital Beverage Coaster*, (pp. 219)

[49] Compare Table 5-1, and (pp. 249)

Figure 6-14. The user interface of a game based DT for Salient Super Prototype Modeling (left). A screenshot of the Speed Reading Game for capturing general cognitive concept structures (right).

Hence, the design methodology described in *DT Web Module Salient Super Prototype Modeling* (pp. 241), provides a basic example, showcasing cognitive prototype oriented DT design. Finally, another tool is described in form of a speed reading game (Figure 6-14 right). The tool is supposed to provide a module for generally extracting cognitive structures of users in their leisure time.

6.1.5 Function Characteristics and Design Guidelines

The research on the concept and working principle of Digital Transformatives provides a solid basis for specific development support. Additionally, the iterative research methodology of this thesis was driven by hypotheses, which were transformed into DT functional features, and complemented by design guidelines. The functional characteristics and design guidelines developed throughout the work are summarized in the DT framework. The main attributes and interrelations among those attributes, of DTs are graphically summarized in Figure 6-15. A chronological summary, including all hypotheses, can be found in section *Appendix B - Hypotheses,* Features, and Guidelines (pp. 292).

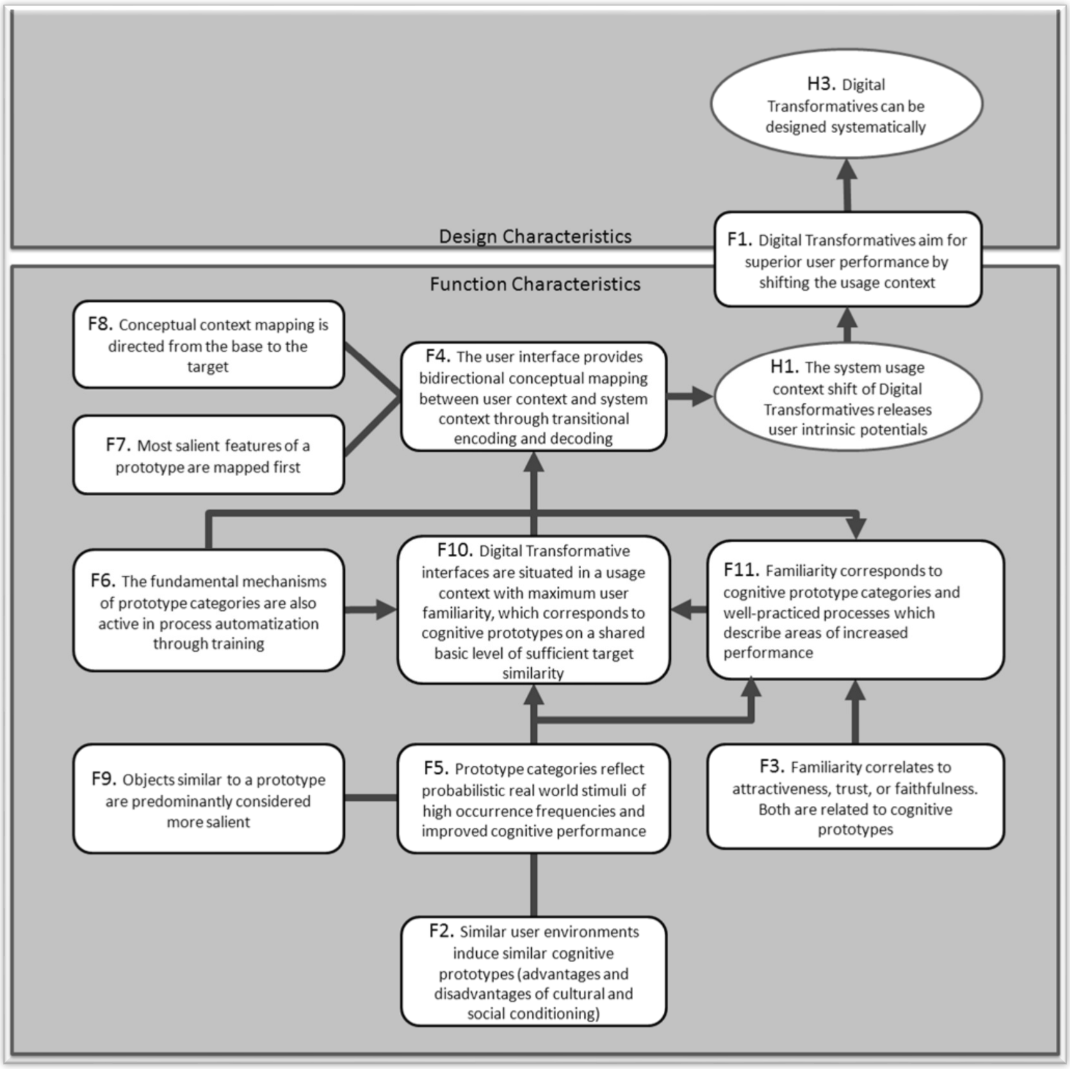


Figure 6-15. Functional characteristics of Digital Transformatives.

<u>FUNCTIONAL CHARACTERISTICS</u>

The fundamental feature of Digital Transformatives is expressed in Feature 1. DTs are able to aim for superior user performance (Feature 1) if a system usage context shift releases user intrinsic potentials (Feature 12). According to the DT concept, DT interfaces are situated in high performance user contexts. System functions need to be encoded into such contexts, while user inputs are decoded back into function contexts (Feature 4).

The concept depends on high performance usage contexts (Feature 10). Hence, Feature 4 is dependent on Feature 10. According to cognitive research, prototype categories mark such areas of high human potentials. They are mainly formed through automatization, learning processes, and frequent occurring environmental stimuli (Feature 5, Feature 6). Salience is a major characteristic of cognitive Prototypes (Feature 9). Moreover, human beings living in similar environments tend to develop similar cognitive prototypes (Feature 2).

Additionally, the practical feature extraction, conducted in chapter 2.2, suggests the importance of familiarity, in relation to cognitive prototypes (Feature 11). The basic concept validation (chapter 3.4.4) provides supplementary empirical evidence for Feature 5 and Feature 6, which, in combination with Feature 3, refine validity of Feature 11.

Feature 11 offers further validation for Feature 10, which is also supported by cognitive evidences expressed in Feature 6, Feature 5, Feature 9, Feature 2. Thus, Feature 10 is based on a comprehensive empirical ground, also validating Feature 4, which provides the basis for Feature 12.

DESIGN GUIDELINES

After the basic concept and the principle of operation of Digital Transformatives have been elaborated, further investigations are driven by the question, whether there is a systematic design methodology for creating Digital Transformatives, expressed in Hypothesis 3.

Design concepts for interactive systems usually evolve from ideas gathered through brainstorming sessions. Such conceptual ideas are assessed based on previously determined requirements, experience based heuristics, and guidelines. Consequently, the basic principles of Digital Transformatives need to be implemented on the brainstorming and requirements level, which is influencing the concept design (Guideline 3).

In a common interactive system development procedure, any concept idea and implementation should be assessed on overall system requirements. Requirements ought to be implemented through familiarity assessment, which, hence, has an outstanding role in the design of DTs (Guideline 4. Assess concepts and implementations by user familiarity.).

The investigations conducted in this work suggest four major methods for assessing familiarity (Guideline 5, Guideline 6, Guideline 7, and Guideline 8)

In an iterative development cycle, Digital Transformatives are first implemented on a conceptual level. The first concept ideas evolve from empirical knowledge and may be complemented through creativity techniques, such as brainstorming. While Know-how may be acquired through experience, creativity is an unreliable factor. The investigations on the working principle of DTs provide an informative basis for creating systematic heuristics, which reduce the dependency on designer creativity and their unpredictable nature.

The working principle is expressed in Guideline 9. Find super salient cognitive concepts in sufficient similarity proximity to original function related concepts.

Since DTs address cognitive concepts within possible end-users, it is ideal to actively involve the users in the concept design process (Guideline 10. Capture information directly from possible end-users. System design concepts should be based on a shared cognitive model of possible end-users.).

Based on the previously given prerequisites, two procedures are proposed to identify DT concepts. The first builds on cognitive studies on categorization and findings on efficient communication, as elaborated in chapter 2.1 (Guideline 11. Traverse superordinate concepts to find cognitive prototypes of close similarity.).

The second procedure builds on a combination of cognitive insights on similarity measures and prototypability, and on measures and results determined from requirements analyses (Guideline 12. Ask for super salient neighbors which share salient sub actions (features) with the functional action (source concept).)

The final section on the practical analysis of use case prototypes revealed important empirical heuristics for inducing behavior change expressed in Guideline 13. Transitional Objects offer valuable anchors to induce behavior change with Digital Transformatives..

An Overview of all elaborated guidelines and features is given in Figure 6-16.

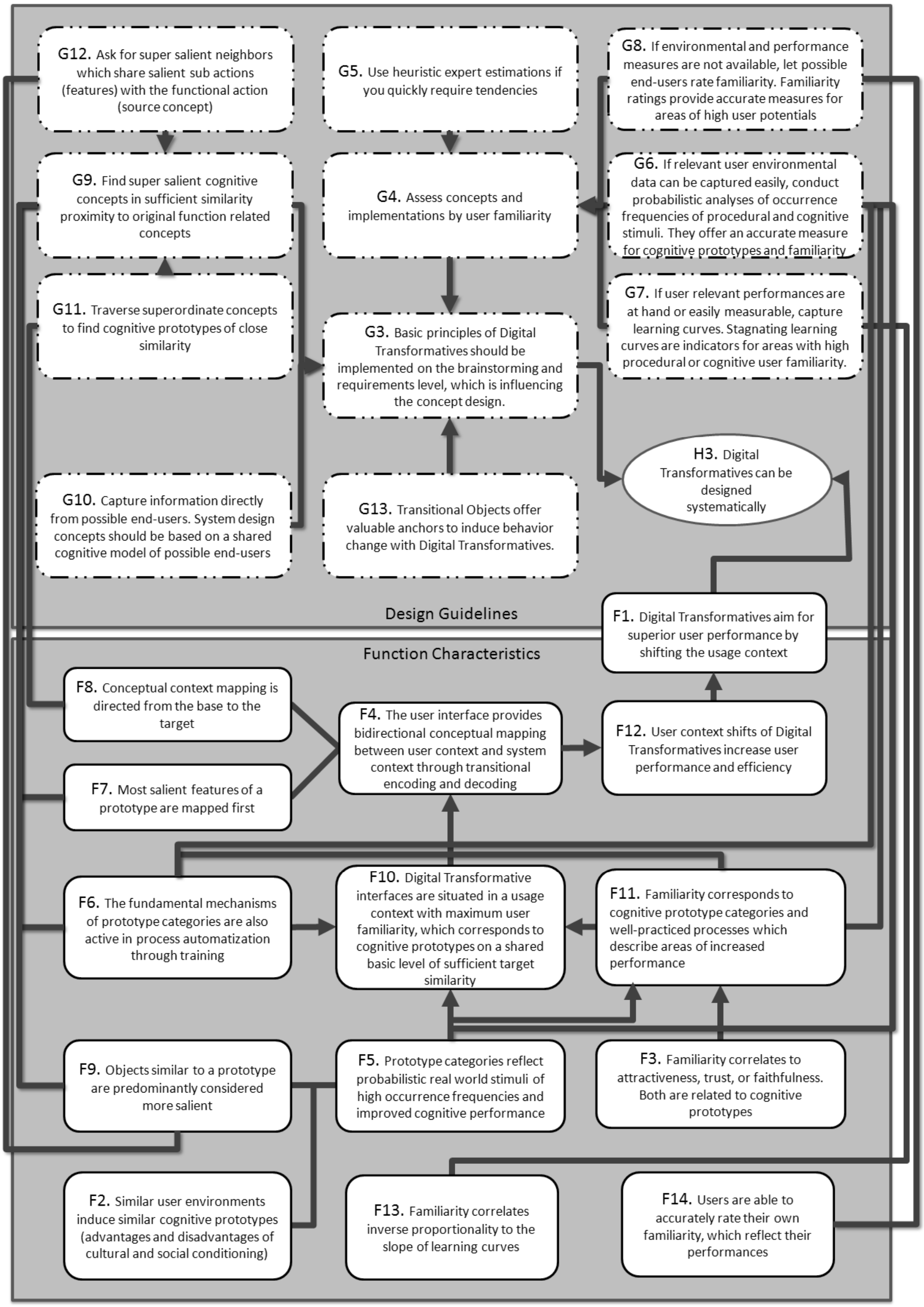


Figure 6-16. Overview of identified DT function characteristics and their relation to DT design guidelines.

6.2 Future Work – Cognitive Prototypes beyond Interactive System Design

The previous text highlights the importance of Cognitive Prototypes (CP) for the design of interactive systems. However, due to the ubiquitous significance of CPs in everyday life, insights, detailed in this work, can be further transferred to other domains. CPs reflect appearances everywhere around us, in countless variations. This omnipresent pulse of occurrences and frequencies has a fundamental reason, survival of the fittest.

Evolution made us become a highly efficient organism within our environment. Food is our fuel, and analogously, our energy is limited. Our system is optimized for achieving the best performance with the resources we have. Hereby, cognitive mechanisms, such as conceptual automatization and prototypization, provide highly efficient solutions for adapting our behavior to the environment we live in. Therefore, our cognitive structures are adapted, bundling our minds power to stimuli or actions, which are occurring on high frequencies, at the expense of less utilized cognitive concepts. Examples of such adaptation processes are habits, learning, or practice. Over time, the processing of frequent stimuli and actions becomes more efficient and pleasant, while processing unfamiliar stimuli strain us. Those effects can be used for communication (e.g. promotion of innovative ideas), may be experienced in recognition (illusions), or the sense of beauty (attractiveness, music, fashion, creativity).

This section will briefly elicit the importance of CPs in the above fields of everyday life, highlighting the chances new cooperative tools may bring to identify CPs from analyzing shared data on the internet.

6.2.1 CPs at Conventions, Standards, Norms, and the Force of Habit

The Digital Transformatives (DT) working principle is based on a concept of efficient communication. Essential elements of highly efficient communication are shared cognitive prototypes (CPs). Good examples of such shared prototypes are conventions.

We all probably have been in a situation where we needed to talk to a stranger who spoke an unknown language. In such cases gestures and mimics turned out to overcome language barriers. Gestures such as waving hands to greet, or thumbs up to express approval are more universal than any language. Those gestures are commonly understood, although they have not been explicitly defined. They are part of social behavior conventions which have been shaped over time and among various cultures. Doubtless, as the above example shows, they often improve efficiency in communication. We are constantly confronted with numerous other conventions of social interaction; even right now, when you read this text top-down from left to right.

More generally, conventions, systems, or products, which achieved a dominant position in a certain culture, are often also referred to as *de facto standards*. Examples for those are Compact Cassettes, HTML, MP3, QWERTY. Analogies between conventions and cognitive prototypes are apparent. They develop at areas which rapidly gain importance through rising frequency of use. One could also say, they develop on demand. For example, the advent of type writers and computers induced a need for keyboards. Multiple solutions for keyboard shapes and layouts are offered, however, with rising frequency of use, certain layouts became dominant. Although, layouts like QWERTY may ergonomically not be the most efficient solution for single individuals, the general agreement on a certain layout definitely increases shared efficiency.

Additionally, the agreement on de facto standards, such as QWERTY, reduces complexity in production. On these grounds technical standards usually also improve efficiency. "A technical standard is an established norm or requirement in regard to technical systems." (Wikipedia.org, 2013e). We implement and employ norms mainly to gain efficiency in engineering, processes, or practices (DIN, 2013; ISO, 2013). A standard based implementation of a system guarantees interaction with other standard based systems, and hereby improves efficiency in production. For example, doors can be created by machines following certain norms, instead of measuring each single door frame to create a tailor-made match.

Obviously, the implementation of standards in our daily environment also re-induces Cognitive Prototypes. Thus, the implementation of a technical standard does not only improve efficiency in production, it also improves cognitive efficiency within every human. In modern houses we do not expect variances in door heights and widths, in the same way we do not expect height variances within a flight of a stairway[50] (Cote & Harrington, 2006; DIN EN ISO 14122-1, 2002; DIN EN ISO 14122-3, 2002). Ancient buildings easily let us experience how much more cautious we have to be if such standards are not followed (compare Figure 6-17 left). Moving around becomes much more stressful and takes more time, if we need to reconsider every step in a stair or every door's height we are passing.

[50] "Variance on riser height and tread depth between steps on the same flight should be very low." (Cote & Harrington, 2006, p. 167).

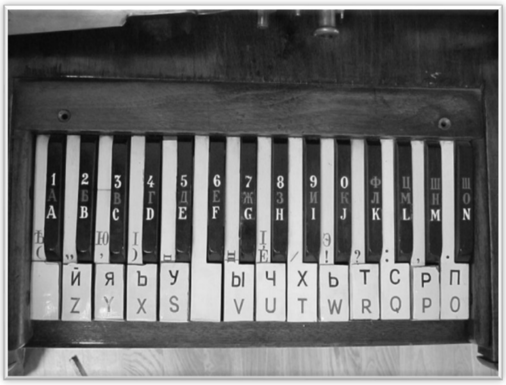


Figure 6-17. Standards, conventions, and the force of habits may improve efficiency. If those standards are broken our usage efficiency decreases, as showcased through low variances in a stair flight (left), or keyboard layouts different to QWERTY (right) (Wikipedia.org, 2012a, 2012b).

Endless further examples could be given, where standards induce cognitive prototypes and vice versa. The efficiency gains become most obvious when habits are broken, such as writing on computer keyboards with foreign language layouts (QERTY vs. QERTZ), alternating between playing tennis and badminton, or driving cars with shift stick or automatic transmission, The high degree of automatization may also become a big risk, for example, when we are driving a car on the left or right side of the street, or using a bicycle with hand-break or back pedal break.

6.2.2 CPs in Perception and Recognition

Cognitive prototypes also influence much more subliminal and fundamental recognition and categorization processes. Knowledge about CPs offers simple but powerful indicators on how recognition works, and how it can be manipulated.

Frequently perceived occurrences shape our recognition. For example, we say that we are social beings. We prefer to live together with friends and socialize with others in cities, instead of encapsulating ourselves, living evenly distributed on mother earth. The most expressive part of social communication is exchanged via faces. We constantly focus on faces around us. Our brain adapts to such high frequencies and forms comparably strong cognitive face prototypes, to optimize general efficiency[51]. As stated in the DT Design Challenge, or through the color perception experiment, such prototypes heavily influence our recognition. Our brain tends to assign stimuli

[51] Faces are highly efficient cognitive areas, which may explain the importance of profile pictures in social networks.

to frequent occurrences. This is why we seem to recognize faces in various random patterns such as clouds, fire, wood, or geological formations.

A prominent example of the influences of cognitive prototypes on perception is given by the Martian face. The Martian face has been spotted by the space ship Viking 2 in 1976 on the surface of Mars (see Figure 6-18).

Figure 6-18. The Martian face, which turned out to be nothing alike (NASA Science, 2001).

The attraction of this appearance was high enough to prioritize its observation during the Mars Global Surveyor mission, conducted eighteen years later. Obviously it turned out to be a common geological formation (NASA Science, 2001). Similar mechanisms might be active in gestalt law effects, as exemplified on grouping laws in Figure 6-19.

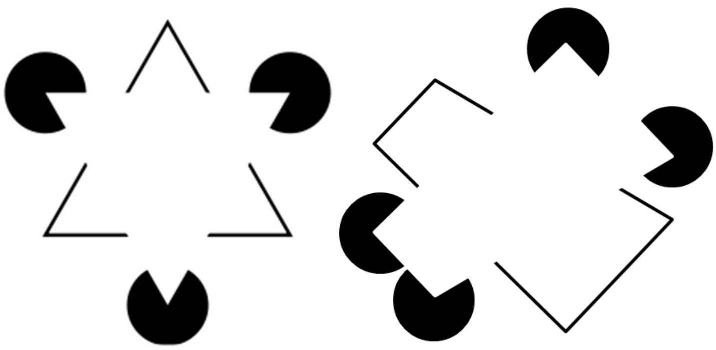


Figure 6-19. Ambiguous images based on Gestalt grouping rules (middle-right) (Wikipedia.org, 2013a).

The recognition of the implicit triangle (left) and the implicit distorted rectangle (right) illustrates the gestalt grouping laws. Triangular shapes are frequent occurrences in our environment, fostering the recognition of such shapes, although they may not even be visible. The shape of a distorted rectangle, however, is less frequent, also reducing the gestalt grouping law effect. Since it is less prototypical it should also be slightly more stressful and less pleasant to look at.

Our favor for occurrences corresponding to highly efficient cognitive prototypes, also explains why we like symmetry. Cognitive prototypes are some form of averages, which often correspond to mean values in normative distributions. Analogously, most cognitive prototypes should be symmetric, as further elaborated in section 6.2.4 (*CPs on Pleasantness and Affection*, pp. 279). This aspect may be exemplified with a person trying to draw a free hand circle. The first try might be quite imperfect, but if we repetitively draw the same circle the average of all layers forms a more an more perfect exemplar.

The above explanations show why it is important to analyze cognitive prototypes in our perceived environment. This way, we are able to understand and reduce misconceptions. On the other hand, this allows us to intentionally implement illusions, or improve subjective-wellbeing by increasing the amount of cognitive prototypes.

6.2.3 CPs Help Promoting an Image

Knowledge on cognitive prototypes can also be very useful for methodical communication of information. Personal or economic success is often directly related to a certain public perception. Cognitive Prototypes can be utilized to systematically communicate a certain image. In advertising, for example, celebrities are being used as cognitive prototypes for subtly inducing a designated association with a product or company.

However, general salient concepts hold further potential for influencing public perception. Let us consider the goal of establishing a remarkable name or logo. For example, the firm name and logo of Apple Inc. makes great use of existing prototypes. The shape of an apple is a general prototype, and the same accounts for the term apple, at least in English speaking cultures. Hence, the logo and name of the company is easily being remembered, because it makes use of a strong existing prototype. In the western world, the apple can be seen as a celebrity of fruits. It is omnipresent and holds many positive associations; healthiness and vitality are two salient nearby prototypes, for instance. Additionally, to guarantee uniqueness, it is important that the prototype of an apple is not already prominently being used by another company in this context.

While it is often easy to ask gut feeling for commonly developed prototypes, it would be helpful to have more objective measures for typicality. Based on the idea that cognitive prototypes form around shared general probabilistic maxima, we need a source which aggregates spontaneous, interest driven, inputs of numerous individuals. Hence, a good indicator on general prototypicality of terms may be found in the amount of queries input in search engines. Hereby, it can be assumed that terms related to more salient concepts, are also entered more frequently. Figure 6-20 displays search query amounts for fruit terms entered in Googles search engine.

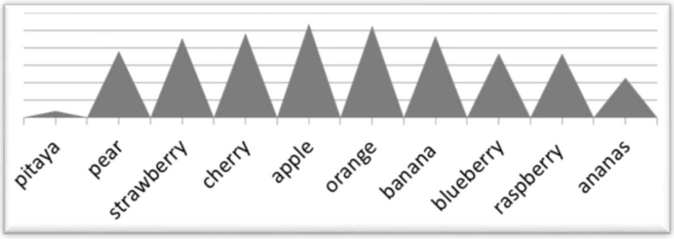


Figure 6-20. Search query amounts (logarithmic scaling) for fruit terms as an indicator for prototypability (suggestqueries.google.com).

If we are taking such amounts as an indicator for prototypability, apple indeed is dominant for fruits; the prototypes of orange, banana, cherry, or strawberry, are slightly less salient. Accordingly, a company name like Pitaya would be comparably unremarkable.

Figure 6-21. Examples for company logos (Apple Inc., 2013; Castrol, 2013; Mercedes-Benz, 2013; Toyota, 2013; Unilever, 2013)

What is true for names, also applies for shapes. Do you recognize all of the logos shown above? In this case, search query terms do not help at first sight. A simple test, whether a logo addresses prototypes, can be conducted by describing it verbally (E. H. Rosch, 1973b). We tend to increase efficiency in our communication by using prototypes. Consequently, the complexity of our description is a strong indicator for prototypability[52]. Additionally, another person could make a drawing based on this

[52]compare *Cognitive Efficiency Catalysts in Communication* (pp. 47)

description. The closer the drawing gets to the source concept, and the shorter the description, the more remarkable the logo.

It is easy to describe the Apple logo, and the Mercedes-Benz logo may be described with a star inside a circle. Circle and star are two very common prototypes. However, the star may not be salient in the form used here. Many individuals might have star prototypes with more than three spikes in mind. On the other hand, in the western world the Mercedes star has become a common prototype itself. This shows that, analogously to de facto standards or conventions, it is possible for big companies to shape own cognitive prototypes within their consumers. Nevertheless, this does not mean that logos of big companies in general develop into cognitive prototypes. In order to shape new prototypes, to a certain degree they should be similar to other existing prototypes.

For example, how would you describe the logo of Castrol Ltd.? The logo of Castrol addresses the cognitive prototype of a circle, but includes a hardly describable shape. The Unilever logo clearly makes use of a U-shape, despite the complex filling, it is remarkable. Contrarily, although the Toyota logo seems less complex, it may also be less recognizable, as long as the two inner ellipses cannot be referred to a prototype. This logo shows quite well the importance of sufficient similarity to cognitive prototypes.

In summary, the creation of a remarkable company name or logo, as described above, is fostered by the occupation of existing prototypes, which is taken from a different context in order to be established in a new context. For example, a logo with a red cross on white ground would hardly be remarkable for a company operating in a medical context. However, the same prototype could be a remarkable feature of a company logo in a technical context, such as a car repair company, Therefore, it is advantageous to occupy strong prototypes in untypical contexts.

In this section it has been briefly elicit, why and how cognitive prototypes can be used to promote an image, and how search query amounts may give an indicator on remarkability. CPs can further help determining a remarkable name and company logo, and they are also related to pleasantness and affection as described in the following section.

6.2.4 CPs on Pleasantness and Affection

If we look at the colors used in the company logos shown in the previous section (Figure 6-21), it becomes apparent that they address major salient color prototypes, compared to results of the World Color Survey (WCS). The WCS, as well as the test on environmental color occurrences[53], show that variations of gray, black, and white are most dominant salient colors in a western world environment, followed by natural frequencies of red, green, and blue. Those colors are most efficiently processed, and should commonly be considered pleasant. Interestingly, most salient colors are also often associated with premium, classical, perfectionist products; in this context one might only think of dresses or cars.

Similar clusters can also be found if we look at fashion. While fashion worn by young people is highly diverse and colorful, we tend to assimilate to each other with rising age. During lifetime a fashion prototype seems to be formed, averaging various trends. These averages correspond to the definition of cognitive prototypes.

A correlation between our sense for affection and averages of occurrences is indicated in various fields, besides fashion. The *Averager*, on faceresearch.org, offers a tool to interactively experience this correlation between attractiveness of faces and averages (DeBruine & Jones, 2013; Langlois & Roggman, 1990; Langlois et al., 2000; Rhodes, 2006).

Figure 6-22. The Averager computes the average of faces from a selection pool (middle). The left face shows an average of the first two female faces in the top row, the right face is the average of the 25 selected faces (DeBruine & Jones, 2013).

The Averager is based on an algorithm, which merges faces to an average. Several random faces are available for selection. The left face in the figure shows the average of two female faces, while the face on the right forms the average of 25 female faces. This clearly illustrates how averages, respectively cognitive prototypes, affect our perception. The greater the average the more smooth and symmetrical the face appears to be. These characteristics, and others, defined through averages may be related to pleasantness, and can also be found in many Gestalt laws.

[53] conducted in the section *Probabilistic Environmental Observation* (pp. 160)

Other studies show, that trustworthiness also corresponds to cognitive prototypes. Those results also correlate with the findings on the correlation between familiarity and prototypicality. We tend to have higher trust in people we have seen more often than others. Imagine, the actor Arnold Schwarzenegger and another stranger are ringing your door bell, trying to sell a vacuum cleaner. It is likely that Schwarzenegger seems to be more trustworthy to you, although you do not actually know this person better than the other. Simply higher occurrence frequencies are sufficient to let persons appear more trustful.

In general, prototype formation is highly dependent on stimulus exposure. The more often we process certain stimuli the more efficient we get, due to prototype formation. The more efficiently we process input patterns, the more pleasure it is to process them. The Mere Exposure Effect may be a strong evidential indicator for this correlation. Independent of the type of stimulus, whether it is a geometrical form, a melody, or a human face, the more often we are presented with certain stimuli, the more we tend to be affected by them (Bornstein & D'Agostino, 1992; Zajonc, 1968). This phenomenon has since been observed in more than two hundred experiments (Bornstein, 1989).

As illustrated above, similarity to probabilistic cognitive prototypes may also serve as a measure for determining and evaluating affection. It also serves as an indicator for familiarity and trustworthiness.

6.2.5 CPs in Communication and Innovation

Extensive exposure of a certain stimuli might also lead to the development of shared cognitive prototypes which may further be used categorical in communication. Like conventional prototype categories, those dynamic categories are essential for efficient communication[54]. For example, one could say: "There is a person who dances like Michael Jackson". The example wonderfully shows the advantages and the limits of utilizing such a category. On the one hand, everyone who saw Michael Jackson's typical moves, understands this complex information with minimal effort. On the other hand, in a decade the typical dancing moves of Michael Jackson might not be commonly known anymore. As illustrated by the example, categories can deliver highly efficient tools for communication; however, they may be highly variable, changing with expertise and context. Therefore, one should be aware to choose the right categories for transporting information. The larger the group of addressees the more general shared prototypes should be used. The logos of global player companies, dis-

[54] elaborated in *Cognitive Efficiency Catalysts in Communication* (pp. 46),

cussed previously, for instance build on very general salient prototypes. More specific, and less salient concepts, could be better if we need to attract a smaller group of addressees.

Knowledge about general prototypes, or expert prototypes within a certain community, provides us with important information on possible acceptance of a product. It gives hint, whether the market is ready for an innovation. The basis for the acceptance of any innovation is its understanding. Numerous inventions came before their time, where possible users did not see a use for it. One of the most prominent examples is the invention of the mechanical movable type printing. Johannes Gutenberg developed and operated the first press of that kind by 1450. However, he could not turn the invention into a profitable business, and got bankrupt in 1456. In 1998 journalists voted Gutenberg to be "Man of the Millenium", because of his influential invention (Wikipedia.org, 2013c). The new economy provides many more similar examples.

If the acceptance of an invention is based on how it is understood, then prototypes help promoting it. Vice versa, prototypes also provide a strong indicator whether it is time for an invention or not. Therefore, one should determine significant features and user interests related to the innovation, and test germane prototypes for their salience. This prototypical finger print for innovation can further be related to prototype landscapes of other successful and unsuccessful products for comparison. The whole procedure is comparable to learning with transitional objects[55]. Hereby, CPs give easy access to similar new information. If sufficient similar cognitive prototypes are established in general public, than it is likely that the value of the innovation will also be understood.

Hence, a finger print for innovations, or new ideas, can be created and matched against existing cognitive prototypes of possible stakeholders. The amount of similarity of this matching gives an indicator on the acceptance of an idea.

6.2.6 CPs in social structures

As exemplified by assimilating fashion with rising age, or the establishment of de facto standards and conventions, cognitive structures are not only reflected by single individuals, but individual cognitive mechanisms are also reflected and implemented in communities and cultures. Hence, we can find cognitive prototypes also in social structures. Typically every community has its social hubs of very active and well

[55] detailed in *Cognitive and Practical Background Constructivistic Learning with Transitional* Objects and Digital Manipulatives, pp. 92

received members, while others are completely inactive. This phenomenon can nowadays be well observed in digital social networks such as Twitter or Facebook. They foster prototype development within communities, by offering tools for easily aggregating and distributing information among group members.

There are many more analogies between cognitive structures of individuals and social structures in communities. For example, in the same way some persons turn into specialists for communication, others follow different passions. A community develops various professions dependent on demanded services. Like a cognitive prototype, single individuals specialize in a certain field, delivering highly efficient work. Our social environment shows structures similar to our internal working principles. Conventions, for example, show how consensus on frequently addressed social aspects is generated, and how this improves general efficiency.

As the above examples illustrate our internal cognitive structures, are reflected into our community, and vice versa. This way, the intrinsic mechanisms of cognitive prototypes also increase efficiency in social communities. Current trends turned the internet into a social place, providing everyone with the chance to share information. Consequently, since the amount of shared information through improved communication services, such as web 2.0 technologies, further increased, the internet gives a more fine granular soil for studying shared cognitive prototypes, and fostering such structures.

6.3 New Chances to Analyze and Utilize Cognitive Prototypes

As highlighted in the previous chapter, CPs are dominating our recognition, which is reflecting our environment. Hence, it is essential for many areas to get a better understanding of CPs, whether this is learning, product branding, innovation, or user interface design. While the analysis of CPs was laborious in times when they were discovered, the internet, and todays' tendency towards ubiquitous computing, provide new possibilities to capture cognitive structures.

Social Networks, search engine analytics, and crowd sourcing mechanisms offer new insights on shared cognitive prototypes. Special tools can be developed to investigate prototypes of user groups and influence all kind of fields that involve human communication.

6.4 Critical Revision

This work heavily suggests cognitive prototype based development of interactive systems. Cognitive prototypes describe declarative and procedural areas of improved performance and can be obtained through familiarity measures.

6.4.1 DTs might Hinder Personal Development

Critically seen, Digital Transformatives might hinder innovation and personal development, since they build on existing skills and abilities. Users are not forced to acquire and learn completely new skills. Therefore, one should always carefully assess whether a given task demands a completely new interface. Alternatively, one could make use of DTs as Transitional Objects[56]. In this case, DTs are designed to serve as tools to acquire new knowledge or behavior.

6.4.2 Failures through the Force of Habit

The design of Digital Transformatives should be conducted with care, especially for systems addressing dangerous tasks. Possibly dangerous problems may occur, if cognitive prototypes are not met properly[57]. This is best illustrated with habits. Habits are procedural super salient prototypes, trained until automatization. Such highly efficient actions demand a minimum amount of attention; they almost happen without conscious control. Consequently, it may become dangerous when habits do not match the new environment anymore. For example, if you cross a street in your home town, you automatically look to a certain side, right before you start walking. In a country with inverted driving directions you might have problems to consciously change this habituated subconscious behavior. Since Digital Transformatives heavily utilize automatized procedures, and declarative cognitive prototypes, a changing environment might have serious consequences. Therefore, it is important to be aware of all salient prototypes used by a Digital Transformative, and validate them against all possible environments the DT might be applied.

[56] Section *Cognitive and Practical Background Constructivistic Learning with Transitional Objects and Digital Manipulatives* (pp. 92)
[57] exemplified in *CPs at Conventions, Standards, Norms, and the Force* of Habit, (pp. 285).

6.5 Wrap-up and Future work

According to the long term goal postulated in the end of section *Basic Idea* (pp. 3), this work seeks to establish a new perspective, helping with the creation of highly efficient applications by fostering intrinsic user performance.

6.5.1 Research Questions

The short term goal of this thesis aimed at validating the basic principles for this achievement, and providing a conceptual model, which is complemented by a methodological framework. This outline was expressed in three major research questions in the beginning of this text:

RQ1. IS IT POSSIBLE TO LEARN FROM MNEMONICS IN ORDER TO IMPROVE HUMAN MACHINE INTERFACES?
RQ1 is mainly addressed in section 3.4 and 4.2.1, which show that the principle active in mnemonic devices also can be used to improve human machine interfaces.

RQ2. ARE THE KEY WORKING PRINCIPLES OF MNEMONIC DEVICES APPLICABLE THROUGH HUMAN MACHINE INTERFACES (HMI) OF INTERACTIVE SYSTEMS?
In chapter 3, the key principle has been further elaborated, and mapped onto a concept of Digital Transformatives. Function characteristics were identified throughout the work. A summary can be found in section 6.1.5.

RQ3. ARE THERE METHODOLOGIES FOR SYSTEMATICALLY APPLYING THE WORKING PRINCIPLES, IN ORDER TO FOSTER THE CREATION OF SUCH ENHANCED SYSTEMS?
A design methodology has been developed for the systematic creation of DTs in chapter 4. Design guidelines have been formulated mainly in the second part of this text, concluded in section 6.1.5.

6.5.2 Outcomes Beyond Research Questions

The work on Digital Transformatives offers fundamental cognitive insights on user interfaces. Especially the importance of cognitive prototypes for user interface design, and human communication in general, is highlighted.

Cognitive prototypes offer a model, which approximates cognitive structures. Comparable to Newton's mechanical laws, they might offer a sufficient approximation of much more complex processes and structures, which cannot be represented and understood, yet. However, due to their probabilistic nature, it is often hard to understand them. This might be one of the reasons why they are only little recognized in many fields involving human communication, such as interactive system design, marketing, or social sciences. Therefore, this text delivers many examples and showcases, trying to concretize access to the abstract idea of prototypes. Understanding cognitive prototypes helps understanding otherwise fuzzy terms.

For example, many users often criticize missing intuitiveness of a user interface or they claim that usage feels unnaturally. But when and why do we call an interface “intuitive”? Although intuition is a fuzzy term it is often the major aspect for purchasing a device. On the other hand, the fuzzy nature of this term makes it hard for system designers to address this point of criticism.

“Intuition is the ability to acquire representation or knowledge about things, apparently without reasoning or usage of reason in general. Cases of intuition are of a great diversity, however processes by which they happen typically remain mostly unknown to the thinker, as opposed to our view of rational thinking. […]” (Wikipedia.org, 2013b)

Hereby, the characteristics of cognitive prototypes suggest correlations with intuition. They describe areas of high efficiency, which are salient and spontaneously activated. Hence intuition may likely be based on cognitive prototypes.

Moreover, this work offers insights, concept, and tools for further socio-technical analyses of cognitive prototypes. The resulting models may be valuable in various fields related to human communication, such as economics, innovation, or information dissemination.

6.5.3 Outlook

This work was accompanied by the implementation of numerous practical test prototypes, which helped refining a conceptual and theoretical framework. In order to transfer this theoretical basis back to practical implementations, a web based framework is being developed trying to support a possible community of practitioners. The implementation of such an online framework started during this work with the implementation of the Action-Modeling tools, ColleaGuess, and the Reading Speed game, and it will further include some sample prototypes for Digital Transformatives, which can be tested, discussed and further developed by interested persons. It will also provide a set of tools, such as a comparative evaluation tool, allowing others to setup familiarity tests for their own Digital Transformative design.

In summary, this work gives a new perspective and understanding of digital systems, helping readers to reconsider existing systems in a new light. Beyond digital design it may also help with further understanding economic or social processes. In future work, it might be interesting to create and investigate tools for analyzing cognitive prototypes, to make use of them in other domains, apart from digital system design.

APPENDIX A - INVERSE DOT PATTERN ILLUSTRATION 290

APPENDIX B - HYPOTHESES, FEATURES, AND GUIDELINES 294

Appendix A - Inverse Dot Pattern Illustration

Figure A-0-1. Inverse dot pattern from Figure 1-1 as it could be perceived through an aperture mask.

Figure A-1 shows the Inverse dot pattern of the introductory part.

Appendix B - Hypotheses, Features, and Guidelines

Each step of iteration raises new research questions, features, and guidelines which lead to hypotheses. Validated hypotheses are transformed into features. The process is started with the expression of the fundamental feature for Digital Transformatives:

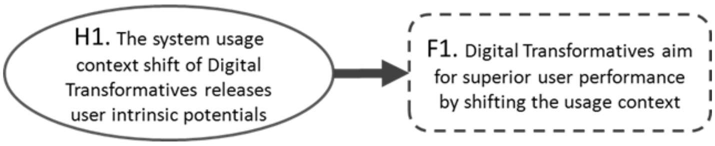
- Feature 1. Digital Transformatives aim for superior user performance by shifting the usage context. (Page 6)

This feature is assumed to be true if the following hypothesis is true.

- Hypothesis 1: The system usage context shift of Digital Transformatives releases user intrinsic potentials. (Page 6)

This fundamental hypothesis is tested comprehensively throughout the work.

Schematically the relation between the initial feature and hypothesis can be expressed in a hypotheses-feature graph.



Feature 1 is not validated, indicated by the dashed outline. The arrow shows a validation dependency. It is being assumed that DTs are able to aim for superior user performance (F1) if a system usage context shift releases user intrinsic potentials (H1). The graph is refined and further extended throughout the work.

Since our cognitive structure is heavily dependent on environmental influences, such influences may be important for Designing Digital Transformatives. Accordingly the findings are expressed in Feature 2:

- Feature 2. Similar user environments induce similar cognitive prototypes (advantages and disadvantages of cultural and social conditioning). (Page 68)
- Feature 3. Familiarity correlates to attractiveness, trust, or faithfulness. (Page 69)

Because cognitive prototypes tend to be hot spots of high user efficiency, and they are related to attractiveness and familiarity, this may be a valuable characteristic for Digital Transformative design.

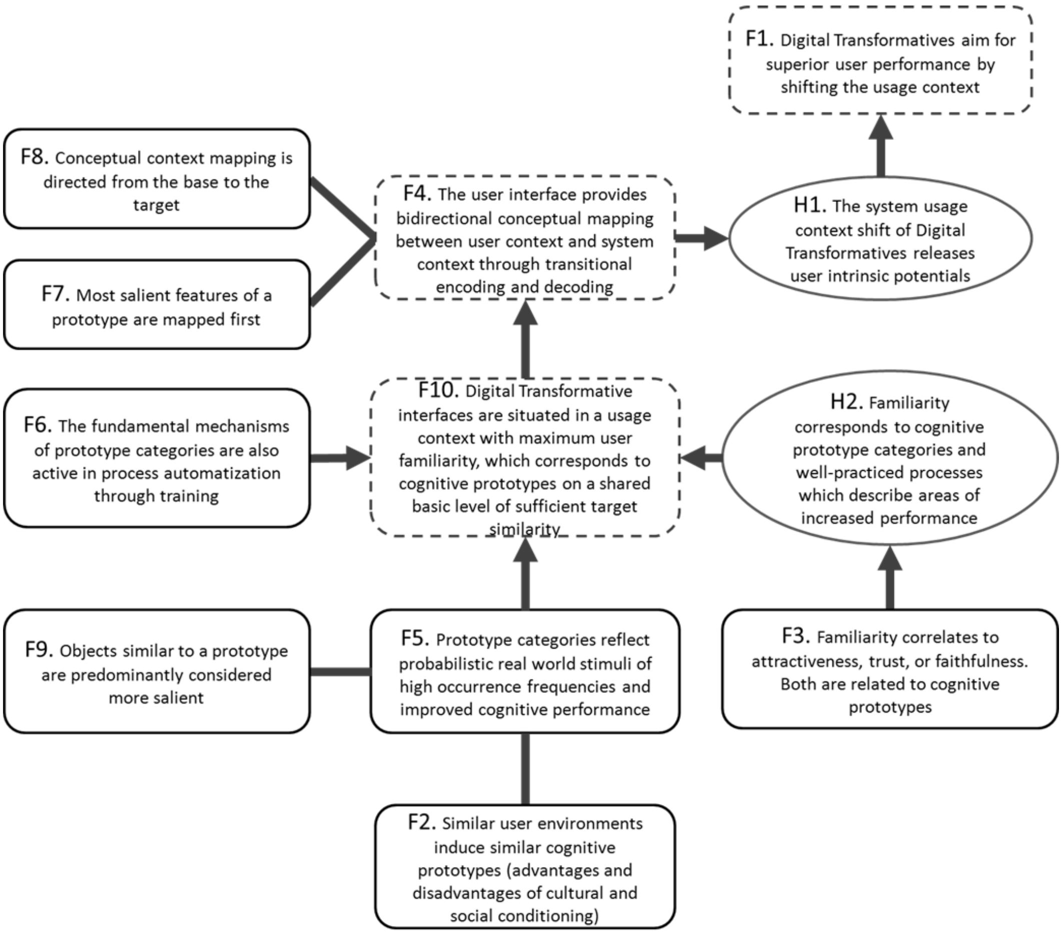
Both, Feature 2 and Feature 3 help refining the picture of Digital Transformatives; however, they are only weak implicit indicators for Hypothesis 1.

The validation of Hypothesis 1 requires an elaborate Digital Transformative concept, which is developed on conceptual and practical findings in chapter 3. According to the concept, Digital Transformative interfaces are situated in high performance user contexts. System functions need to be encoded into such contexts, while user inputs are decoded back into function contexts. The concept is expressed in Feature 4.

- Feature 4. The user interface provides a bidirectional conceptual mapping between user context and system context through transitional encoding and decoding. (Page 108)

High performance user contexts correspond to cognitive prototypes categories in human thinking. Hence, several evidential findings for cognitive research can be applied as Digital Transformative features:

- Feature 5. Prototype categories reflect probabilistic real world stimuli of high occurrence frequencies and improved cognitive performance. (Page 109)
- Feature 6. The fundamental mechanisms of prototype categories are also active in process automatization through training. (Page 109)
- Feature 7. Most salient or familiar features of a prototype are mapped first. (Page 110)
- Feature 8. Conceptual context mapping is directed from the base to the target. (Page 110)
- Feature 9. Objects similar to a prototype are predominantly considered more salient. (Page 110)
- Hypothesis 2. Familiarity corresponds to cognitive prototype categories and well-practiced processes which describe areas of increased performance. (Page 111)
- Feature 10. Digital Transformative interfaces are situated in a context with maximum user familiarity, which corresponds to cognitive prototypes on a shared basic level of sufficient target similarity. (Page 111)



The assessment of Hypothesis 1 is essential for validating Feature 1. The evaluation of Hypothesis 1 is dependent on the concept expressed in Feature 4. Feature 8 and Feature 7 are defining sub-features of Feature 4.

The concept depends on high performance usage contexts, expressed in Feature 10. Hence, Feature 4 is dependent on Feature 10. According to cognitive research, prototypes categories mark such areas of high human potentials. They are mainly formed through automatization, learning processes, and frequent occurring environmental stimuli (Feature 5, Feature 6). Salience is a major characteristic of cognitive Prototypes (Feature 9), and human beings living in similar environments tend to develop similar cognitive prototypes (Feature 2).

Moreover, the practical feature extraction, conducted in chapter 2.2, suggests the importance of familiarity, in relation to cognitive prototypes (Hypothesis 2).

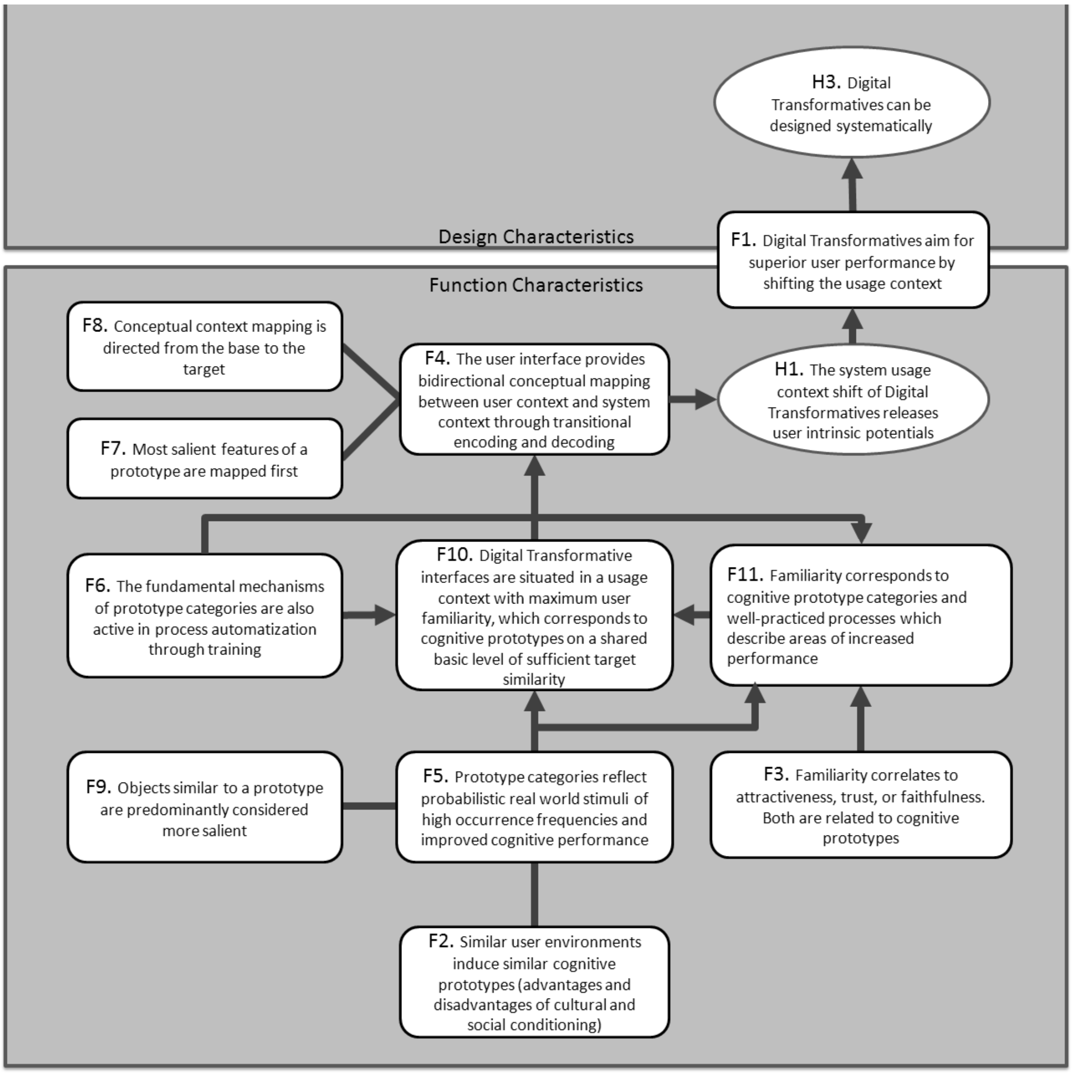
The basic concept validation of chapter 3.4.4 provided further empirical evidence on Feature 5 and Feature 6, which, in combination with Feature 3, refine validity of Hypothesis 2. Hence, Hypothesis 2 is transformed into Feature 11.

- Feature 11. Familiarity corresponds to cognitive prototype categories and well-practiced processes which describe areas of increased performance. (Page 128)

Feature 11 offers further validation for Feature 10, which is also supported by cognitive evidences expressed in Feature 6, Feature 5, Feature 9, Feature 2. Thus, Feature 10 is based on a comprehensive empirical ground, further validating Feature 4, which provides the basis for Hypothesis 1. Consequently, Hypothesis 1 becomes Feature 12:

- Feature 12. User context shifts of Digital Transformatives increase user performance and efficiency. (Page 128)

This provides us with a more refined feature graph.



After the basic concept of Digital Transformatives, and the principle of operation has been elaborated, further investigations are driven by the question, whether there is a systematic design methodology for creating Digital Transformatives, expressed in Hypothesis 3:

- Hypothesis 3: Digital Transformatives can be designed systematically. (Page 129)

Design concepts for interactive systems usually evolve from ideas gathered through brainstorming sessions. Such conceptual ideas are assessed based on previously determined requirements, experience based heuristics, and guidelines. Consequently, the basic principles of Digital Transformatives need to be implemented on the brainstorming and requirements level, which is influencing the concept design.

- Guideline 3. Basic principles of Digital Transformatives should be implemented on the brainstorming and requirements level, which is influencing the concept design. (Page 152)

Concept design guidelines

In a common interactive system development procedure, any concept idea and implementation should be assessed on overall system requirements. Requirements ought to be implemented through familiarity assessment, which, hence, has an outstanding role in the design of DTs.

- Guideline 4. Assess concepts and implementations by user familiarity. (Page 154)

The investigations conducted in this work suggest four major methods for assessing familiarity:

- Guideline 5. Use heuristic expert estimations if you quickly require tendencies. (Page 154)
- Guideline 6. If relevant user environmental data can be captured easily, conduct probabilistic analyses of occurrence frequencies of procedural and cognitive stimuli. They offer an accurate measure for cognitive prototypes and familiarity. (Page 162)
- Guideline 7. If user relevant performances are at hand or easily measurable, capture learning curves. Stagnating learning curves are indicators for areas with high procedural or cognitive user familiarity. (Page 163)
    - Feature 13. Familiarity correlates inverse proportionality to the slope of learning curves. (Page 163)
- Guideline 8. If environmental and performance measures are not available, let possible end-users rate familiarity. Familiarity ratings provide accurate measures for areas of high user potentials. (Page 171)
    - Hypothesis 4. Users are able to accurately rate their own familiarity, which reflect their performances. (Page 163)
    - Feature 14. Users are able to accurately rate their own familiarity, which reflect their performances. (Page 171)

In an iterative development cycle, Digital Transformatives are first implemented on a conceptual level. The first concept ideas evolve from empirical knowledge and may be complemented through creativity techniques, such as brainstorming. While Know-how may be acquired through experience, creativity is an unreliable factor. The investigations on the working principle of DTs provide an informative basis for creating systematic heuristics, which reduce the dependency on designer creativity and their unpredictable nature.

The working principle is expressed in Guideline 9:

- Guideline 9. Find super salient cognitive concepts in sufficient similarity proximity to original function related concepts. (Page 173)

Since DTs address cognitive concepts within possible end-users, it is ideal to actively involve the users in the concept design process.

- Guideline 10. Capture information directly from possible end-users. System design concepts should be based on a shared cognitive model of possible end-users. (Page 178)

Based on the previously given prerequisites, two procedures are proposed to identify DT concepts. The first builds on cognitive studies on categorization and findings on efficient communication, as elaborated in chapter 2.1.

- Guideline 11. Traverse superordinate concepts to find cognitive prototypes of close similarity. (Page 178)

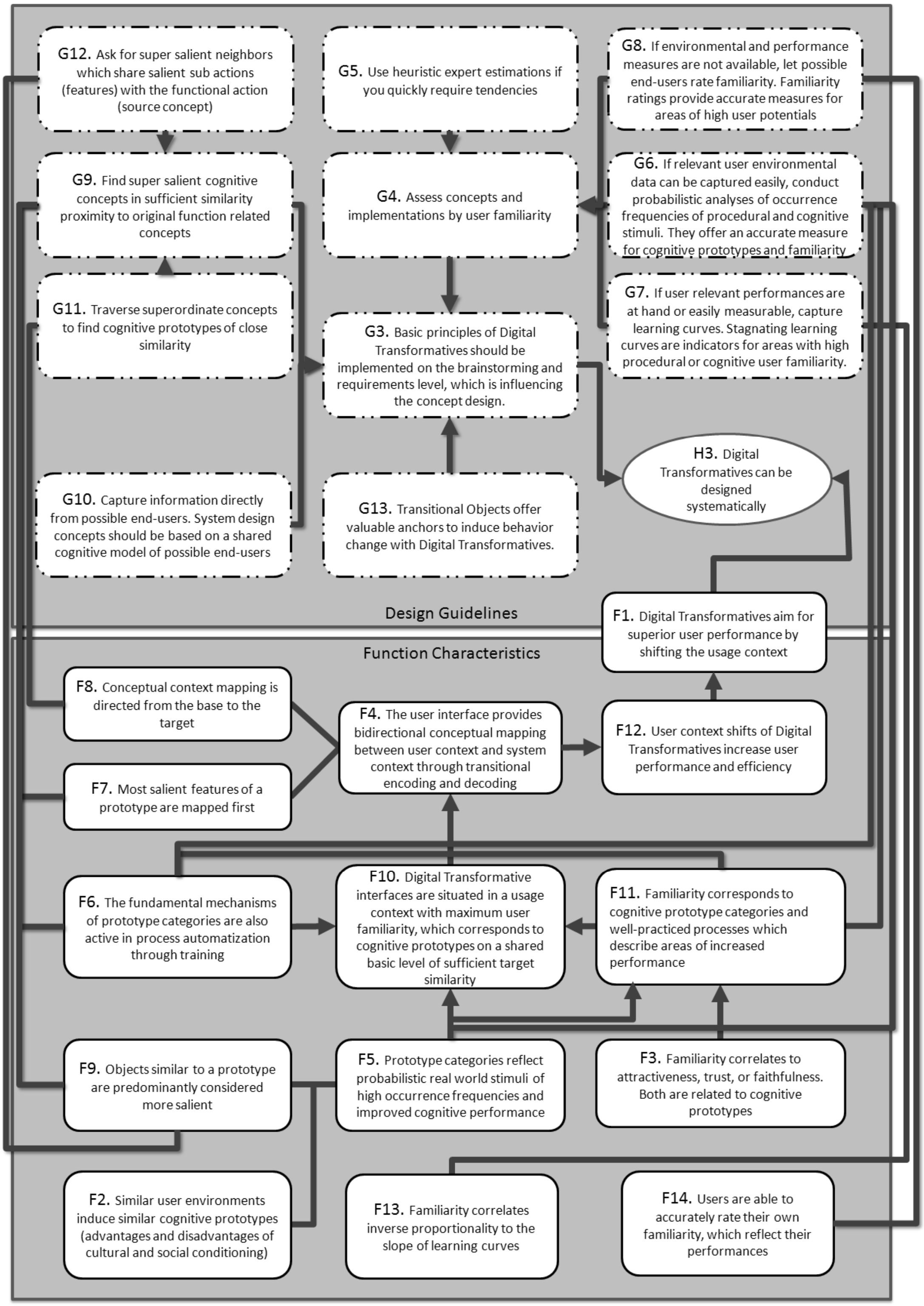
The second procedure builds on a combination of cognitive insights on similarity measures and prototypability, and on measures and results determined from requirements analyses.

- Guideline 12. Ask for super salient neighbors which share salient sub actions (features) with the functional action (source concept). (Page 183)

The final section on the practical analysis of use case prototypes revealed important empirical heuristics for inducing behavior change expressed in the last guideline.

- Guideline 13. Transitional Objects offer valuable anchors to induce behavior change with Digital Transformatives. (Page 210)

An Overview of all guidelines and features is given in the following.


G12. Ask for super salient neighbors which share salient sub actions (features) with the functional action (source concept)
G5. Use heuristic expert estimations if you quickly require tendencies
G8. If environmental and performance measures are not available, let possible end-users rate familiarity. Familiarity ratings provide accurate measures for areas of high user potentials
G9. Find super salient cognitive concepts in sufficient similarity proximity to original function related concepts
G4. Assess concepts and implementations by user familiarity
G6. If relevant user environmental data can be captured easily, conduct probabilistic analyses of occurrence frequencies of procedural and cognitive stimuli. They offer an accurate measure for cognitive prototypes and familiarity
G11. Traverse superordinate concepts to find cognitive prototypes of close similarity
G3. Basic principles of Digital Transformatives should be implemented on the brainstorming and requirements level, which is influencing the concept design.
G7. If user relevant performances are at hand or easily measurable, capture learning curves. Stagnating learning curves are indicators for areas with high procedural or cognitive user familiarity.
H3. Digital Transformatives can be designed systematically
G10. Capture information directly from possible end-users. System design concepts should be based on a shared cognitive model of possible end-users
G13. Transitional Objects offer valuable anchors to induce behavior change with Digital Transformatives.
Design Guidelines
F1. Digital Transformatives aim for superior user performance by shifting the usage context
Function Characteristics
F8. Conceptual context mapping is directed from the base to the target
F4. The user interface provides bidirectional conceptual mapping between user context and system context through transitional encoding and decoding
F12. User context shifts of Digital Transformatives increase user performance and efficiency
F7. Most salient features of a prototype are mapped first
F6. The fundamental mechanisms of prototype categories are also active in process automatization through training
F10. Digital Transformative interfaces are situated in a usage context with maximum user familiarity, which corresponds to cognitive prototypes on a shared basic level of sufficient target similarity
F11. Familiarity corresponds to cognitive prototype categories and well-practiced processes which describe areas of increased performance
F9. Objects similar to a prototype are predominantly considered more salient
F5. Prototype categories reflect probabilistic real world stimuli of high occurrence frequencies and improved cognitive performance
F3. Familiarity correlates to attractiveness, trust, or faithfulness. Both are related to cognitive prototypes
F2. Similar user environments induce similar cognitive prototypes (advantages and disadvantages of cultural and social conditioning)
F13. Familiarity correlates inverse proportionality to the slope of learning curves
F14. Users are able to accurately rate their own familiarity, which reflect their performances

Formulas

Formula 1. Similarity measure by Hampton (Hampton, 1995). 33
Formula 2. Mathematical description of the power law of practice, from (Gordon D. Logan, 2002). 39
Formula 3. Formula for determining familiarity adapted from power law of practice by (Gordon D. Logan, 2002). 162
Formula 4. Determine the rolling direction of the double cone in the double cone experiment. 200
Formula 5. A Behavior is triggered when the three factors of motivation, ability, and external triggers are exceeding the threshold of this behavior. 207

Tables

Table 2-1. The systems described in this chapter, and their position in the spectrum of tangibles described by (Ishii et al., 2012). 98
Table 4-1. Six terms for addressing all relevant interaction attributes of a usability specification by (Whiteside et al., 1988). 137
Table 4-2 Quantitative system usability measures (after (Whiteside et al., 1988) adapted by (Dix, 2004)). 138
Table 4-3. Overview of requirements data-gathering techniques, from (Rudman & Engelbeck, 1995). 140
Table 4-4. Effectiveness of methods for acquiring different types of knowledge, from (Maiden & Rugg, 1996) 142
Table 4-5. Conditions and resource constraints for different requirement acquisition methods, from (Maiden & Rugg, 1996) 142
Table 4-6. Example of a HTA task hierarchy. Notion proposed by (Annett & Duncan, 1967) 144
Table 4-7. Example of a HTA plan. Notion proposed by (Annett & Duncan, 1967). 145
Table 4-8. Questions supporting the HTA process of finding sub goals in training design (Piso, 1981) 145
Table 4-9. Questions supporting the HTA process of finding sub goals in interface design (Hodgkinson & Crawshaw, 1985) 146
Table 4-10. Questions supporting the HTA process of finding sub goals in job design (Bruseberg & Shepherd, 1997). 146
Table 5-1. Initial comparative assessment guidelines. 239
Table 5-2. Experiences of the action modeling web module 241
Table 5-3. Guidelines on how to create tools for capturing cognitive prototypes. 251
Table 6-1. Methods for assessing familiarity. 262

Figures

Figure 1-1. Random dots on a sphere. 1
Figure 1-2. Connection between system functions and system user interface of a classical mechanical system (left). More system design flexibility through digital connection (middle) allows for digital transformations (right). 4
Figure 1-3. Schematic view of the iterative research methodology of this work. A reduction of the fuzziness of the initial idea correlates to increasing elaboration of the conceptual framework. 5
Figure 1-4. Initial hypotheses-feature graph. 6
Figure 1-5. Reducing effort through information encoding and decoding as applied in mnemonic devices. The length of the arrows indicates the effort. The transformation steps of the marble sphere give a concrete example. 8
Figure 1-6. Problem based system user interface providing access to support functionality. Illustrating the task of memorizing a numeric access code. 9
Figure 1-7. The schema of a Digital Transformative. Opposed to common digital systems, shown in Figure 1-6. 10
Figure 2-1. Fundamental cognitive (top-left) and practical investigations (bottom-left) for developing a Digital Transformative concept. 13
Figure 2-2. Cognitive mechanisms related to Digital Transformatives. 14
Figure 2-3. Classical Schema of Mnemonic Devices after (Voigt, 2001). 16
Figure 2-4. Achillea Millefolium – for Beniowski looked like a house covered with snow. The Eagle (Aquila) with thousand (mille) feathers / leafs (folium). 17
Figure 2-5. The Digital Transformatives schema transferred to human communication. 20
Figure 2-6. Working principle of conceptual metaphors. The source concept (vehicle) inherits most salient features to the target concept (tenor). Some features are more salient than others, indicated through different line-thickness. 23
Figure 2-7. Ratings of members of the category tree, the more typical the closer to the center; adopted from (Malt & Smith, 1984). 25
Figure 2-8. Focal color hierarchy determined by (Berlin & Kay, 1969). This hierarchy was later revised, modified, extended and transformed into a set of categories (P. Kay, Berlin, Maffi, Merrifield, & others, 1997). 26
Figure 2-9. Focal colors distribution, visualized on a Munsell Grid, determined through the first experiments conducted by (Berlin & Kay, 1969). Visualization adapted from (Fred Hatt, 2011); also compare color chart provided by (Richard Cook et al., 2012). 27
Figure 2-10. The color term centroids of the tested 110 languages in CIE L*a*b* coordinates. A 3D visualization (left). A top view relief map, with 100

occurrences per line (right). The dots mark the English color terms. From (P. Kay & Regier, 2003). 29
Figure 2-11. Ambiguous image for testing cognitive representations. From (Chambers & Reisberg, 1985). 31
Figure 2-12. Prototype faces used in the experiments of (Solso & McCarthy, 1981) (left). Confidence rating of the faces directly after (filled markers) and after six weeks (unfilled markers) (right). From (Solso & McCarthy, 1981). 32
Figure 2-13. Varying reaction times for assessment category affiliation after priming with different levels of category typicality from (E. Rosch, 1975b). 36
Figure 2-14. Performances on remembering actual and random chess positions distinguished between experts and novice performances (Chase & Simon, 1973). 37
Figure 2-15. Reading speed and error rate change over time during training. From (LaBerge & Samuels, 1974). 38
Figure 2-16. Visualization of the power law of practice (Sternberg, 2008). 39
Figure 2-17. Reading speed latency of 3 graders to college students for words of 3 to 6 characters length from (Samuels et al., 1978). 40
Figure 2-18. Initial Projection: temporal asymmetry of metaphors from (D. Gentner et al., 2001). 45
Figure 2-19. Two, partly different conceptual models, of categories shown in hierarchical organization. Sub-ordinates inherit salient features from super ordinates. Black connecting lines illustrate possible informative richness at different matching levels in the concept hierarchy. 47
Figure 2-20. Drawing used by Rosch (E. H. Rosch, 1973b) for evaluating prototype categories in the recognition of shapes. From (Ungerer & Schmid, 2006). 48
Figure 2-21. Property extraction methodology. 55
Figure 2-22. Digital system types addressing major learning categories of Blooms Taxonomy. 59
Figure 2-23. JogLab online word finder supporting the creation of acrostics and backronyms. 60
Figure 2-24. Screenshot of the 2know software for supporting the major memorization technique (right). 60
Figure 2-25. Supporting imagery link provided by the rebus system for the abstract password “kou-ce-hur” (King, 1991). 61
Figure 2-26. The Draw A Secret user input (left). The grid based interpretation of the system (middle). The transformed bit representation (right); from (Jermyn et al., 1999). 62
Figure 2-27. PassFaces™ learning screen (left) and selection screen (right) (passfaces.com, 2012). 63
Figure 2-28. Cognitive Authentication Scheme (Weinshall, 2006). 63

Figure 2-29. PassPoints interface, showing feedback after points have been chosen (left) (Biddle et al., 2011). Blonder patent for graphical passwords (right) (Blonder, 1996). 64
Figure 2-30. Story system proposed by (left) (Davis et al., 2004). Password encoding through a "Repeated Sequence of Actions" (right) (Abdullah et al., 2008). 65
Figure 2-31. Visualizations of four digit passwords frequencies from (DataGenetics, 2013). 67
Figure 2-32. Screenshot of the first commercially sold Xerox 8010, in 1981 (guidebookgallery.org, 2012). 70
Figure 2-33. Screenshot Monkey Island 2: LeChuck's Revenge (left) (Warren, 2003). Screenshot Löwenzahn (right). 72
Figure 2-34. Screen capture of Microsoft Bob (Rose, 2008). 72
Figure 2-35. Screen capture of Packard Bell Navigator 3.5 (left) (Toastytech.com, 2012b), and Magic Cap (right) (Halfhill & Reinhardt, 1994). 74
Figure 2-36. Typical computer game design pattern. The user interface is created based on a game concept. The game concept is developed for delivering a joyful experience in the game context. This experience is enabled through an underlying hidden simulation system. 77
Figure 2-37. Screenshot of The World Water Game (left), and Power of Research (right) (Adib, 2006; Powerofresearch.eu, 2012). 80
Figure 2-38. Online stock market platform Wall Street Survivor (left). Screenshot of the multiplayer online 3D ego shooter America's Army (right) (Wallstreetsurvivor.com, 2012) (US Army, 2012). 81
Figure 2-39. Games with a Purpose: The ESP Game (left), Tagatune (right)(Google Inc., 2011; E. L. M. Law et al., 2007; Von Ahn & Dabbish, 2004). 83
Figure 2-40. Screenshot of Attent with Serios as an Email extension (left) and in general (right) (Seriosity Inc, 2010). 84
Figure 2-41. The Nethernet toolbar (thenethernet.com, 2012). 84
Figure 2-42. Bottle Bank Arcade systems for changing behavior with fun elements (Volkswagen & thefuntheory.com, 2009a). 85
Figure 2-43. Real world systems for changing behavior with fun elements. Piano Staircase (left), The Speed Camera Lottery (right) (Volkswagen & thefuntheory.com, 2009b, 2010). 86
Figure 2-44. Children playing with a LOGO Turtle (Logo Foundation, 2000). 90
Figure 2-45. Topobo interactive learning toy (Raffle, Parkes, & Ishii, 2004). 92
Figure 2-46. The transitional object interfacing the learners knowledge and new target domain. The area of intersection should be sufficiently big. 93
Figure 2-47. ARGolf (left), RobertAR(right) (Buchholz & Brosda, 2012). 93
Figure 2-48. The tangible user interface systems metaDESK (Ullmer & Ishii, 1997) 95

Figure 2-49. Phoxel-Space (left), and Glume (right) (Parkes et al., 2006; Ratti et al., 2004). 96
Figure 2-50. Kinetic TUI Relief (left). Augmented Reality based TUIs Tiles (right) (Daniel Leithinger, Kumpf, & Ishii, 2009; Poupyrev et al., 2001). 97
Figure 2-51. Augmented Reality based TUIs in ARTHUR (Broll et al., 2004). 98
Figure 2-52. Accumulated context shift properties. The fundamental cognitive properties of meaningfulness (increased knowledge elaboration) and familiarity (typicality) can be found in all context shifts. 102
Figure 3-1: Schema of a mnemonic device (left): Familiar known information is linked onto new target information, predominantly based on salient features (middle). This allows us to use cognitive categories to improve efficiency in communication (right) 106
Figure 3-2. Refinement of the Digital Transformatives schema in human communication. 107
Figure 3-3. Concept for efficient human machine interfaces, derived from efficiency enhanced cognitive processes. 108
Figure 3-4. Context shift property tag cloud sorted by frequency. Familiarity and meaningfulness appeared to be common characteristics of all context shifts. 110
Figure 3-5. Initial features, hypotheses, and their evidential dependencies. Features with little evidence are marked with dotted lines. Dependencies are marked through arrows, and defining sub-features are connected through lines. 114
Figure 3-6. Age distribution of the test group and control group. 118
Figure 3-7. The game interface of the memory test environment. The screenshots display the interface for a visual set of cards (left), and an audio set (right). 118
Figure 3-8. Average turns needed by the participants of the test group, who should be sensible to familiarity differences. 119
Figure 3-9. Average turns needed by the participants of the control group. All sets were unfamiliar to those members. 120
Figure 3-10. The results of the familiar test group (dashed line) and the control group (solid line) distinguished between familiar and unfamiliar. 121
Figure 3-11. Correlation between performance and un-familiarity of teams of the 1. Bundesliga (not all shown teams are labeled). 125
Figure 3-12. Correlation between performance and un-familiarity of teams of the 2. Bundesliga (not all shown teams are labeled). 126
Figure 3-13. Unfamiliarity regresses logarithmic reciprocal to a linear progression of the relation between performance and un-familiarity, with a reliability of over 90%. 127
Figure 3-14. Digital Transformative Function characteristics feature graph. 129
Figure 4-1. Iterative individualized product development and quality assurance before industrialization (top). Product development right after

industrialization started (bottom). Back to four step assurance (top), after three step approach proofed to fail. 132
Figure 4-2. Interaction design process as a water fall model including an iteration cycle (from Dix, 2004). 134
Figure 4-3. Interval of acceptance of measurement levels in a usability specification (adapted from (Dix, 2004; Whiteside et al., 1988)). 149
Figure 4-4. General steps of Requirements Engineering. 150
Figure 4-5. Iterative interaction design process as a water fall model including an iteration cycle (adapted from Dix, 2004). 151
Figure 4-6. First adaption of the interactive system design process to foster the design of Digital Transformatives. 153
Figure 4-7. Universal focal colors determined by (Berlin & Kay, 1969). 155
Figure 4-8. A screenshot of the Color Counter application. 158
Figure 4-9. Quantized approximation of the accumulated perception of color frequencies in daily life (logarithmic scaling). 159
Figure 4-10. 3D representation (top) and relief (bottom) of perceived frequencies color on a 8 bit RGB (hue; brightness; frequency) color space; all logarithmic scaling. 160
Figure 4-11. Clusters of color frequencies stimulating perception in an environment in Germany, compared to English focal colors determined in San Fransisco bay area by (Berlin & Kay, 1969). 161
Figure 4-12: User determined familiarity test based on memorizing faces. Screencasts showing the memorization phase (left), and the assignment phase (right). 165
Figure 4-13. Drag and drop familiarity assessment on a vertical scale from bottom, for unfamiliar, to top, for familiar items. 166
Figure 4-14. The second test with generally known and unknown faces including the statistics shown after each run (left). The final high-score and voucher for the users (right). 167
Figure 4-15. Comparative familiarity assessment of items in a memory test. Separated between familiar and unfamiliar (left). 168
Figure 4-16. Average normalized error rate related to unfamiliarity. 169
Figure 4-17. Error distribution of items rated with familiarity of 0-40, 40-60, and 60-100. 169
Figure 4-18. Familiar items were first set corresponding to lower error rates. 170
Figure 4-19. Visualization of the cognitive design principle of Digital Transformatives. 173
Figure 4-20. A visualization of the challenge of finding similar high performance cognitive concepts similar to the source concept. 174
Figure 4-21. Accumulated English search queries at google.com. The query frequencies may represent a model for shared cognitive concepts. 177
Figure 4-22. Visualization of relation between source concept (marked as a blue glowing sphere) and demanded salient prototypes and super prototypes. 179

Figure 4-23. Visualization of determining super salient prototypes. 180
Figure 4-24. Schematic procedural overview of the salient super prototype identification approach. 181
Figure 4-25. Sub-Action Modeling Approach. Four steps to find target contexts of high familiarity and sufficient similarity to the source context. 183
Figure 4-26. Schema of a possible Digital Transformatives design process. 186
Figure 5-1. Illustration of the implementation of the overall research methodology. 189
Figure 5-2. Greetings from Chilogne: Illustration of the One World One City Digital Transformative design concept. 191
Figure 5-3. Digital Transformative concept for navigation through familiar landmarks. 192
Figure 5-4. Three steps from a virtual to a tangible interactive Augmented Reality model 194
Figure 5-5. 2nd generation airfoil wing in a box (middle). Augmentation on a mobile Device (bottom). 195
Figure 5-6. The Science Center To Go setup for the Double Cone experiment. 196
Figure 5-7. The Mini Wing exhibit augmented through the airflow. Two vectors are displaying the lift and drag. 197
Figure 5-8. Sound waves of a fire truck displayed at the Doppler exhibit. 198
Figure 5-9. The double slit experiment simulating a wave field. 199
Figure 5-10. The double cone exhibit - consisting of two rails, two ramps, and four roller objects. In the double cone selection interface buttons are shown in the bottom, the resulting formula is displayed at the top of the screen. 200
Figure 5-11. The Boltzmann exhibit: Heating plate in the front refrigerator in the back. The user holds the thermometer into an area of low energy, the molecules here are moving slower. The Boltzmann distribution is shown in in the bottom left. 201
Figure 5-12. The Mixed Reality context of the Science Center To Go, aiming for a natural tangible interface to formulas. 203
Figure 5-13. Evaluation results of students rating the motivation and usability of the Mini Wing exhibit (Larsen & Bogner, 2012; Larsen et al., 2012). 205
Figure 5-14. Teachers' rating of the general usage of the system (Larsen & Bogner, 2012; Larsen et al., 2012). 205
Figure 5-15. Persuasive Technologies aim at influencing human behavior while Digital Transformatives are trying to adapt computer interfaces to human behavior. 209
Figure 5-16. Concept of a system for direct reflection and logging of drinking behavior. 211
Figure 5-17. First concept drawing showing the setup of the digital beverage coaster. 212
Figure 5-18. The graphical user interface of the digital beverage coaster. Only peripheral information is shown on the desktop if the beverage is on the

coaster (left). An beverage change interface is shown as soon as the glass is lifted off the coaster (right). 213
Figure 5-19. Drinking behavior of test person T1 with drinking amount display during the test phase (bright bars). Control values were gathered without giving any feedback on drinking amount (dark bars). The polynomial trend line of the test phase with stability index R^2=0,7672. 214
Figure 5-20 Drinking behavior of test person T2 with drinking amount display during the test phase (bright bars). Control values were gathered without giving any feedback on drinking amount (dark bars). The polynomial trend line of the test phase with stability index R^2=0,2771. 215
Figure 5-21. Drinking behavior of test person T3 with drinking amount display during the test phase (bright bars). Control values were gathered without giving any feedback on drinking amount (dark bars). The polynomial trend line of the test phase with stability index R^2=0,3849. 216
Figure 5-22. The Drinking Garden user interface. For raising long time awareness, users of the system need to care for a virtual garden. They water virtual plants by drinking. 217
Figure 5-23. Workload model of a non-comparative assessment. 222
Figure 5-24. Workload model of a comparative assessment. 223
Figure 5-25. DT principles involved in deriving absolute user assessments from comparative inputs. 224
Figure 5-26. Pre-Test: Comparative and non-comparative assessment of brightness using a pen and paper prototype. 225
Figure 5-27. Test User interface: Five sliders had to be arranged from top (dark) to bottom (bright). The sliders appeared sequentially from left to right. The two consecutive states between setting slider 4 (left) and showing slider 5 (right) are shown above. Users were engaged to adapt former sliders whenever desired. 226
Figure 5-28. High score, history and general distribution statistics were provided after each assessment run. 228
Figure 5-29. Absolute average errors for each slider at first assessment and after corrections. 228
Figure 5-30. Time to continue with the next slider (including a 1 second fade in effect of the user interface). 229
Figure 5-31. How many users continued their input without correcting previous sliders, determined relatively to those who continued. 230
Figure 5-32. Number of corrections per slider, and improvement per correction in percent. 231
Figure 5-33. User input error for each control value; with and without corrections. Trend lines are of 4th degree. 232
Figure 5-34. Accumulated distribution of user inputs before (stitched line), after corrections (bold line), and control values (gray bars). 233

Figure 5-35. average positive and negative absolute error per slider number with 4th order polynomial trend lines. 234
Figure 5-36. Comparative distribution of the error in respect to the relative difference of the control values. 236
Figure 5-37. Iterative development of the Sub-Action modeling web module, starting off with paper prototypes (top-left), an experimental online interface showing cascaded action lines (top-right), and the final hierarchical bullet point based interface for modeling incomplete action chains (bottom). 240
Figure 5-38. ColleaGuess user interface consisting of an activity stream, and competitive game elements, 243
Figure 5-39. The game logic adapted from super modeling approach procedure shown in Figure 4-24. Gray dashed lines indicate transitions of the original procedure. 244
Figure 5-40. Transformative mechanisms of the super salient concept modeling game. 246
Figure 5-41. Diagram visualizing results of multiple studies on the relation of reading speed and age, from (Hunziker, 2006). 248
Figure 5-42. Screenshots of a cooperative online speed reading game for capturing cognitive concept structures. Session setup interface (left) and in game interface (right). 249
Figure 6-1. Digital Transformative basic schema for reducing user effort by offering an interface in a user efficient context. Information encoding and decoding becomes part of the design, and needs to be implemented digitally by the system. 253
Figure 6-2. Examples for DTs: Android Pattern Lock (left) and Papert's LOGO Turtle (right) (Logo Foundation, 2000; Meacham, 2013). 254
Figure 6-3. Context shift property tag cloud sorted by frequency. Familiarity and meaningfulness appeared in all context shifts of the investigated DT systems. 255
Figure 6-4. Efficiency model of human communication. Deduced from evidential cognitive research. 256
Figure 6-5. Schematic representation of the Digital Transformatives concept and its working principle. 257
Figure 6-6. The working principle of Digital Transformatives referred to the visualization of a salience-similarity space of cognitive concepts. 258
Figure 6-7. The DT design challenge (left). Salient Super Prototypes Approach for systematically identifying target contexts of high salience and sufficient similarity to the source context (right). 259
Figure 6-8. Iterative identification procedure of salient super prototypes. 260
Figure 6-9. Iterative identification of similar salient prototypes through sub-action modeling. 261
Figure 6-10. User interface of an application developed for analyzing the probabilistic environmental color stimuli of test-persons (top). Test results

(middle) compared to the a distribution found through the Basic Color Terms study (Berlin & Kay, 1969). 263
Figure 6-11. Extensive test on accuracy of user rated familiarity and its relation to performance. 264
Figure 6-12. Indicators for familiarity corresponding to practice. Familiarity in team sports in relation to performance (left). Familiarity and performance in the face memory test (right). 265
Figure 6-13. An adaption of the common interactive system design approach fostering systematic Digital Transformatives design. New aspects are highlighted in blue. 266
Figure 6-14. The user interface of a game based DT for Salient Super Prototype Modeling (left). A screenshot of the Speed Reading Game for capturing general cognitive concept structures (right). 268
Figure 6-15. Functional characteristics of Digital Transformatives. 269
Figure 6-16. Overview of identified DT function characteristics and their relation to DT design guidelines. 272
Figure 6-17. Standards, conventions, and the force of habits may improve efficiency. If those standards are broken our usage efficiency decreases, as showcased through low variances in a stair flight (left), or keyboard layouts different to QWERTY (right) (Wikipedia.org, 2012a, 2012b). 275
Figure 6-18. The Martian face, which turned out to be nothing alike (NASA Science, 2001). 276
Figure 6-19. Ambiguous images based on Gestalt grouping rules (middle-right) (Wikipedia.org, 2013a). 276
Figure 6-20. Search query amounts (logarithmic scaling) for fruit terms as an indicator for prototypability (suggestqueries.google.com). 278
Figure 6-21. Examples for company logos (Apple Inc., 2013; Castrol, 2013; Mercedes-Benz, 2013; Toyota, 2013; Unilever, 2013) 278
Figure 6-22. The Averager computes the average of faces from a selection pool (middle). The left face shows an average of the first two female faces in the top row, the right face is the average of the 25 selected faces (DeBruine & Jones, 2013). 280
Figure A-0-1. Inverse dot pattern from Figure 1-1 as it could be perceived through an aperture mask. 290

Bibliography

Abdullah, M. D. H., Abdullah, A. H., Ithnin, N., & Mammi, H. K. (2008). Towards Identifying Usability and Security Features of Graphical Password in Knowledge Based Authentication Technique (pp. 396–403). IEEE. doi:10.1109/AMS.2008.136

Abt, C. C. (1966). *Games for learning*. Educational Services Cambridge, MA.

Abt, C. C. (1970). Serious games: The art and science of games that simulate life. *New York*.

Abt, C. C. (1987). *Serious games*. Univ Pr of Amer.

Adib, R. (2006). World Water Game. Retrieved June 21, 2012, from http://world-water-game.softonic.de/

Allwood, C. M., & Eliasson, M. (1987). Analogy and other sources of difficulty in novices' very first text-editing. *International journal of man-machine studies*, *27*(1), 1–22.

Alsted, J. H. (1610). *Systema mnemonicum duplex*. Frankfurt am Main.

Anderson, J. R. (1983). *The architecture of cognition*. Harward University Press.

Anderson, L. W., Sosniak, L. A., & Bloom, B. S. (1994). *Bloom's taxonomy: A forty-year retrospective* (2nd ed., Vol. 93). Chicago: University of Chicago Press.

Annett, J., & Duncan, K. (1967). Task analysis and training design.

Annett, J., Duncan, K., Stammers, R., & Gray, M. (1971). *Task analysis*. HM Stationery Office.

Apple Inc. (2011). *SIRI (Speech Interpretation and Recognition Interface)*.

Apple Inc. (2012). Apple - OS X Lion - Das fortschrittlichste Computer-Betriebssystem der Welt. Retrieved July 22, 2012, from http://www.apple.com/de/macosx/

Apple Inc. (2013). Apple. Retrieved April 21, 2013, from http://www.apple.com/

Armstrong, S. L., Gleitman, L. R., & Gleitman, H. (1983). What some concepts might not be. *Cognition*, *13*(3), 263–308.

Artelt, C., Naumann, J., & Schneider, W. (2010). Kapitel 3 Lesemotivation und Lernstrategien. *PISA 2009*, 73.

Astley, J., & Stammers, R. (1987). Adapting hierarchical task analysis for user-system interface design. *New Methods in Applied Ergonomics. Taylor and Francis, London (175-184)*.

Atkin, A. (2010). Peirce's theory of signs.

Atkinson, R. C. (1975). Mnemotechnics in second-language learning. *American Psychologist*, *30*(8), 821.

Atkinson, R. C., & Raugh, M. R. (1975). An application of the mnemonic keyword method to the acquisition of a Russian vocabulary. *Journal of Experimental Psychology: Human Learning and Memory; Journal of Experimental Psychology: Human Learning and Memory*, *1*(2), 126.

Avedon, E. M., Sutton-Smith, B., & Sutton, B. (1971). *The study of games*. J. Wiley.

Avons, S. (1999). Effects of visual similarity on serial report and item recognition. *The Quarterly Journal of Experimental Psychology: Section A*, *52*(1), 217–240.

Bagnall, B. (2007). *Maximum Lego NXT: Building Robots with Java Brains*. Variant Press.

Banker, R. D., Datar, S. M., & Kemerer, C. F. (1987). Factors affecting software maintenance productivity: An exploratory study. In *Proceedings of the 8th International Conference on Information Systems* (pp. 160–175). Retrieved from http://www.pitt.edu/~ckemerer/CK research papers/FactorsSwMaintenanceProductivity_BankerDatarKemerer87.pdf

Baron, J. (2008). *Thinking and deciding .* (4th Edition.). Cambridge University Press.

Baydin Inc. (2010). The Email Game. Retrieved August 24, 2012, from http://emailga.me/learn.html

Becker-Carus, C., & Herbring, H. (2004). *Allgemeine Psychologie*. Spektrum Akad. Verl.

Belk, R. W., Bahn, K. D., & Mayer, R. N. (1982). Developmental recognition of consumption symbolism. *Journal of Consumer Research*, 4–17.

Beniowski, B. (1842). *A handbook of phrenotypics for teachers and students*. The Author.

Berlin, B., & Kay, P. (1969). Basic color terms: Their universality and evolution. *Berkeley and Los Angeles: University of California*.

Biddle, R., Chiasson, S., & Van Oorschot, P. C. (2011). Graphical passwords: Learning from the first twelve years. *ACM Computing Surveys*, *44*(4).

Black, A., Ducasse, S., Nierstrasz, O., Pollet, D., Cassou, D., Denker, M., & others. (2009). Pharo by example.

Black, M. L. (2012). Narrative and Spatial Form in Digital Media: A Platform Study of the SCUMM Engine and Ron Gilbert's The Secret of Monkey Island. *Games and Culture*.

Blonder, G. E. (1996). *Graphical password*. Google Patents.

Bock, R. D., & Jones, J. (1968). The measurement and prediction of judgment and choice.

Bogardus, E. S. (1925). Measuring social distance. *Journal of Applied Sociology*, *9*(January-February), 299–308.

Bolzoni, L. (2004). *The web of images: vernacular preaching from its origins to St. Bernardino of Siena*. Ashgate Pub Ltd.

Bornstein, R. F. (1989). Exposure and affect: Overview and meta-analysis of research, 1968–1987. *Psychological bulletin*, *106*(2), 265.

Bornstein, R. F., & D'Agostino, P. R. (1992). Stimulus recognition and the mere exposure effect. *Journal of personality and social psychology*, *63*, 545–545.

Bower, G. D. (2008). Chunky Experts who Devide to Conquer. In *Cognitive Psychology* (5th ed., pp. 14–15). Wadsworth Pub Co.

Bower, G. H. (1970). Analysis of a Mnemonic Device: Modern psychology uncovers the powerful components of an ancient system for improving memory. *American Scientist*, *58*(5), 496–510.

Bower, G. H. (1972). Mental imagery and associative learning.

Bower, G. H., & Clark, M. C. (1969). Narrative stories as mediators for serial learning. *Psychonomic Science; Psychonomic Science*.

Bradley, R. A., & Terry, M. E. (1952). Rank analysis of incomplete block designs: I. The method of paired comparisons. *Biometrika*, *39*(3/4), 324–345.

Broll, W., Lindt, I., Ohlenburg, J., Herbst, I., Wittkamper, M., & Novotny, T. (2005). An infrastructure for realizing custom-tailored augmented reality user interfaces. *Visualization and Computer Graphics, IEEE Transactions on*, *11*(6), 722–733.

Broll, W., Lindt, I., Ohlenburg, J., Wittkämper, M., Yuan, C., Novotny, T., … Strothman, A. (2004). Arthur: A collaborative augmented environment for architectural design and urban planning. Retrieved from http://eprints.ucl.ac.uk/7853/1/7853.pdf

Brooke, J. (1996). SUS-A quick and dirty usability scale. *Usability evaluation in industry*, *189*, 194.

Brostoff, S., & Sasse, M. A. (2000). Are Passfaces more usable than passwords? A field trial investigation. *PEOPLE AND COMPUTERS*, 405–424.

Brown, R. (1958). How shall a thing be called? *Psychological review*, *65*(1), 14.

Brown, R. (1965). Social psychology.

Brown, R. W., & Lenneberg, E. H. (1954). A study in language and cognition. *The Journal of Abnormal and Social Psychology*, *49*(3), 454.

Bruner, J. (1986). The inspiration of Vygotsky. *Actual Minds Possible Worlds*.

Bruner, JS, Goodnow, J. J., & Austin, G. A. (1956). A study of thinking. *Science Editions*.

Bruseberg, A., & Shepherd, A. (1997). Job design in integrated mail processing. *Engineering Psychology and Cognitive Ergonomics. Volume Two: Job Design and Product Design. Ashgate Publishing, Aldershot, Hampshire,(25-32).*

Bryan, W. L., & Harter, N. (1899). Studies on the telegraphic language: The acquisition of a hierarchy of habits. *Psychological Review*, *6*(4), 345.

Bryson, M., Bereiter, C., Scardamalia, M., & Joram, E. (1991). Going beyond the problem as given: Problem solving in expert and novice writers. *Complex problem solving: Principles and mechanisms*, 61–84.

Buchholz, H. (2005). *Konzept einer intelligenten Mikrowelt:Eine explorative adaptive Lernumgebung zum konstruktivistischen spielerischen Lernen durch wissensbasierte Simulation.* Otto-Von-Guericke Universität Magdeburg, Magdeburg.

Buchholz, H., & Brosda, C. (2012). Melting Interfaces – Learning in Mixed Realities. A Retrospective on Transitional Objects. In A. Lazoudis, H. Salmi, & S. Sotiriou (Eds.), *Augmented Reality in Education - Proceedings of the "Science Center To Go" Workshops* (pp. 69–78). Athens, Greece: EPINOLA S.A.

Buchholz, H., Brosda, C., & Wetzel, R. (2010). Science Center To Go - A Mixed Reality Learning Environment of Miniature. In *Proceedings of the "Learning with ATLAS@CERN" Workshops Inspiring Science Learning EPINOIA* (pp. 85–96). Crete.

Buchholz, H., & Wetzel, R. (2009). Introducing Science Centre TOGO - A Mixed Reality Learning Environment for Everyone's Pocket (pp. 104–110). Presented at the IADIS International Conference Mobile Learning, Barcelona, Spain.

Burgener, R. (1999). Twenty Questions. *The neural-net on the Internet*. Retrieved June 28, 2012, from http://web.archive.org/web/20000229091147/http://www.20q.net/index.html

Buswell, G. T. (1922). Fundamental reading habits: a study of their development.

Carruthers, M. (2000). *The craft of thought: Meditation, rhetoric, and the making of images, 400-1200* (Vol. 34). Cambridge University Press.

Carruthers, M. J. (1992). *The book of memory: A study of memory in medieval culture* (Vol. 10). Cambridge University Press.

Castrol. (2013). Castrol Motor Oils and Lubricants. Retrieved April 21, 2013, from http://www.castrol.com/castrol/castrolglobalhomeflash.do

Cavallari, B., Hedberg, J., & Harper, B. (1992). Adventure games in education: A review. *Australian Journal of Educational Technology*, *8*(2), 172–184.

Chambers, D., & Reisberg, D. (1985). Can mental images be ambiguous? *Journal of Experimental Psychology: Human Perception and Performance*, *11*(3), 317.

Chambers, D., & Reisberg, D. (1992). What an image depicts depends on what an image means. *Cognitive Psychology*, *24*(2), 145–174.

Chase, W. G., & Simon, H. A. (1973). The mind's eye in chess.

Cheifet, S., & Kildall, G. (1994). The Computer Chronicals: CES Special Las Vegas 1994 - Arabic. Las Vegas. Retrieved from http://www.youtube.com/watch?v=jQggTfZIstk

Chi, M. T. H., Glaser, R., & Rees, E. (1982). Expertise in problem solving. In R. J. Sternberg (Ed.), *Advances in the psychology of human intelligence* (Vol. 1.). Hillsdale, NJ: Erlbaum.

Cicero, M. T., & Caplan, H. (1954). *Cicero Ad C. Herennium de ratione dicendi:(Rhetorica ad Herennium)* (Vol. 403). Harvard University Press.

Clark, H. H., & Clark, E. V. (1977). *Psychology and language*. Cambridge Univ Press.

Clark, J. M., & Paivio, A. (1991). Dual coding theory and education. *Educational psychology review*, *3*(3), 149–210.

Coley, J. D., Medin, D. L., & Atran, S. (1997). Does rank have its privilege? Inductive inferences within folkbiological taxonomies. *Cognition*, *64*(1), 73–112.

Collier, G. A. (1973). Review of Basic Color Terms. *Language*, *49*, 245–248.

Collier, M. (2005). Hume and cognitive science: The current status of the controversy over abstract ideas. *Phenomenology and the Cognitive Sciences*, *4*(2), 197–207.

Conklin, H. C. (1973). Color categorization. *American Anthropologist*, *75*(4), 931–942.

Cook, R. S., Kay, P., & Regier, T. (2005). The World Color Survey database. *Handbook of categorization in cognitive science*, 223–241.

Cote, R., & Harrington, G. (2006). *NFPA 101 Life Safety Code Handbook*. National Fire Protection Association, Inc., Quincy, MA.

Coulson, S. (2001). *Semantic leaps: Frame-shifting and conceptual blending in meaning construction*. Cambridge Univ Pr.

Craik, F. I. M., Govoni, R., Naveh-Benjamin, M., & Anderson, N. D. (1996). The effects of divided attention on encoding and retrieval processes in human memory. *Journal of Experimental Psychology: General*, *125*(2), 159.

Craik, F. I. M., & Lockhart, R. S. (1972). Levels of processing: A framework for memory research1. *Journal of verbal learning and verbal behavior*, *11*(6), 671–684.

Craik, F. I. M., & Tulving, E. (1975). Depth of processing and the retention of words in episodic memory. *Journal of Experimental Psychology: General*, *104*(3), 268.

Craik, K. J. W. (1943). *The nature of explanation*. Cambridge University Press.

Crowley, S., & Hawhee, D. (2004). *Ancient rhetorics for contemporary students* (3rd Edition.). Longman Publishing Group.

Da Signa, B. (1892). Rhetorica novissima. *Scripta anecdota glossatoria*.

Darby, S. (2006). The effectiveness of feedback on energy consumption. *A Review for DEFRA of the Literature on Metering, Billing and direct Displays*, *486*. Retrieved from http://powerwatch.biz/site/wp-content/uploads/2012/02/smart-metering-report.pdf

DataGenetics. (2013). PIN number analysis. Retrieved April 15, 2013, from http://www.datagenetics.com/blog/september32012/

Davidoff, J. (2001). Language and perceptual categorisation. *Trends in Cognitive Sciences*, *5*(9), 382–387.

Davidoff, J., Davies, I., & Roberson, D. (1999). Colour categories in a stone-age tribe. *Nature*, *398*(6724), 203–204.

Davis, D., Monrose, F., & Reiter, M. K. (2004). On user choice in graphical password schemes. In *13th USENIX Security Symposium* (pp. 151–164).

De Angeli, A., Coventry, L., Johnson, G., & Renaud, K. (2005). Is a picture really worth a thousand words? Exploring the feasibility of graphical authentication systems. *International Journal of Human-Computer Studies*, *63*(1), 128–152.

De Bruin, W. B., Fischhoff, B., Millstein, S. G., & Halpern-Felsher, B. L. (2000). Verbal and Numerical Expressions of Probability: "It's a Fifty–Fifty Chance." *Organizational Behavior and Human Decision Processes*, *81*(1), 115–131. doi:10.1006/obhd.1999.2868

De Groot, A. D. (1946). *Het denken van den schaker*. Noord-Holl. Uitg. Mij.

De Groot, A. D. (1978). *Thought and choice in chess* (Vol. 4). Mouton De Gruyter.

DeBruine, L., & Jones, B. (2013). Face Research ⇒ Demos ⇒ Make An Average. *Face Research ⇒ Experiments about face and voice preferences*. Retrieved April 13, 2013, from http://faceresearch.org/demos/average

Deming, W. E. (2000). *Out of the Crisis*. The MIT Press.

Derose, S. J. (2005). "Eskimo" words for snow. Retrieved July 23, 2012, from http://www.derose.net/steve/guides/snowwords/

Deterding, S., Dixon, D., Khaled, R., & Nacke, L. (2011). From game design elements to gamefulness: defining gamification. In *Proceedings of the 15th International Academic MindTrek Conference: Envisioning Future Media Environments* (pp. 9–15).

Deutsche Fußball Liga GmbH. (2012). Bundesliga - official website. Retrieved September 12, 2012, from http://www.bundesliga.com/en/

Dillon, T. (2004). *Adventure games for learning and storytelling*. Citeseer.

DIN. (2013). Deutsches Institut für Normung : Homepage EN. Retrieved April 22, 2013, from http://www.din.de/cmd;jsessionid=607F17E92F875ADF43D7BE47B3D8772B.1?level=tpl-home&languageid=en

DIN EN ISO 14122-1. Deutsches Institut für Normierung (2002).

DIN EN ISO 14122-3. Deutsches Institut für Normierung (2002).

Dirik, A. E., Memon, N., & Birget, J. C. (2007). Modeling user choice in the PassPoints graphical password scheme. In *Proceedings of the 3rd symposium on Usable privacy and security* (pp. 20–28).

Dix, A. (2004). *Human-computer interaction*. Prentice hall.

Dommerich, J. C. (1765). *Die Mnemonik und Heuristik nach ihren ersten Zügen entworfen*. Halle / Helmstedt.

Douglas, J. H. (1987). Task appropriateness of mnemonic techniques. *Perceptual and Motor Skills*, *64*(1), 171–178.

Douglas, S. A., & Moran, T. P. (1983). Learning text editor semantics by analogy. In *Proceedings of the SIGCHI conference on Human Factors in Computing Systems* (pp. 207–211).

Ducasse, S., Renggli, L., Shaffer, C. D., & others. (2010). *Dynamic Web Development with Seaside*. Lulu Enterprises.

Edwards, W. (1968). Conservatism in human information processing. *Formal representation of human judgment*, 17–52.

Engle, R. W., Tuholski, S. W., Laughlin, J. E., & Conway, A. R. A. (1999). Working memory, short-term memory, and general fluid intelligence: a latent-variable approach. *Journal of Experimental Psychology: General*, *128*(3), 309.

Enos, R. L. (2005). Rhetorica Ad Herennium. *Classical rhetorics and rhetoricians: critical studies and sources*, 331.

Ericcson, K., Chase, W. G., & Faloon, S. (1980). Acquisition of a memory skill. *Science*, *208*(4448), 1181–1182.

Espinosa, J. A., Kraut, R. E., Slaughter, S. A., Lerch, J. F., Herbsleb, J. D., & Mockus, A. (2001). Shared Mental Models, Familiarity and Coordination: A Mulit-Method Study of Distributed Software Teams. Retrieved from http://repository.cmu.edu/hcii/95/

Fauconnier, G., & Turner, M. (2003). *The way we think: Conceptual blending and the mind's hidden complexities*. Basic Books.

Fehr, B., & Russell, J. A. (1984). Concept of emotion viewed from a prototype perspective. *Journal of experimental psychology: General*, *113*(3), 464.

Feldmeier, D., & Karn, P. (1990). Unix password security-ten years later. In *Advances in Cryptology—CRYPTO'89 Proceedings* (pp. 44–63).

Fischhoff, B., & Bruine De Bruin, W. (1999). Fifty–fifty= 50%? *Journal of Behavioral Decision Making*, *12*(2), 149–163.

Fodor, J. (1994). Concepts: A potboiler. *Cognition*, *50*(1-3), 95–113.

Fodor, J. A., & Katz, J. J. (1963). The structure of a semantic theory. *Language*, *39*(2), 170–210.

Fogg, B. (2009a). A behavior model for persuasive design. In *Proceedings of the 4th international Conference on Persuasive Technology* (p. 40).

Fogg, B. (2009b). The Behavior Grid: 35 ways behavior can change. In *Proceedings of the 4th international Conference on Persuasive Technology* (p. 42).

Fogg, B. (2009c). The new rules of persuasion. *RSA Journal*, 24–29.

Fogg, B. J. (2008). *Fogg on Simplicity*. Retrieved from http://vimeo.com/2094487

Fogg, B. J. (2011). BJ Fogg's Behavior Model. Retrieved September 27, 2012, from http://www.behaviormodel.org/index.html

Fogg, B. J., Cuellar, G., & Danielson, D. (2007). Motivating, influencing, and persuading users. In A. Sear & J. A. Jack (Eds.), *The human-computer interaction handbook* (pp. 133–147). CRC. Retrieved from http://bjfogg.com/hci.pdf

Franks, J. J., & Bransford, J. D. (1971). Abstraction of visual patterns. *Journal of Experimental Psychology*, *90*(1), 65.

Fraunhofer FIT. (2011a). Energie-Monitoring - Fraunhofer-Institut für Angewandte Informationstechnik. Retrieved September 28, 2012, from http://www.fit.fraunhofer.de/de/fb/cscw/projects/energie-monitoring.html

Fraunhofer FIT. (2011b). Intelligente Smart Metering-Systeme - Fraunhofer-Institut für Angewandte Informationstechnik. Retrieved September 28, 2012, from http://www.fit.fraunhofer.de/de/smartmetering.html

Fraunhofer IAIS. (2012). Fraunhofer IAIS: Roberta. Retrieved August 28, 2012, from http://www.iais.fraunhofer.de/index.php?id=4822&L=1

Fred Hatt. (2011). The Full Gamut fredhatt-b&k-american-english-color- – drawing life. Retrieved July 13, 2012, from http://fredhatt.com/blog/2011/10/23/the-full-gamut/fredhatt-bk-american-english-color/

Fullerton, T., Swain, C., & Hoffman, S. (2008). *Game design workshop: a playcentric approach to creating innovative games*. Morgan Kaufmann.

Fussballdaten Verlags GmbH. (2012). Die aktuelle Bundesliga 2012/2013 - Der 3. Spieltag - Fussballdaten - Die Fußballdatenbank. Retrieved September 12, 2012, from http://www.fussballdaten.de/bundesliga/

Galloway, A. R. (2004). Social realism in gaming. *Game Studies*, *4*(1).

Garcia, G., & Diener, D. (1993). *Do you remember ... A comparison of mnemonic strrategies.* Presented at the Annual meeting of the WPA, Phoenix, AZ.

Gaver, W. W. (1991). Technology affordances. In *Proceedings of the SIGCHI conference on Human factors in computing systems: Reaching through technology* (pp. 79–84). Retrieved from http://dl.acm.org/citation.cfm?id=108856

General Magic. (1994). *Magic Cap.*

Gentener, D., Bowdle, B., & Ortony, A. (2008). Metaphor as structure-mapping. *Metaphor and thought*, 109–128.

Gentner, D. (1983). Structure-mapping: A theoretical framework for analogy. *Cognitive science*, *7*(2), 155–170.

Gentner, D. (1988). Metaphor as structure mapping: The relational shift. *Child development*, 47–59.

Gentner, D. (2003). Why we're so smart. In D. Gentner & S. Goldin-Meadow (Eds.), *Language in mind: Advances in the study of language and cognition* (pp. 195–236). Cambridge, MA:: MIT Press.

Gentner, D., Bowdle, B., Wolff, P., & Boronat, C. (2001). Metaphor is like analogy. *The analogical mind: Perspectives from cognitive science*, 199–253.

Gentner, D., & Christie, S. (2008). Relational language supports relational cognition in humans and apes. *Behavioral and Brain Sciences*, *31*(02), 136–137.

Gentner, D., & Clement, C. (1988). Evidence for relational selectivity in the interpretation of analogy and metaphor. *The psychology of learning and motivation*, *22*, 307–358.

Gentner, D., & Goldin-Meadow, S. (2003). Whither whorf. *Language in mind: Advances in the study of language and cognition*, 3–14.

Gentner, D., & Markman, A. B. (1997). Structure mapping in analogy and similarity. *American Psychologist*, *52*(1), 45.

Gentner, Dedre, & Gentner, D. R. (1982). *Flowing Waters or Teeming Crowds: Mental Models of Electricity.* DTIC Document.

Gentner, Don, & Nielsen, J. (1996). The anti-Mac interface. *Communications of the ACM*, *39*(8), 70–82.

Gerstadt, C. L., Hong, Y. J., & Diamond, A. (1994). The relationship between cognition and action: performance of children-7 years old on a stroop-like day-night test. *Cognition*, *53*(2), 129–153.

Gesellschaft für DFB-Online mbH. (2012). DFB - Deutscher Fußball-Bund e.V. - Archiv. Retrieved September 12, 2012, from http://www.dfb.de/index.php?id=82918

Gibbs, R. (2008). The Cambridge handbook of metaphor and thought. *Recherche*, *67*, 02.

Gibson, J. (1979). The ecological approach to visual perception. *of Experimental Psychology Human Perception and*, *39*(2).

Gibson, J. J. (1977). The theory of affordances. *Perceiving acting and knowing Toward an ecological psychology*, 67–82.

Gilbert, L. C. (1953). Functional motor efficiency of the eyes and its relation to reading. *University of California Publications in Education*.

Ginsburg, H. P., & Opper, S. (1988). *Piaget's theory of intellectual development* . Prentice-Hall, Inc.

Glonnegger, E. (1988). *Das Spiele-Buch: Brett-und Legespiele aus aller Welt: Herkunft, Regeln und Geschichte*. H. Hugendubel, O. Maier, c1988.

Glucksberg, S., & Keysar, B. (1990). Understanding metaphorical comparisons: Beyond similarity. *Psychological review*, *97*(1), 3.

Glucksberg, S., McGlone, M. S., & Manfredi, D. (1997). Property attribution in metaphor comprehension. *Journal of memory and language*, *36*, 50–67.

Gobet, F., Lane, P. C. R., Croker, S., Cheng, P. C. H., Jones, G., Oliver, I., & Pine, J. M. (2001). Chunking mechanisms in human learning. *Trends in cognitive sciences*, *5*(6), 236–243.

Gobet, F., & Simon, H. A. (1996a). Recall of random and distorted chess positions: Implications for the theory of expertise. *Memory & cognition*, *24*(4), 493–503.

Gobet, F., & Simon, H. A. (1996b). Templates in chess memory: A mechanism for recalling several boards. Retrieved from http://bura.brunel.ac.uk/handle/2438/1339

Gobet, F., & Simon, H. A. (1996c). The roles of recognition processes and look-ahead search in time-constrained expert problem solving: Evidence from grandmaster-level chess. *Psychological Science*, 52–55.

Goodman, D. (1987). *Hypercard — The complete handbook*. Bantam Books, New York, USA.

Goodman, P. S., & Garber, S. (1988). Absenteeism and accidents in a dangerous environment: Empirical analysis of underground coal mines. *Journal of Applied Psychology; Journal of Applied Psychology*, *73*(1), 81.

Goodman, P. S., & Leyden, D. P. (1991). Familiarity and group productivity. *Journal of Applied Psychology; Journal of Applied Psychology*, *76*(4), 578.

Google Inc. (2011). *Google Image Labeler*.

Gordon, M. E., & Fitzgibbons, W. J. (1982). Empirical test of the validity of seniority as a factor in staffing decisions. *Journal of Applied Psychology*, *67*(3), 311.

got2know.net. (2012). 2Know Mnemonic Software. Retrieved May 3, 2012, from http://www.got2know.net/

Gould, J. D., & Lewis, C. (1985). Designing for usability: key principles and what designers think. *Communications of the ACM*, *28*(3), 300–311.

Grey, R. (1756). *Memoria Technica: Or, a New Method of Artificial Memory,... By Richard Grey,...* printed for D. Browne, J. Whiston and B. White, and J. Rivington and J. Fletcher.

Grissemann, H. (1981). ZLT, Zürcher Lesetest. *Förderdiagnostik bei gestörtem Schriftspracherwerb*, *5*.

Gross, M. D., & Green, K. E. (2012). Architectural robotics, inevitably. *interactions*, *19*(1), 28–33.

Gross, T. (2003). Ambient interfaces: design challenges and recommendations. *Human computer interaction: theory and practice*, 68–72.

Gruenfeld, D. H., Mannix, E. A., Williams, K. Y., & Neale, M. A. (1996). Group composition and decision making: How member familiarity and information distribution affect process and performance. *Organizational Behavior and Human Decision Processes*, *67*(1), 1–15.

guidebookgallery.org. (2012). GUIdebook > GUIs. Retrieved May 22, 2012, from http://www.guidebookgallery.org/guis

Halasz, F., & Moran, T. P. (1982). Analogy considered harmful. In *Proceedings of the 1982 conference on Human factors in computing systems* (pp. 383–386).

Halfhill, T. R., & Reinhardt, A. (1994). Just like magic?, *Byte*(February), 22–23.

Hampton, J. A. (1979). Polymorphous concepts in semantic memory. *Journal of Verbal Learning and Verbal Behavior*, *18*(4), 441–461.

Hampton, J. A. (1995). Testing the prototype theory of concepts. *Journal of Memory and Language*, *34*(5), 686–708.

Hampton, J. A. (1997a). Emergent attributes in combined concepts. *Creative thought: An investigation of conceptual structures and processes*, 83–110.

Hampton, J. A. (1997b). Psychological representation of concepts. *Cognitive models of memory*, 81–110.

Harrington, M. (2009). Innovation: The lessons of Bob. Retrieved May 29, 2012, from http://www.techflash.com/seattle/2009/08/Innovation_The_lessons_of_Bob_53605837.html

Harslem, E., & Nelson, L. R. E. (1982). A retrospective on the development of Star. In *Proceedings of the 6th international conference on Software engineering* (pp. 377–383).

Heider, E. R. (1971a). “ Focal” color areas and the development of color names. *Developmental Psychology; Developmental Psychology*, *4*(3), 447.

Heider, E. R. (1971b). Universals in color memory and name. In *meeting of the Eastern Psychological Association, New York*.

Heider, E. R. (1972). Universals in color naming and memory. *Journal of Experimental Psychology; Journal of Experimental Psychology*, *93*(1), 10.

Helson, H. (1964). Current trends and issues in adaptation-level theory. *American Psychologist*, *19*(1), 26.

Helvetius, C. A. (1758). De l’etendue de la mémoire. In *De L’esprit* (pp. 354–367). Amsterdam.

Herley, C., van Oorschot, P., & Patrick, A. (2009). Passwords: If We’re So Smart, Why Are We Still Using Them? *Financial Cryptography and Data Security*, 230–237.

Hesse, M. B. (1966). *Models and analogies in science* (Vol. 7). University of Notre Dame Press Notre Dame.

Hickerson, N. (1971). Review of Berlin and Kay (1969). *International Journal of American Linguistics*, *37*, 257–270.

Higbee, K. L. (2001). *Your memory: How it works and how to improve it.* Da Capo Press.

Hinds, P. J., Carley, K. M., Krackhardt, D., & Wholey, D. (2000). Choosing work group members: Balancing similarity, competence, and familiarity. *Organizational Behavior and Human Decision Processes*, *81*(2), 226–251.

Hodgkinson, G., & Crawshaw, C. (1985). Hierarchical task analysis for ergonomics research: An application of the method to the design and evaluation of sound mixing consoles. *Applied Ergonomics*, *16*(4), 289–299.

Honeck, R. P., Kibler, C., & Firment, M. J. (1987). Figurative language and psychological views of categorization: Two ships in the night. *Cognition and symbolic structures: The psychology of metaphoric transformation. Norwood, NJ: Ablex*.

Hunziker, H.-W. (2006). *Im Auge des Lesers: foveale und periphere Wahrnehmung-vom Buchstabieren zur Lesefreude*. Zürich: Transmedia, Stäubli Verlag.

Hyde, T. S., & Jenkins, J. J. (1973). Recall for words as a function of semantic, graphic, and syntactic orienting tasks. *Journal of Verbal Learning and Verbal Behavior*, *12*(5), 471–480.

HypoSurface Corp. (2012). HypoSurface. Retrieved August 29, 2012, from http://hyposurface.org/

Iidaka, T., Anderson, N. D., Kapur, S., Cabez, R., & Craik, F. I. (2000). The effect of divided attention on encoding and retrieval in episodic memory revealed by positron emission tomography. *Journal of Cognitive Neuroscience*, *12*(2), 267–280.

International Standards Organisation. (1999). *ISO 13407: Human-centred Design Processes for Interactive Systems*. ISO, Geneva, Switzerland.

International Standards Organisation. (2008). *ISO 9241-210: Ergonomics of human-system interaction -- Part 210: Human-centred design for interactive systems* (ISO). Geneva, Switzerland.

Ishii, H. (2008). Tangible bits: beyond pixels. In *Proceedings of the 2nd international conference on Tangible and embedded interaction* (pp. xv–xxv). Retrieved from http://dl.acm.org/citation.cfm?id=1347392

Ishii, H., Lakatos, D., Bonanni, L., & Labrune, J. B. (2012). Radical atoms: beyond tangible bits, toward transformable materials. *interactions*, *19*(1), 38–51.

Ishii, H., & Ullmer, B. (1997). Tangible bits: Towards seamless interfaces between people, bits and atoms. In *Proceedings of Human Factors in Computing Systems: CHI 97* (pp. 234–241). Denver (Colorado).

ISO. (2013). ISO - International Organization for Standardization. Retrieved April 22, 2013, from http://www.iso.org/iso/home.html

Jermyn, I., Mayer, A., Monrose, F., Reiter, M. K., & Rubin, A. D. (1999). The design and analysis of graphical passwords. In *Proceedings of the 8th USENIX Security Symposium* (pp. 1–14).

joglab.com. (2012). Acrostic Poem, Mnemonic and Backronym Creation at JogLab. Fast, Free, Fun. Retrieved May 3, 2012, from http://www.joglab.com/

Johnson-Laird, P. N. (1983). *Mental models: Towards a cognitive science of language, inference, and consciousness*. Harvard University Press.

Johnson-Laird, P. N. (2005). Mental models and thought. *The Cambridge handbook of thinking and reasoning*, 185–208.

Juul, J. (2001). The repeatedly lost art of studying games. *Game Studies*, *1*(1).

Kahneman, D., & Tversky, A. (1979). Prospect theory: An analysis of decision under risk. *Econometrica: Journal of the Econometric Society*, 263–291.

Kahneman, D., & Tversky, A. (1984). Choices, values, and frames. *American psychologist*, *39*(4), 341.

Kant, I. (1792). *Vom Gedächtnis; in Anthropologie I.*

Kant, I. (1800). *Anthropologie in pragmatischer hinsicht abgefaßt* (2te verbesserte Auflage.). Königsberg: Friedrich Nicolovius.

Kaplan, A. (1964). The Conduct of Inquiry, Methodology for Behavioral Science. *San Francisco, CA*.

Kästner, C. A. L. (1804). *Mnemonik oder System der Gedächtnißkunst der Alten*. Leipzig.

Kästner, C. A. L. (1805). *Mnemonik oder die Gedächtnisskunst der Alten systematisch bearbeitet*. Paul Gotthelf Kummer.

Kato, H., & Billinghurst, M. (1999). Marker tracking and hmd calibration for a video-based augmented reality conferencing system. In *Augmented Reality, 1999.(IWAR'99) Proceedings. 2nd IEEE and ACM International Workshop on* (pp. 85–94). Retrieved from http://ieeexplore.ieee.org/xpls/abs_all.jsp?arnumber=803809

Katz, J. J. (1972). *Semantic theory*. Harper & Row New York.

Katz, R. (1982). The effects of group longevity on project communication and performance. *Administrative Science Quarterly*, 81–104.

Kay, A. (2005). Squeak Etoys Authoring & Media. *Viewpoints Research Institute*. Retrieved from http://vpri.org/pdf/rn2005002_authoring.pdf

Kay, A. C., & Ram, S. (2003). Dr. Alan Kay on the Meaning of "Object-Oriented Programming." Retrieved August 21, 2012, from http://userpage.fu-berlin.de/~ram/pub/pub_jf47ht81Ht/doc_kay_oop_en

Kay, AlanC. (1996). *The early history of Smalltalk*. ACM.

Kay, P. (1971). Taxonomy and semantic contrast. *Language*, 866–887.

Kay, P., Berlin, B., L. Maffi, Merrifield, W. R., & Cook, R. (2009). *World Color Survey*. Sanford: CSLI Publications.

Kay, P., Berlin, B., Maffi, L., Merrifield, W., & others. (1997). Color naming across languages. *Color categories in thought and language*, *21*, 58.

Kay, P., Berlin, B., Maffi, L., & Merrifield, W. R. (2007). World color survey.

Kay, P., & Regier, T. (2003). Resolving the question of color naming universals. *Proceedings of the National Academy of Sciences*, *100*(15), 9085.

Kay, Paul, & Cook, R. S. (2011). The World Color Survey. Retrieved July 15, 2012, from http://www1.icsi.berkeley.edu/wcs/

Keil, F. C. (1992). *Concepts, kinds, and cognitive development*. the MIT Press.

Kennedy, J. M. (1990). Metaphor–Its Intellectual Basis. *Metaphor and Symbol*, *5*(2), 115.

Kerres, M. (2001). *Multimediale und telemediale Lernumgebungen: Konzeption und Entwicklung*. Oldenbourg Wissenschaftsverlag.

King, M. M. (1991). Rebus passwords. In *Computer Security Applications Conference, 1991. Proceedings., Seventh Annual* (pp. 239–243).

Kirkham, N. Z., Slemmer, J. A., & Johnson, S. P. (2002). Visual statistical learning in infancy: Evidence for a domain general learning mechanism. *Cognition*, *83*(2), B35–B42.

Klein, D. V. (1990). Foiling the cracker: A survey of, and improvements to, password security. In *Proceedings of the 2nd USENIX Security Workshop* (pp. 5–14).

Kruschke, J. K. (2003). Concept learning and categorization: Models. *Encyclopedia of cognitive science*.

LaBerge, D. (1975). Acquisition of automatic processing in perceptual and associative learning. *Attention and performance*, *5*, 50–64.

LaBerge, D. (1976). Perceptual learning and attention. *Handbook of learning and cognitive processes*, *4*, 237–273.

LaBerge, D., & Samuels, S. J. (1974). Toward a theory of automatic information processing in reading. *Cognitive psychology*, *6*(2), 293–323.

Labov, W. (1973). *The boundaries of words and their meanings. New ways of analyzing variation in English*. (C.-J. Bailey & R. Shuy, Eds.). Washington, DC: Georgetown University Press.

Labov, W. (1978). Denotational structure. *Parasession on the Lexicon*, 220–260.

Lakoff. (2012). Index of /lakoff/metaphors. Retrieved May 13, 2012, from http://cogsci.berkeley.edu/lakoff/metaphors/

Lakoff, G. (1987). Women, fire, and dangerous things. *What categories reveal*.

Lakoff, G., & Johnson, M. (1980). *Metaphors we live by* (Vol. 111). Chicago London.

Lakoff, G., & Turner, M. (1989). More than cool reason: A field guide to poetic metaphor. *Chicago/London*.

Langlois, J. H., Kalakanis, L., Rubenstein, A. J., Larson, A., Hallam, M., & Smoot, M. (2000). Maxims or myths of beauty? A meta-analytic and theoretical review. *Psychological bulletin*, *126*(3), 390.

Langlois, J. H., & Roggman, L. A. (1990). Attractive faces are only average. *Psychological science*, *1*(2), 115.

Lansdale, M. W., & Ormerod, T. C. (1994). *Understanding interfaces: a handbook of human-computer dialogue*. Academic Press Professional, Inc.

Larkin, J., McDermott, J., Simon, D. P., & Simon, H. A. (1980). Expert and novice performance in solving physics problems. *Science*, *208*(4450), 1335–1342.

Larsen, Y. C., & Bogner, F. X. (2012). *D52 Final Report of the Validation Phase*.

Larsen, Y. C., Buchholz, H., Brosda, C., & Bogner, F. X. (2012). Evaluation of a portable and interactive augmented reality learning system by teachers and students. In A. Lazoudis, H. Salmi, & S. Sotiriou (Eds.), *Augmented Reality in Education - Proceedings of the "Science Center To Go" Workshops* (pp. 47–56). Athens, Greece: EPINOLA S.A.

Law, E. L. M., Von Ahn, L., Dannenberg, R., & Crawford, M. (2007). Tagatune: A game for music and sound annotation. *Proceedings of ISMIR (Vienna, Austria, 2007)*.

Law, E., West, K., Mandel, M., Bay, M., & Downie, J. S. (2009). Evaluation of algorithms using games: The case of music tagging. In *Proceedings of the 10th International Conference on Music Information Retrieval (ISMIR).(October 2009)* (pp. 387–392).

Lazoudis, A., Salmi, H., Sotiriou, S., Bogner, F. X., Owen, M., Owen, S., … and others. (2012). *Science Center To Go: Guide of Good Practice*. Athens, Greece: EPINOLA S.A.

Leech, G. N. (1969). *A linguistic guide to English poetry*. Longmans.

Lefavrais, P. (1967). *Manuel du test de l'Alouette: test d'analyse de la lecture et de la dyslexie*. Éditions du Centre de psychologie appliquée.

LEGO.com. (2012). MINDSTORMS : Home. Retrieved August 27, 2012, from http://mindstorms.lego.com

Leithinger, D., & Ishii, H. (2010). Relief: a scalable actuated shape display. In *Proceedings of the fourth international conference on Tangible, embedded, and embodied interaction* (pp. 221–222). Retrieved from http://dl.acm.org/citation.cfm?id=1709928

Leithinger, D., Lakatos, D., DeVincenzi, A., Blackshaw, M., & Ishii, H. (2011). Direct and gestural interaction with relief: A 2.5 D shape display. In *Proceedings of the 24th annual ACM symposium on User interface software and technology* (pp. 541–548).

Leithinger, Daniel, Kumpf, A., & Ishii, H. (2009). Tangible Media Group - Project Relief. Retrieved August 29, 2012, from http://tangible.media.mit.edu/project/relief/

Lenneberg, E. H. (1953). Cognition in ethnolinguistics. *Language*, *29*(4), 463–471.

Lenneberg, E. H., & Roberts, J. M. (1991). *The language of experience: A study in methodology*. UMI.

Lerner, R. M. (1998). Theories of human development: Contemporary perspectives.

Lesgold, A. (1988). Problem solving. *The psychology of human thought*, 188–213.

Levin, D. T. (2000). Race as a visual feature: using visual search and perceptual discrimination tasks to understand face categories and the cross-race recognition deficit. *Journal of Experimental Psychology: General*, *129*(4), 559.

Levin, J. R., Levin, M. E., Glasman, L. D., & Nordwall, M. B. (1992). Mnemonic vocabulary instruction: Additional effectiveness evidence. *Contemporary Educational Psychology*, *17*(2), 156–174.

Li, Z., Cheng, K., Wang, Y., Hiltz, S. R., & Turoff, M. (2001). Thurstone's Law of Comparative Judgment for Group Support. In *Proceedings of the Americas Conference on Information Systems* (pp. 2–5). Retrieved from http://www.dss.dpem.tuc.gr/pdf/BB004.pdf

Lieberman, H., Smith, D., & Teeters, A. (2007). Common Consensus: a web-based game for collecting commonsense goals. In *ACM Workshop on Common Sense for Intelligent Interfaces*. Retrieved from http://www.media.mit.edu/~lieber/Publications/Common-Consensus.pdf

Linder, M., & Grissemann, H. (2003). *Zürcher Lesetest: ZLT; Förderdiagnostik bei gestörtem Schriftspracherwerb; Manual*. Huber.

Logan, G.D. (1988). Toward an instance theory of automatization. *Psychological review*, *95*(4), 492.

Logan, Gordon D. (2002). An instance theory of attention and memory. *Psychological Review*, *109*(2), 376–400. doi:10.1037//0033-295X.109.2.376

Logo Foundation. (2000). A Logo Primer. Retrieved August 28, 2012, from http://el.media.mit.edu/logo-foundation/logo/turtle.html

Love, B. C. (2003). Concept learning. *Encyclopedia of cognitive science*. Retrieved from http://onlinelibrary.wiley.com/doi/10.1002/0470018860.s00499/full

LucasArts. (1990). *The Secret of Monkey Island*.

Luce, R. D. (1959). Individual choice behavior. A theoretical analysis.

Luce, R. D. (1994). Thurstone and sensory scaling: Then and now. Retrieved from http://psycnet.apa.org/psycinfo/1994-28134-001

Luce, T. (1974). Blacks, whites and yellows: They all look alike to me. *Psychology Today*, *8*(6), 105–106.

Ma, H., Chandrasekar, R., Quirk, C., & Gupta, A. (2009). Improving search engines using human computation games. In *Proceeding of the 18th ACM conference on Information and knowledge management* (pp. 275–284). Retrieved from http://dl.acm.org/citation.cfm?id=1645990

Maiden, N. A. M., & Rugg, G. (1996). ACRE: selecting methods for requirements acquisition. *Software Engineering Journal*, *11*(3), 183–192.

Malpass, R. (1992). They all look alike to me. *The undaunted psychologist*, 74–88.

Malt, B. C., & Smith, E. E. (1984). Correlated properties in natural categories. *Journal of verbal learning and verbal behavior*, *23*(2), 250–269.

Maslow, A. (1962). *Toward a psychology of being*. Princeton, NJ:: D. Van Nostrana Co Inc.

Maslow, A. H. (1966). *The psychology of science: a reconnaissance*. Harper & Row.

Mayringer, H., & Wimmer, H. (2003). *SLS 1-4: Salzburger Lese-Screening für die Klassenstufen 1-4: Manual*. H. Huber.

Mayringer, H., & Wimmer, H. (2008). *Salzburger Lese-Screening für die Klassenstufen 1-4 (SLS 1-4)*. Hans Huber.

McCracken, H. (2010). The Bob Chronicles - On its fifteenth anniversary, a look back at a legendary software flop. Retrieved May 29, 2012, from http://technologizer.com/2010/03/29/microsoft-bob/

McIver, J. P., & Carmines, E. G. (1981). *Unidimensional scaling* (Vol. 24). Sage Publications, Incorporated.

McNamara, D. S., & Kintsch, W. (1996). Learning from texts: Effects of prior knowledge and text coherence. *Discourse processes*, *22*(3), 247–288.

Meacham, J. W. (2013, August). Alternative unlocking patterns.

Medin, D. L. (1998). Concepts and conceptual structure. *Mind readings*, 93–126.

Medin, D. L., & Heit, E. (1999). Categorization. *Handbook of cognition and perception*, 99–143.

Medin, D. L., Lynch, E. B., & Solomon, K. O. (2000). Are there kinds of concepts? *Annual review of psychology*, *51*(1), 121–147.

Medin, D. L., Proffitt, J. B., & Schwartz, H. C. (2000). Concepts: An overview.

Mercedes-Benz. (2013). Mercedes-Benz international: Sports Fashion TV Design Vehicles History. Retrieved April 21, 2013, from http://www5.mercedes-benz.com/en/

Mervis, C. B., Catlin, J., & Rosch, E. (1976). Relationships among goodness-of-example, category norms, and word frequency. *Bulletin of the Psychonomic Society; Bulletin of the Psychonomic Society*.

Microsoft. (1995). *Microsoft Bob*.

Microsoft. (2012). A history of Windows - Microsoft Windows. Retrieved July 22, 2012, from http://windows.microsoft.com/en-US/windows/history

Microsoft Corporation. (2011). Ribbon Hero 2. Retrieved August 24, 2012, from http://www.ribbonhero.com/

Microsoft Corporation. (2012). Microsoft Office entdecken, testen und kaufen - Office.com. Retrieved August 24, 2012, from http://office.microsoft.com/de-de/

Miller, G. A. (1956). The magical number seven, plus or minus two: some limits on our capacity for processing information. *Psychological review*, *63*(2), 81.

MIT Media Lab. (2012). Tangible Media Group. Retrieved August 29, 2012, from http://tangible.media.mit.edu/

mongodb.org. (2013). MongoDB. Retrieved April 26, 2013, from http://www.mongodb.org/

Montessori, M. (1946). *Education for a new world*. Kalakshetra Madras.

Moroney, N., Fairchild, M., Hunt, R., Li, C., Luo, M. R., & Newman, T. (2002). The CIECAM02 color appearance model.

Morris, C. D., Bransford, J. D., & Franks, J. J. (1977). Levels of processing versus transfer appropriate processing*. *Journal of verbal learning and verbal behavior*, *16*(5), 519–533.

Morris, R., & Thompson, K. (1979). Password security: A case history. *Communications of the ACM*, *22*(11), 594–597.

Murphy, G. L. (1993). Theories and concept formation.

Mussen, P. H., Flavell, J. H., Carmichael, L., & Markman, E. M. (1983). *Handbook of Child Psychology: Cognitive development* (Vol. 3). John Wiley & Sons.

Nacke, L. (2005). Facilitating the education of game development. *Master's thesis, University of Magdeburg*.

NASA Science. (2001). Unmasking the Face on Mars. Retrieved April 21, 2013, from http://science.nasa.gov/science-news/science-at-nasa/2001/ast24may_1/

Neely, J. H. (2003). Priming. *Encyclopedia of cognitive science*.

Neumann, P. G. (1977). Visual prototype formation with discontinuous representation of dimensions of variability. *Memory & Cognition*, *5*(2), 187–197.

Nieborg, D. (2004). *America's Army: More than a game*.

Nielsen, J. (1992). The usability engineering life cycle. *Computer*, *25*(3), 12–22.

Nielsen, J. (1993). Noncommand user interfaces. *Communications of the ACM*, *36*(4), 83–99.

Nielsen, Jakob. (1989). Usability engineering at a discount. In *Proceedings of the third international conference on human-computer interaction on Designing and using human-computer interfaces and knowledge based systems (2nd ed.)* (pp. 394–401). New York, NY, USA: Elsevier Science Inc.

Nielsen, Jakob. (1994a). *Usability engineering*. Morgan Kaufmann.

Nielsen, Jakob. (1994b). Usability inspection methods. In *Conference companion on Human factors in computing systems* (pp. 413–414).

Nielsen, Jakob, Mack, R. L., & Shirk, H. N. (1996). Usability inspection methods. *Journal of Technical Writing and Communication*, *26*(1), 97.

Norman, D. (1998). *The invisible computer: Why good products can fail, the personal computer is so complex, and information appliances are the answer*. Cambridge, MA: The MIT Press.

Norman, D. A. (1988). The Psychology of Everyday Things. *Psychopathology*.

Norman, D. A. (1999). Affordance, conventions, and design. *interactions*, *6*(3), 38–43.

Norwich, K. H. (1983). To perceive is to doubt: the relativity of perception. *Journal of theoretical biology*, *102*(2), 175–190.

O'Malley, C., & Stanton Fraser, D. (2004). Literature review in learning with tangible technologies, (report 12).

Oberauer, K., Mayr, U., & Kluwe, R. H. (2005). Gedächtnis und Wissen. In H. Spada (Ed.), *Allgemeine Psychologie* (pp. 115–196). Bern: Huber.

Olson, J. S., & Moran, T. P. (1995). *Mapping the Method Muddle: Guidance in Using Methods for User Interface Design.* DTIC Document. Retrieved from http://oai.dtic.mil/oai/oai?verb=getRecord&metadataPrefix=html&identifier=ADA328582

Oracle Corporation. (2013). MySQL - The world's most popular open source database. Retrieved April 26, 2013, from http://www.mysql.com/

Ormerod, T. C., & Shepherd, A. (1998). Using Task Analysis for Information Requirements Specification: The Sub-Goal Template (SGT) Method. *The handbook of task analysis for human-computer interaction*, 347–366.

Ortony, A. (1979). Beyond literal similarity. *Psychological review*, *86*(3), 161.

Ortony, A., Vondruska, R. J., Foss, M. A., & Jones, L. E. (1985). Salience, similes, and the asymmetry of similarity. *Journal of memory and language*, *24*(5), 569–594.

Packard Bell. (1995). *Packard Bell Navigator 3.5*.

Paivio, A., Rogers, T., & Smythe, P. C. (1968). Why are pictures easier to recall than words? *Psychonomic Science*.

Palmer, S. (1999). Gestalt perception. *The MIT encyclopedia of the cognitive sciences*, 344–346.

Palmer, S. E. (1999). *Vision science: Photons to phenomenology* (Vol. 1). MIT press Cambridge, MA.

Palmer, S. E. (2000). Perceptual organization.

Palmeri, T. J. (1999). Theories of automaticity and the power law of practice.

Papert, S. (1980). *Mindstorms: Children, computers, and powerful ideas*. Basic Books, Inc.

Park, S. K., & Miller, K. W. (1988). Random number generators: good ones are hard to find. *Communications of the ACM*, *31*(10), 1192–1201.

Parkes, A., LeClerc, V., & Ishii, H. (2006). Glume: exploring materiality in a soft augmented modular modeling system. In *CHI'06 extended abstracts on Human factors in computing systems* (pp. 1211–1216). Retrieved from http://dl.acm.org/citation.cfm?id=1125678

Parlament Republik Österreich. (2006). Roth-Halvax: Jugendliche sollen sich politisch stärker einbringen - Bundesratspräsidentin mit Schulprojekt gegen Politikverdrossenheit. Retrieved June 21, 2012, from http://www.parlament.gv.at/PAKT/PR/JAHR_2006/PK0392/index.shtml

passfaces.com. (2012). Two Factor Authentication, Graphical Passwords - Passfaces. Retrieved May 11, 2012, from http://www.passfaces.com/

Penn, D. C., Holyoak, K. J., & Povinelli, D. J. (2008). Darwin's mistake: Explaining the discontinuity between human and nonhuman minds. *Behavioral and Brain Sciences*, *31*(02). doi:10.1017/S0140525X08003543

Persuasive Technology Lab. (2010). WHAT IS CAPTOLOGY? – Persuasive Tech. Retrieved September 27, 2012, from http://captology.stanford.edu/about/what-is-capatology.html

Peterson, M. A., Kihlstrom, J. F., Rose, P. M., & Glisky, M. L. (1992). Mental images can be ambiguous: Reconstruals and reference-frame reversals. *Memory & cognition*, *20*(2), 107–123.

pharo-project.org. (2012). Pharo Open Source Smalltalk. Retrieved July 22, 2012, from http://www.pharo-project.org/home

Piaget, J. (1954). *The construction of reality in the child*. New York: Basic Books.

Piaget, J. (1962). *Play, dreams and imitation* (Vol. 24). New York: Norton.

Piaget, J. (2007). *The child's conception of the world*. Rowman & Littlefield Pub Inc.

Piaget, J., & Mays, W. (1972). *The principles of genetic epistemology*. Basic Books New York.

Piaget, J., Wedgwood, S., & Blanchet, A. (1976). *The grasp of consciousness: Action and concept in the young child*. Harvard University Press Cambridge, MA.

Piso, E. (1981). Task analysis for process-control tasks: The method of Annett et al. applied. *Journal of occupational psychology*, *54*(4), 247–254.

Posner, M. I., Goldsmith, R., & Welton Jr, K. E. (1967). Perceived distance and the classification of distorted patterns. *Journal of Experimental Psychology*, *73*(1), 28.

Posner, M. I., & Keele, S. W. (1967). Decay of visual information from a single letter. *Science*, *158*(3797), 137–139.

Posner, M. I., & Snyder, C. R. R. (2004). Attention and Cognitive Control. *Cognitive psychology: Key readings*, 205.

Poupyrev, I., Tan, D., Billinghurst, M., Kato, H., Regenbrecht, H., & Tetsutani, N. (2001). Tiles: A mixed reality authoring interface. In *INTERACT 2001 Conference on Human Computer Interaction* (pp. 334–341). Retrieved from http://www.ivanpoupyrev.com/e-library/2001/Interact2001_tiles.pdf

Poupyrev, I., Tan, D. S., Billinghurst, M., Kato, H., Regenbrecht, H., & Tetsutani, N. (2002). Developing a generic augmented-reality interface. *Computer*, *35*(3), 44–50.

Powerofpolitics.com. (2012). Power of Politics. Retrieved June 21, 2012, from http://www.powerofpolitics.com

Powerofresearch.eu. (2012). Power of Research - the free science online game. Retrieved June 22, 2012, from http://www.powerofresearch.eu/About.aspx?lan=en&red=true

Poynor, R. (1995). The Hand That Rocks the Cradle: Gillian Crampton Smith is making the Royal College of Art's Computer Related Design program a multimedia powerhouse. *ID-NEW YORK-DESIGN PUBLICATIONS-*, *42*, 60–65.

Preim, B. (1999). *Entwicklung interaktiver Systeme*. Springer Berlin, Heidelberg, New York.

Prelinger, E. (1959). Extension and structure of the self. *The Journal of Psychology*, *47*(1), 13–23.

Preuster, T. (2010a). 5 Jahre Power of Politics - Ein Rückblick und ein Ausblick. Retrieved June 22, 2012, from http://blog.tpmgames.com/post/5-Jahre-Power-of-Politics-Ein-Ruckblick-und-ein-Ausblick.aspx

Preuster, T. (2010b). Power of Research : Beta offiziell released. Retrieved June 21, 2012, from http://blog.tpmgames.com/post/Power-of-Research-Beta-offiziell-released.aspx

Prinz, W. (99). NESSIE: An Awareness Environment for Cooperative Settings. In S. Bødker, M. Kyng, & K. Schmidt (Eds.), *Proceedings of The Sixth European Conference on Computer Supported Cooperative Work - ECSCW'99* (pp. 391–410). Copenhagen, Denmark: Kluwer Academic Publishers. Retrieved from http://www.springerlink.com/index/n14vh152q2002p10.pdf

Quintilian. (2006). *Institutes of Oratory.* (J. S. Watson, Trans., L. Honeycutt, Ed.). Iowa State. Retrieved from http://rhetoric.eserver.org/quintilian/

Rachna Dhamija, & Adrian Perrig. (2000). D´ej`a Vu: A User Study Using Images for Authentication. In *13th USENIX Security Symposium.*

Raffle, H. S., Parkes, A. J., & Ishii, H. (2004). Topobo: a constructive assembly system with kinetic memory. In *Proceedings of the SIGCHI conference on Human factors in computing systems* (pp. 647–654). Retrieved from http://dl.acm.org/citation.cfm?id=985774

Rasch, G. (1960). *Probabilistic models for some intelligence and attainnment tests* (Vol. 1). Danmarks paedagogiske Institut.

Ratti, C., Wang, Y., Piper, B., Ishii, H., & Biderman, A. (2004). PHOXEL-SPACE: an interface for exploring volumetric data with physical voxels. In *Proceedings of the 5th conference on Designing interactive systems: processes, practices, methods, and techniques* (pp. 289–296). Retrieved from http://dl.acm.org/citation.cfm?id=1013156

Raugh, M. R., & Atkinson, R. C. (1974). *A mnemonic method for the acquisition of a second-language vocabulary.* Stanford Univ. Inst. for Math. Studies in the Social Sciences.

Raugh, M. R., & Atkinson, R. C. (1975). A Mnemonic Method for Learning a Second-Language Vocabulary. *Journal of Educational Psychology, 67*(1), 1.

Rauterberg, M., Spinas, P., Strohm, O., Ulich, E., & Waeber, D. (1994). Benutzerorientierte Software-Entwicklung–Konzepte. *Methoden und Vorgehen zur Benutzerbeteiligung, Zürich/Stuttgart.*

Reed, S. K. (1972). Pattern recognition and categorization. *Cognitive psychology, 3*(3), 382–407.

Reeves, B., & Read, J. L. (2009). *Total engagement: using games and virtual worlds to change the way people work and businesses compete*. Harvard Business School Press.

Regier, T., Kay, P., & Cook, R. S. (2005). Focal colors are universal after all. *Proceedings of the National Academy of Sciences of the United States of America*, *102*(23), 8386.

Reitman, J. S. (1976). Skilled perception in Go: Deducing memory structures from inter-response times. *Cognitive psychology*, *8*(3), 336–356.

Resnick, M. (2004). Edutainment? No thanks. I prefer playful learning. *Associazione Civita Report on Edutainment*, *14*. Retrieved from http://www.roboludens.net/Edut_Articoli/Playful_Learning.pdf

Resnick, M., Maloney, J., Monroy-Hernández, A., Rusk, N., Eastmond, E., Brennan, K., … others. (2009). Scratch: programming for all. *Communications of the ACM*, *52*(11), 60–67.

Resnick, M., Martin, F., Berg, R., Borovoy, R., Colella, V., Kramer, K., & Silverman, B. (1998). Digital manipulatives: new toys to think with. In *Proceedings of the SIGCHI conference on Human factors in computing systems* (pp. 281–287). Retrieved from http://dl.acm.org/citation.cfm?id=274644.274684

Resnick, Mitchel, & Silverman, B. (2003). Going in Circles. Retrieved August 27, 2012, from http://llk.media.mit.edu/projects/circles/circles.html

Rhodes, G. (2006). The evolutionary psychology of facial beauty. *Annu. Rev. Psychol.*, *57*, 199–226.

Richard Cook, Paul Kay, & Terry Regier. (2012). World Color Survey (WCS) Data Archives. *WCS Data Archives*. Retrieved July 15, 2012, from http://www1.icsi.berkeley.edu/wcs/data.html

Rocard, M., Csermely, P., Jorde, D., Lenzen, D., & Wallberg-Henriksson, H. (2007). *Science Education Now: A renewed Pedagogy for the Future of Europe*. Belgium: European Commision. Retrieved from http://ec.europa.eu/research/science-society/

Rogers, Y., Sharp, H., & Preece, J. (2011). *Interaction design: beyond human-computer interaction*. Wiley Publishing.

Rosch, E. (1975a). Cognitive reference points. *Cognitive psychology*, *7*(4), 532–547.

Rosch, E. (1975b). Cognitive representations of semantic categories. *Journal of Experimental Psychology: General; Journal of Experimental Psychology: General*, *104*(3), 192.

Rosch, E. (1975c). The nature of mental codes for color categories. *Journal of Experimental Psychology: Human Perception and Performance*, *1*(4), 303.

Rosch, E. (1978). Principles of categorization. *Fuzzy grammar: a reader*, 91–108.

Rosch, E. H. (1973a). Natural categories. *Cognitive psychology*, *4*(3), 328–350.

Rosch, E. H. (1973b). On the internal structure of perceptual and semantic categories.

Rosch, E., & Mervis, C. B. (1975). Family resemblances: Studies in the internal structure of categories. *Cognitive psychology*, *7*(4), 573–605.

Rosch, E., Mervis, C. B., Gray, W. D., Johnson, D. M., & Boyes-Braem, P. (1976). Basic objects in natural categories. *Cognitive psychology*, *8*(3), 382–439.

Rosch, E., & others. (1977). Human categorization. *Studies in cross-cultural psychology*, *1*, 1–49.

Rose, D. (2008). Daniel's Legacy Computer Collections - Microsoft® Bob™ Exhibit - Bob Home. Retrieved May 30, 2012, from http://www.danielsays.com/ms-bob-06-bob-home.html

Ross, B. H. (2000). Concepts: Learning, (2), 248–251.

Ross, B. H., & Spalding, T. L. (1994). *Concepts and categories*. Academic Press.

Rossaint, R., Werner, C., & Zwissler, B. (2008). Allgemeine und spezielle Anästhesiologie, Schmerztherapie und Intensivmedizin. *Die Anästhesiologie*.

Royce, W. W. (1970). Managing the development of large software systems. In *proceedings of IEEE WESCON* (Vol. 26). Retrieved from http://leadinganswers.typepad.com/leading_answers/files/original_waterfall_paper_winston_royce.pdf

Rudman, C., & Engelbeck, G. (1995). Lessons in choosing methods for designing complex graphical user interfaces. In *Proceedings of a workshop on Human-computer interface design: success stories, emerging methods, and real-world context: success stories, emerging methods, and real-world context* (pp. 198–228).

Russell, B. C., Torralba, A., Murphy, K. P., & Freeman, W. T. (2008). LabelMe: a database and web-based tool for image annotation. *International journal of computer vision*, *77*(1), 157–173.

Ryan, T. (1999). *The Anatomy of a Design Document, Part 1: Documentation Guidelines for the Game Concept and Proposal*. Gamasutra.

Salmi, H., Sotiriou, S., & Bogner, F. (2009). Visualising the invisible in science centres and science museums: Augmented Reality (AR) technology application and science teaching. *Web-Based Learning Solutions for Communities of Practice: Virtual Environments for Social and Pedagogical Advancement*, 185–210.

Samuels, S. J., LaBerge, D., & Bremer, C. D. (1978). Units of word recognition: Evidence for developmental changes. *Journal of Verbal Learning and Verbal Behavior*, *17*(6), 715–720.

Saunders, B. A. C., & Van Brakel, J. (1997). Colour: An exosomatic organ? *Behavioral and Brain Sciences*, *20*(02), 212–220.

SCeTGo Consortium. (2012). SCeTGo: Science Center To Go – A mixed reality learning environment of miniature exhibits. project website. Retrieved September 27, 2012, from http://www.sctg.eu/

Schmitz, J., Lehrl, S., Schröder, U., & Wagner, G. (2003). Einfluss von Dehydratation auf die kognitive Leistungsfähigkeit im Rahmen der Rosbacher Trinkstudie (RTS). Presented at the 40. Wissenschaftlicher Kongress der Deutschen Gesellschaft für Ernährung eV in Potsdam, Potsdam.

Schneider, W., & Shiffrin, R. M. (1977). Controlled and automatic human information processing: I. Detection, search, and attention. *Psychological review*, *84*(1), 1.

Schulmeister, R. (2007). *Grundlagen hypermedialer Lernsysteme: Theorie-Didaktik-Design*. Oldenbourg Wissenschaftsverlag.

Seaside.st. (2012). seaside - The framework for developing sophisticated web applications in Smalltalk. Retrieved September 23, 2012, from http://seaside.st/

Secord, P. F. (1968). Consistency theory and self-referent behavior. *Theories of cognitive consistency: A sourcebook*, 349–354.

Seriosity Inc. (2010). Seriosity Attent: Email Management Software Solution for the Enterprise. Retrieved August 24, 2012, from http://www.seriosity.com/attent.html

Seybold, J. (1981). Xerox's' Star'. *The Seybold Report*, *10*, 16.

Shaer, O. (2009). Tangible User Interfaces: Past, Present, and Future Directions. *Foundations and Trends® in Human–Computer Interaction*, *3*(1-2), 1–137. doi:10.1561/1100000026

Shane, H. G. (1981). Significant writings that have influenced the curriculum: 1906-81. *The Phi Delta Kappan*, *62*(5), 311–314.

Shewhart, W. A. (1939). *Statistical method from the viewpoint of quality control*. Dover Publications.

Shneiderman, B., & Plaisant, C. (2010). *Designing the User Interface: Strategies for Effective Human-Computer Interaction (Fifth Edition)*. Boston, MA: Addison Wesley.

Sierra Entertainment. (2006). *Leisure Suit Larry Collection*.

Smgww.org. (2012). The Stock Market Game™ - Home. Retrieved June 22, 2012, from http://www.smgww.org/

Smith, D. C., Irby, C., Kimball, R., Berplank, B., & Harslem, E. (1990). Designing the Star user interface. In *Human-computer interaction* (pp. 237–259).

Smith, D. C., Irby, C., Kimball, R., & Harslem, E. (1982). The Star user interface: An overview. *Proceedings of the June*, 7–10.

Smith, E. E., & Medin, D. L. (1981). *Categories and concepts*. Harvard University Press Cambridge, MA.

Smith, E. E., Osherson, D. N., Rips, L. J., & Keane, M. (1988). Combining prototypes: A selective modification model. *Cognitive Science*, *12*(4), 485–527.

Smith, E. E., Shoben, E. J., & Rips, L. J. (1974). Structure and process in semantic memory: A featural model for semantic decisions. *Psychological review*, *81*(3), 214.

Smith, J. D. (2005). Wanted: A New Psychology of Exemplars. *Canadian Journal of Experimental Psychology/Revue canadienne de psychologie expÃ\copyright rimentale*, *59*(1), 47.

Solso, R. L. (2005). *Cognitive Psychology* (Vol. 7th. Edition). New York: Pearson Education, inc.

Solso, R. L., & Biersdorff, K. K. (1975). Recall under conditions of cumulative cues. *Journal of General Psychology; Journal of General Psychology*.

Solso, R. L., & McCarthy, J. E. (1981). Prototype formation of faces: A case of pseudo-memory. *British Journal of Psychology*, *72*(4), 499–503.

Sommerville, I. (2001). *Software Engineering (6. Aufl.).* Pearson Studium.

Sotiriou, S., Anastopoulou, S., Rosenfeld, S., Aharoni, O., Hofstein, A., Bogner, F., … Hoeksema, K. (2006). Visualizing the invisible: the CONNECT approach for teaching science. In *Advanced Learning Technologies, 2006. Sixth International Conference on* (pp. 1084–1086).

Spence, J. D. (1984). *The memory palace of Matteo Ricci.* New York, NY, USA: Viking Penguin.

Spiegel.de. (2008). Surftipp: Kostenlos-Erweiterung macht das Web zum Rollenspiel. Retrieved June 10, 2012, from http://www.spiegel.de/netzwelt/spielzeug/surftipp-kostenlos-erweiterung-macht-das-web-zum-rollenspiel-a-550341.html

Stanton, N. A., & Young, M. S. (1999). What price ergonomics? *Nature*, *399*, 197–198.

Stanton, Neville A. (2006). Hierarchical task analysis: Developments, applications, and extensions. *Applied Ergonomics*, *37*(1), 55–79. doi:10.1016/j.apergo.2005.06.003

Starbatty, J. (2009). *MAKRO - Benutzerhandbuch für das makroökonomische Planspiel.* Tübingen.

Sternberg, R. J. (2008). *Cognitive psychology* (5th ed.). Wadsworth Pub Co.

Stockwatch.com.au. (2012). StockWatch Share Trading Game. Play Now! Retrieved June 22, 2012, from http://www.stockwatch.com.au/

Stroop, J. R. (1935). Studies of interference in serial verbal reactions, JEP 18 task. *Can. J. Psych*, *36*, 684–700.

Stroop, J. R. (1992). Studies of interference in serial verbal reactions. *Journal of Experimental Psychology: General*, *121*(1), 15.

Strube, G., Becker, B., Freksa, C., Hahn, U., Opwis, K., & Palm, G. (1996). *Wörterbuch der Kognitionswissenschaft.* Klett-Cotta.

sugar labs. (2010). Sugar Labs—learning software for children. Retrieved September 10, 2012, from http://www.sugarlabs.org/

Sutherland, I. E. (1964). Sketch pad a man-machine graphical communication system. In *Proceedings of the SHARE design automation workshop* (pp. 6–329).

Tanaka, J. W., & Taylor, M. (1991). Object categories and expertise: Is the basic level in the eye of the beholder? *Cognitive psychology*, *23*(3), 457–482.

Taylor, A. (2000). IT projects: sink or swim. *The computer bulletin*, *42*(1), 24–26.

Taylor, Stanford E. (2006). Fluency in silent reading. *Retrieved January*, *15*, 2007.

Taylor, Stanford Earl, Frackenpohl, H., & Pettee, J. L. (1960). *Grade level norms for the components of the fundamental reading skill*. Educational Developmental Laboratories.

The Behavior Wizard. (2012). The Behavior Wizard. Retrieved September 27, 2012, from http://www.behaviorwizard.org/wp/

thenethernet.com. (2012). Welcome to The Nethernet. Retrieved June 10, 2012, from http://thenethernet.com/

Thompson-Schill, S. L., D'Esposito, M., Aguirre, G. K., & Farah, M. J. (1997). Role of left inferior prefrontal cortex in retrieval of semantic knowledge: A reevaluation. *Proceedings of the National Academy of Sciences*, *94*(26), 14792–14797.

Thurstone, L. L. (1927). A law of comparative judgment. *Psychological Review; Psychological Review*, *34*(4), 273.

Thurstone, L. L. (1945). The prediction of choice. *Psychometrika*, *10*(4), 237–253.

Toastytech.com. (2012a). Microsoft Bob. Retrieved May 30, 2012, from http://toastytech.com/guis/bob.html

Toastytech.com. (2012b). Packard Bell Navigator 3.5. Retrieved May 30, 2012, from http://toastytech.com/guis/pbnav35.html

TopWindowsTutorials. (2009). *Microsoft BOB - A tour of Microsoft's forgotten desktop "enhancement."* Retrieved from http://www.youtube.com/watch?v=5teG6ou8mWU

Toyota. (2013). Toyota Cars, Trucks, SUVs & Hybrids | Toyota Official Site. Retrieved April 21, 2013, from http://www.toyota.com/

Turkle, S., & Papert, S. (1990). Epistemological pluralism: Styles and voices within the computer culture. *Signs*, *16*(1), 128–157.

Tversky, A. (1977). Features of similarity. *Psychological review*, *84*(4), 327.

Tversky, A., & Kahneman, D. (1981). The framing of decisions and the psychology of choice. *Science*, *211*(4481), 453–458.

Twitter. (2013). Bootstrap. Retrieved April 26, 2013, from http://twitter.github.io/bootstrap/

Ullmer, B., & Ishii, H. (1997). The metaDESK: models and prototypes for tangible user interfaces. In *Proceedings of the 10th annual ACM symposium on User interface software and technology* (pp. 223–232).

Ullmer, B., & Ishii, H. (2000). Emerging frameworks for tangible user interfaces. *IBM systems journal*, *39*(3.4), 915–931.

Ungerer, F., & Schmid, H. J. (2006). *An introduction to cognitive linguistics*. Pearson Education.

Unilever. (2013). Unilever global company website | Unilever Global. Retrieved April 21, 2013, from http://www.unilever.com/

US Army. (2012). America's Army Official Website. Retrieved June 9, 2012, from http://www.americasarmy.com/

Valentine, T., & Endo, M. (1992). Towards an exemplar model of face processing: The effects of race and distinctiveness. *The Quarterly Journal of Experimental Psychology*, *44*(4), 671–703.

Van Bever, A., & others. (1923). *Maximes et pensées de Chamfort; suivies de Dialogues philosophiques; text revu sur l'édition originale et publié avec des notes et un index par Ad. van Bever*. G. Crès.

Vicente, K. J., & De Groot, A. D. (1990). The memory recall paradigm: Straightening out the historical record. *American Psychologist*, *45*(2), 285.

Vincent, A., Craik, F. I., & Furedy, J. J. (1996). Relations among memory performance, mental workload and cardiovascular responses. *International Journal of Psychophysiology*, *23*(3), 181–198.

Voigt, U. (2001). *Esels Welt: Mnemotechnik zwischen Simonides und Harry Lorayne*. Likanas.

Volkswagen, & thefuntheory.com. (2009a). Bottle Bank Arcade Machine | The Fun Theory. Retrieved August 24, 2012, from http://www.thefuntheory.com/bottle-bank-arcade-machine

Volkswagen, & thefuntheory.com. (2009b). Piano Staircase | The Fun Theory. Retrieved August 24, 2012, from http://www.thefuntheory.com/piano-staircase

Volkswagen, & thefuntheory.com. (2009c). The World's Deepest Bin | The Fun Theory. Retrieved August 24, 2012, from http://www.thefuntheory.com/worlds-deepest-bin

Volkswagen, & thefuntheory.com. (2010). The Speed Camera Lottery | The Fun Theory. Retrieved August 24, 2012, from http://www.thefuntheory.com/speed-camera-lottery-0

Von Ahn, L. (2006). Games with a purpose. *Computer*, *39*(6), 92–94.

Von Ahn, L., & Dabbish, L. (2004). Labeling images with a computer game. In *Proceedings of the SIGCHI conference on Human factors in computing systems* (pp. 319–326).

Von Ahn, L., & Dabbish, L. (2008). Designing games with a purpose. *Communications of the ACM*, *51*(8), 58–67.

Von Ahn, L., Ginosar, S., Kedia, M., & Blum, M. (2007). Improving image search with phetch. In *Acoustics, Speech and Signal Processing, 2007. ICASSP 2007. IEEE International Conference on* (Vol. 4, pp. IV–1209).

Von Ahn, L., Kedia, M., & Blum, M. (2006). Verbosity: a game for collecting common-sense facts. In *Proceedings of the SIGCHI conference on Human Factors in computing systems* (pp. 75–78).

Von Ahn, L., Liu, R., & Blum, M. (2006). Peekaboom: a game for locating objects in images. In *Proceedings of the SIGCHI conference on Human Factors in computing systems* (pp. 55–64).

Vygotsky, L. S. (1964). Thought and language. *Annals of Dyslexia*, *14*(1), 97–98.

Wagner, D., & Schmalstieg, D. (2007). Artoolkitplus for pose tracking on mobile devices. In *Proceedings of 12th Computer Vision Winter Workshop (CVWW'07)* (pp. 139–146). Retrieved from http://clairethesis.googlecode.com/svn-history/r53/trunk/relatedWork/ARToolKitPlusMobilePoseTracking.pdf

Walker, P. M., Tanaka, J. W., & others. (2003). An encoding advantage for own-race versus other-race faces. *PERCEPTION-LONDON-*, *32*(9), 1117–1126.

Wallstreetsurvivor.com. (2012). Wall Street Survivor uses fantasy stock market stock game to teach how to invest and how to buy stocks. Retrieved June 22, 2012, from http://www.wallstreetsurvivor.com/

Warren, C. (2003). LucasArts and the Design of Successful Adventure Games: The True Secret of Monkey Island.

Wattenmaker, W. D. (1995). Knowledge structures and linear separability: Integrating information in object and social categorization. *Cognitive Psychology*, *28*(3), 274–328.

Watts, A. F. (1944). The language and mental development of children.

Weidenmann, B. (2001). Lernen mit Medien. In A. Krapp & Weidenmann (Eds.), *Pädagogische Psychologie: Ein Lehrbuch* (4. vollständig überarbeitete Auflage., pp. 417–466). Weinheim: Beltz, PVU.

Weinshall, D. (2006). Cognitive authentication schemes safe against spyware. In *Security and Privacy, 2006 IEEE Symposium on* (p. 6–pp).

Weiser, M. (1991). The computer for the 21st century. *Scientific American*, *265*(3), 94–104.

Wertheimer, M. (1938). Numbers and numerical concepts in primitive peoples.

Whiteside, J., Bennett, J., & Holtzblatt, K. (1988). Usability engineering: Our experience and evolution. *Handbook of Human–Computer Interaction, edited by M. Helander, Amsterdam: North Holland*.

Whorf, B. L., & Carroll, J. B. (1956). *Language, thought, and reality: Selected writings* (Vol. 5). the MIT Press.

Wiedenbeck, S., Waters, J., Birget, J. C., Brodskiy, A., & Memon, N. (2005a). Authentication using graphical passwords: effects of tolerance and image choice. In *Proceedings of the 2005 symposium on Usable privacy and security* (pp. 1–12).

Wiedenbeck, S., Waters, J., Birget, J. C., Brodskiy, A., & Memon, N. (2005b). PassPoints: Design and longitudinal evaluation of a graphical password system. *International Journal of Human-Computer Studies*, *63*(1), 102–127.

Wikipedia.org. (2012a). Dashboard. Retrieved April 30, 2012, from http://en.wikipedia.org/wiki/Dashboard

Wikipedia.org. (2012b). Keyboard layout. Retrieved April 30, 2012, from http://en.wikipedia.org/wiki/Keyboard_layout

Wikipedia.org. (2012c). Mac OS X – Wikipedia. Retrieved July 22, 2012, from http://de.wikipedia.org/wiki/Mac_OS_X

Wikipedia.org. (2012d). Microsoft Windows - Wikipedia, the free encyclopedia. Retrieved July 22, 2012, from http://en.wikipedia.org/wiki/Microsoft_Windows

Wikipedia.org. (2012e). Minivan - Wikipedia, the free encyclopedia. Retrieved August 8, 2012, from http://en.wikipedia.org/wiki/Minivan

Wikipedia.org. (2012f). Mnemonic. Retrieved April 25, 2012, from http://en.wikipedia.org/wiki/Mnemonic

Wikipedia.org. (2013a). Ambiguous image. Retrieved April 21, 2013, from http://en.wikipedia.org/wiki/Ambiguous_image

Wikipedia.org. (2013b). Intuition (psychology). Retrieved September 25, 2013, from http://en.wikipedia.org/wiki/Intuition_(psychology)#cite_note-1

Wikipedia.org. (2013c). Johannes Gutenberg – Wikipedia. Retrieved April 22, 2013, from http://de.wikipedia.org/wiki/Johannes_Gutenberg

Wikipedia.org. (2013d). Taboo (game) - Wikipedia, the free encyclopedia. Retrieved May 29, 2013, from http://en.wikipedia.org/wiki/Taboo_(game)

Wikipedia.org. (2013e). Technical standard - Wikipedia, the free encyclopedia. Retrieved June 9, 2013, from http://en.wikipedia.org/wiki/Technical_standard

Wikiquote.org. (2012). Nicolas Chamfort - Wikiquote. Retrieved May 13, 2012, from http://en.wikiquote.org/wiki/Nicolas_Chamfort

Wisniewski, E. J. (1997). When concepts combine. *Psychonomic Bulletin & Review*, *4*(2), 167–183.

Wisniewski, E. J., & Medin, D. L. (1994). On the interaction of theory and data in concept learning. *Cognitive Science*, *18*(2), 221–281.

Wittgenstein, L. (1953). Philosophical investigations. *GEM Anscombe*, *261*, 49.

Wittkämper, M., Braun, A. K., Herbst, I., & Herling, J. (2007). A distributed system for augmented reality experiences in science centers and museums. *Technologies for E-Learning and Digital Entertainment*, 936–946.

Woodbury, A. (1991). Eskimo Words for "Snow." *Linguist List*, 5–1239.

World Memory Sports Council. (2011). Current Memory Records 2011 v4. Retrieved from http://www.worldmemorychampionships.com/TDE_CMS/database/userfiles//Current%20Memory%20Records%202011%20v4.pdf

Wu, T. (1999). A real-world analysis of Kerberos password security. In *Proceedings of the 1999 ISOC Network and Distributed System Security Symposium, http://www. isoc. org/isoc/conferences/ndss/99/proceedings/papers/wu. pdf.*

www.sueddeutsche.de. (2005). Clemens Mayer: Gedächtnis-Weltmeister und Abiturient. Retrieved April 28, 2012, from http://www.sueddeutsche.de/panorama/profil-clemens-mayer-gedaechtnis-weltmeister-und-abiturientpp-1.851018

Wyszecki, G., & Stiles, W. S. (1967). *Color science*. Wiley New York.

Yates, F. A. (1966). *The art of memory*. Chicago: University of Chicago Press.

Zajonc, R. B. (1968). Attitudinal effects of mere exposure. *Journal of Personality and Social Psychology*, *9*(2p2), 1.

Related Publications

The following publications are directly related to Digital Transformatives.

Buchholz, H., & Prinz, W. (2011). Der Digitale Untersetzer-Änderung des Trinkverhaltens durch Reflexion. I-com Zeitschrift Für Interaktive Und Kooperative Medien 10, 10(2), 16–20.

Buchholz, Hagen. (2005). Konzept einer intelligenten Mikrowelt:Eine explorative adaptive Lernumgebung zum konstruktivistischen spielerischen Lernen durch wissensbasierte Simulation. Diplomarbeit, Otto-Von-Guericke Universität Magdeburg, Magdeburg.

Buchholz, Hagen, & Brosda, C. (2012). Melting Interfaces – Learning in Mixed Realities. A Retrospective on Transitional Objects. In A. Lazoudis, H. Salmi, & S. Sotiriou (Eds.), Augmented Reality in Education - Proceedings of the "Science Center To Go" Workshops (pp. 69–78). Athens, Greece: EPINOLA S.A.

Buchholz, Hagen, Brosda, C., & Wetzel, R. (2010). Science Center To Go - A Mixed Reality Learning Environment of Miniature. In Proceedings of the "Learning with ATLAS@CERN" Workshops Inspiring Science Learning EPINOIA (pp. 85–96). Crete.

Buchholz, Hagen, & Wetzel, R. (2009). Introducing Science Centre TOGO - A Mixed Reality Learning Environment for Everyone's Pocket (pp. 104–110). Presented at the IADIS International Conference Mobile Learning, Barcelona, Spain.

Hassenzahl, M., Prinz, W., Buchholz, H., & Laschke, M. (2011). Zwischen Können und Tun liegt ein großes Meer und auf seinem Grunde die gescheiterte Willenskraft. I-com Zeitschrift Für Interaktive Und Kooperative Medien 10, 10(2).

Larsen, Y. C., Buchholz, H., Brosda, C., & Bogner, F. X. (2012). Evaluation of a portable and interactive augmented reality learning system by teachers and students. In A. Lazoudis, H. Salmi, & S. Sotiriou (Eds.), Augmented Reality in Education - Proceedings of the "Science Center To Go" Workshops (pp. 47–56). Athens, Greece: EPINOLA S.A.

Lazoudis, A., Salmi, H., Sotiriou, S., Bogner, F. X., Owen, M., Owen, S., … and others. (2012). Science Center To Go: Guide of Good Practice. Athens, Greece: EPINOLA S.A.

Curriculum Vitae

Hagen Buchholz studied Computational Visualistics at the Otto-von-Guericke University of Magdeburg. From 09/2006 until 06/2013 he was working at FIT's Collaborative Virtual and Augmented Environments department. At his time at Fraunhofer, he worked in several national and international projects. His main research field at Fraunhofer comprised computer supported learning and contextual computer systems. He was responsible for several projects related to the "Science Center To Go" and FIT's rapid prototyping lab.

Dipl.-Ing. Hagen Buchholz

Geboren am 27.11.1979 in Wiesbaden

Ledig

seit 11/13	Stipendiat des „EXIST-Gründerstipendiums Unitab“ an der RWTH Aachen in Kooperation mit Fraunhofer FIT (gefördert durch das Bundesministerium für Wirtschaft und Technologie)
03/10-08/2014	Promotionsstudent der RWTH Aachen
09/06-06/13	Wissenschaftlicher Angestellter am Fraunhofer-Institut für Angewandte Informationstechnik (FIT) – Arbeitsgruppe Kollaborative Virtuelle und Augmentierte Umgebungen, später Mixed and Augmented Reality Solutions

	Mitarbeit, System Implementierung und Projektleitung in mehreren nationalen und internationalen Forschungs- sowie Industrieprojekten
	Verantwortlich für die Entwicklung einer Mobilen Augmented Reality (AR) Lernumgebung; Projektakquise & Verantwortung (EXPLOAR, SCeTGo, PATHWAY)
	2008/2009 Lehrtätigkeit am BIT (Bonn-Aachen International Center for Information Technology) und der RWTH Aachen: VR/AR; Game Design and Development;
	01/2007-12/2009 Mitglied des FIT Institutslenkungsausschuss
	2008-6/2013 Verantwortlicher des Rapid Prototyping Labs am Fraunhofer FIT
	Seit 2013 Gutachter für internationale Konferenzen u.a. ISMAR und NordiCHI
04/06-09/06	Research Scholar am Virtual Reality Lab der University of Michigan
01/06-03/06	Freiberuflich: Entwicklung interaktiver Lerneinheiten für die Impara GmbH
10/00-12/05	Studium der Computervisualistik an der Fakultät für Informatik der Otto-von-Guericke-Universität Magdeburg mit dem Schwerpunkt „Konstruktion und Fertigung“ (Abschlussnote: 1.4);
	2004: Erwerb des Bakkalaureats Abschlusses.
09/03-02/04	Auslandssemester am Virtual Reality Lab der University of Michigan (Ann Arbor, US)
10/99–08/00	Zivildienst Verwaltungstätigkeiten in der Zivildienstgruppe Frankfurt
08/90-06/99	Schulausbildung abgeschlossen mit Abitur am Wilhelm-Hofman-Gymnasium in St. Goarshausen