
The effect of feedback in adaptive toolbars

Worksheets

Title:

The effect of feedback in
adaptive toolbars
Worksheets

Theme:

Master's thesis

Project period:

PDP10, spring 2011

Project group:

1077

Authors:

Anders Grønne
Brian Hansen

Supervisor:

Lars Bo Larsen

Copies: 4

Number of pages: 74

Article: 11

Worksheets: 47

Appendix: 13

Attachment: CD

Finished: 31st of May, 2011

Abstract:

Adaptive User Interfaces have been a subject of research since 1985. No existing research has attempted to draw the user's attention towards changes in the adaptive elements. An adaptive toolbar was implemented in a simulated text editor application. Feedback was used to inform the participant of updates in the adaptive toolbar. The effect of this feedback was tested using an eye-gaze tracker. It was found that feedback affected both the toolbar used by the participants to solve tasks, and in which toolbar the participants first sought the solution.

Preface

This study is the result of a 10th semester project written by group 1077 studying Engineering Psychology at Aalborg University under the School of Information and Communication Technology.

All references in this report are written according to the Harvard referencing method, though sometimes with added information, e.g. a page number. The attached CD in the back of these works sheets contains electronic appendices and a digital copy of the article produced and these work sheets. See Chapter 7 on page 48 for a list of contents on the CD.

School of Information and Communication Technology
Aalborg University, 31st of May 2011

Anders Grønne

Brian Hansen

The content of this report/these worksheets is freely available, but publication of the report/worksheets and the content only be done with the consent of the authors.

Contents

1	Introduction	1
2	Previous research	3
2.1	User action prediction algorithms	3
2.2	Guidelines for adaptive user interface design	3
2.3	Tests of adaptive user interfaces	5
2.4	Eye tracking	7
3	Theory of learning	8
3.1	Dreyfus and Dreyfus - Five stages of skill acquisition	8
3.2	David A. Kolb - Experiential learning	10
3.3	Learning in Adaptive User Interfaces	12
3.4	Feedback	14
4	Method and materials	17
4.1	Hypothesis	17
4.2	Purpose	17
4.3	Test application	19
4.4	Experiment software GUI	23
4.5	Method	29
4.6	Materials	33
5	Results	35
5.1	Participants	35
5.2	Toolbar used	35
5.3	Eye-gaze tracking results	39
5.4	Task Duration	41
5.5	Questionnaire analysis	42
5.6	Summary	45
6	Discussion	47
7	Appendices on the attached CD	48
	Bibliography	48
I	Appendices	51
1	Literary review notes	52

Introduction

As everyday software applications become increasingly numerous and include an increasing amount of features, a user is required to adapt to an increasing amount of visually different graphical user interfaces (GUIs) and the possible ambiguity of the functions they contain. In addition to this, it is often sought to allow for the highest degree of usability in software applications for any user regardless of skill level. One solution, which also somewhat demonstrate how frustrating it can be to achieve these goals, was suggested by Benyon [1993a]:

Attempting to selectively breed humans so that particular characteristics become dominant may be one way of improving usability, but one that would probably be deemed politically unacceptable.

Should selective breeding indeed prove politically unacceptable, another possible solution exist in adaptive user interfaces (AUI). The purpose of an AUI is to adapt the GUI and possibly its functions based on information gathered from a user. Benyon [1993b] defined a general model for adaptive user interfaces, consisting of three parts. The model consist of a model of the system, a model of the interaction, and a model of another system, which the system can adapt to. He furthermore describes a more comprehensive model of an adaptive human-computer system, which is capable of adapting to a wide range of parameters, e.g. user, interaction, and domain. Lastly he describes how the comprehensive model can be used to change a “normal” user interface to an adaptive user interface. This method includes a system analysis which investigates relevant parameters of a human-computer system. The exact form of an AUI depends highly on the software application in question, but almost any generic software application is a possible client for at least some AUI features. However a certain degree of complexity is required in order for it to make sense to implement an AUI. Benyon [1993a] states:

A theory of adaptivity should help us to decide when an adaptive system solution to usability problems is appropriate. There is little to be gained if expensive adaptivity mechanisms are used to achieve a minimal improvement in usability. There is no point in adapting to some user characteristic if that characteristic cannot be reliably and unobtrusively inferred from the interaction.

How this adaptation manifests itself in the application is both affected by the type of software, but also the algorithm used for controlling the adaption. Algorithms may vary greatly in complexity, and can range from a simple gathering of the “most frequently used functions” and to a prediction algorithm so advanced, that it might be impossible for the user to derive its functionality. This also raises one of the known usability-issues when working with AUIs. As transparency of interaction is one of the success criteria for high usability (as suggested

by e.g. Wensveen et al. [2004]), i.e. allowing the user to understand what goes on “behind the scene”, an algorithm predicting the function the user requires next with a high accuracy, but appear inscrutable in doing so, could be rated as a poor algorithm by the user. Another important part of transparency of interaction is feedback and feed-forward [Wensveen et al., 2004]. If the user is to learn the interface, the user must understand how the interface works. One way for the user to learn how the interface works, is to provide him with feedback as to why - in regards to AUIs - the interface updated as it did. Likewise feedforward could be utilised. Contrary to feedback, which occurs during or after interaction, feedforward is the information on its functionality provided by the product. Wensveen et al. [2004] describes three different types of feedforward/feedback. These are:

Functional The action generated by the product when performing its intended action.

Augmented Additional information provided by the product besides its intended action

Inherent The information provided as a natural consequence of performing the action.

Inherent feedback cannot be applied in a computer desktop environment, as the only inherent feedback provided is the feel of the keyboard and mouse buttons. However the remaining two information channels are available in Human-Computer Interaction.

AUIs separate into two major groups. Full adaptive interface where almost the entire interface adapts itself to the user and adaptive toolbars, where just a small part of the interface is adaptive. This project will use University of Strathclyde [2011]’s definition of a toolbar. They define a toolbar as:

In the graphical user interface (GUI), a toolbar is a horizontal row or vertical column of selectable images (buttons) that carry out certain frequently used actions when clicked.

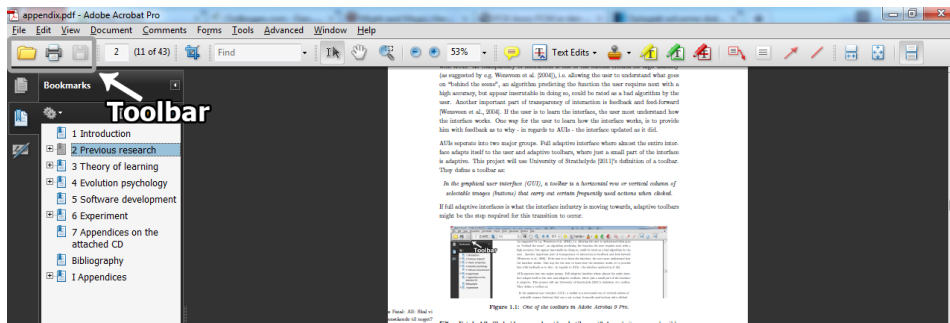


Figure 1.1: One of the toolbars in Adobe Acrobat 9 Pro.

Previous research

Adaptive User Interfaces has been an area of research since the mid 1980s, but research in the area did not really pick up till the turn of the millennium. However, in spite of its young age the field already holds various areas of interest/focus. This chapter will outline the studies conducted in some of the different areas.

2.1 User action prediction algorithms

An essential part of any AUI is the algorithm which analyses the user actions and adapts the user interface accordingly. The algorithm selected often depends on how the AUI is sought to adapt. A simple example of this, is the *most frequently used* algorithm, which can be found in the Start menu in the Windows operating system (versions: XP/Vista/7). This algorithm will simply count the amount of times a user has launched each of the installed application and combine this with some sort of time measure, then update the Start-menu to display the ones used most frequently. This is relatively simple algorithm, and most others are far more complicated.

Nazemi et al. [2010] presented an extended version of the KO-algorithm (called KO*/19) which attempted to predict the user's intended action in order to make the task easier/more efficient for the user. Nazemi et al. [2010]'s method records every interaction event observed in the UI and sends these to be analysed to capture user information. The user information, which the analysis is capable of capturing, is sent back to the UI module again and used for user-centered adaption.

2.2 Guidelines for adaptive user interface design

“Improving the usability of computer systems is perhaps the most important goal of human-computer interaction research.” [Benyon, 1993b]

Adaptive User Interfaces can be applied on a wide range of devices, ranging from monitors on mobile devices to systems designed for 100 inch displays. The number of user interfaces, which could be changed to an adaptive user interface, is staggering when just considering the wide range of typical applications used on a normal desktop computer. As a result hereof, the literature regarding user interface design for adaptive user interfaces is often split into several, different categories depending on the device and application the interface is designed for.

Thomas [1993] defined six general guidelines which adaptive user interfaces should adhere to. These include: always let the user be in complete control, prevent system suggestions from disturbing the user unnecessarily, and allow the user to easily reset adaptations. Thomas [1993] applied these guidelines to an adaptive user interface built on top of Microsoft Excel 3.0 for MacIntosh® called Flexcel. The Flexcel adaptation tool tip can be seen in Figure 2.1. A user test showed that participant unanimously appreciated Flexcel as it “[...] provides a comfortable access to (different) user/task specific profiles and for reasons of privacy.” [Thomas, 1993] Gajos et al. [2006] designed three different adaptive techniques, which were tested against a non-adaptive interface. These were:

Moving Interface: Items were moved from their original placement to the main toolbar - Figure 2.2a on page 5.

Split Interface: Items were copied from their original placement to an adaptive toolbar - Figure 2.2b on page 5.

Visual Popout Interface: Items were highlighted with a different colour - Figure 2.2c on page 5.

Gajos et al. [2006] adjusted two parameters to investigate which had the biggest influence on user acceptance of adaptive user interfaces. These were precision of predictability and frequency of adaption. They found that designer should strive for a high precision and that the frequency of adaptations must not be too high, furthermore they came to the following three conclusions:

- The frequency of interaction and task complexity decided which parts of the adaptive user interface users found relevant
- Users appreciated all relevant information being grouped
- User interfaces with high complexity benefit from drastic adaptations

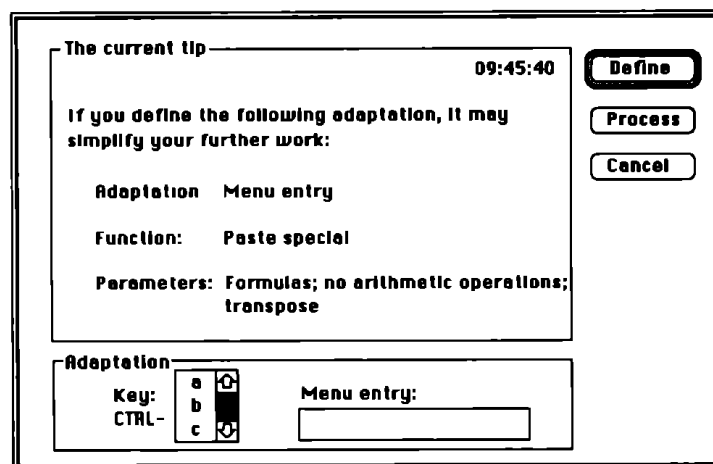


Figure 2.1: The Flexcel adaptation tool tip. [Thomas, 1993]

Liu et al. [2003] used *Episode Identification and Association* to design an adaptive user interface which employed a five step approach to recognise user actions:

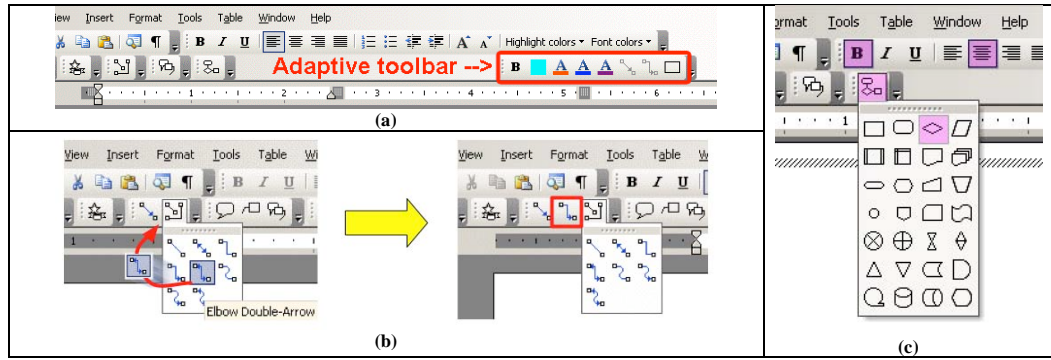


Figure 2.2: *The interfaces used by Gajos et al. [2006]. (a) Split Interface, (b) Moving Interface, and (c) Visual Popout Interface*

1. Observe the interaction
2. Identify episodes
3. Recognise user behaviour patterns
4. Assist users in their actions (plan)
5. Build user profile

Pilot studies indicated higher productivity with the designed interface.

2.3 Tests of adaptive user interfaces

Several studies have concerned themselves with the specific parameters which influence user acceptance of adaptive user interfaces.

Gajos et al. [2008] conducted a user experiment where the Split Interface (see Section 2.2 on page 3 for a description of the Split Interface) integrated into Microsoft Word and found that increased accuracy and predictability improves user satisfaction and increased accuracy had larger positive effect on performance, utilisation, and some satisfaction ratings than increased predictability.

Miller and Hannen [1998] conducted an experiment with a adaptive Cockpit Information Manager system, where actual pilots flew extensive mission simulations. The subjective data indicates a perceived efficiency increase and workload decrease.

Tsandilas and Schraefel [2005] tested two techniques for adaptive textual lists. The NORMAL technique (Figure 2.3a on 6) suggested items by changing the background colour and the SHRINK technique (Figure 2.3b on 6) used a fish-eye effect around the cursor, but kept suggested items in the same font size as the items around the cursor. They found that accuracy had an influence on user performance both when the target was available in the adaptive toolbar and when it was not. They argue that this was due to the decreased reliability of the low accuracy condition.

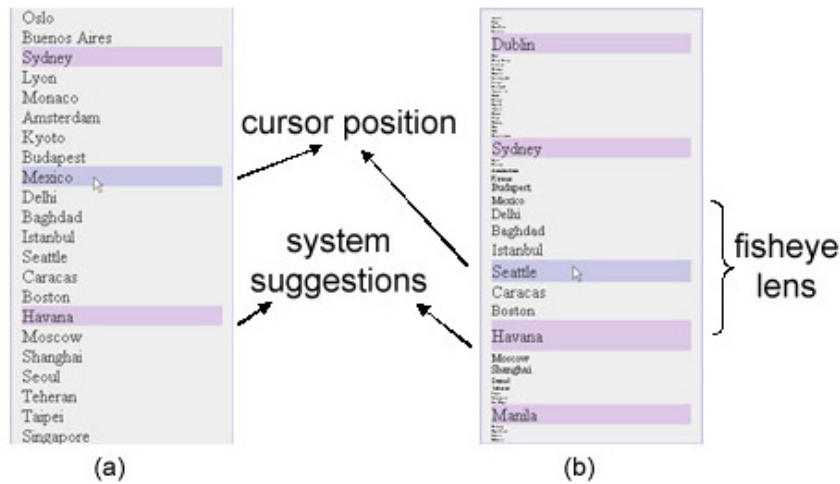


Figure 2.3: The two adaptation techniques used by Tsandilas and Schraefel [2005]. (a) *NORMAL* and (b) *SHRINK*

Findlater and McGrenere [2008] compared AUIs on small and large monitor sizes. They tested three user interface conditions (static, high accuracy AUI, and low accuracy AUI). The test application used for the large monitor conditions with an adaptive toolbar can be seen in Figure 2.4. The adaptive menus followed a *most recently used* algorithm. For the large screen they found no difference in efficiency between the static and the high accuracy (78.5 %) interfaces, but found the low accuracy (50 %) interface to be significantly less efficient than the static interface.

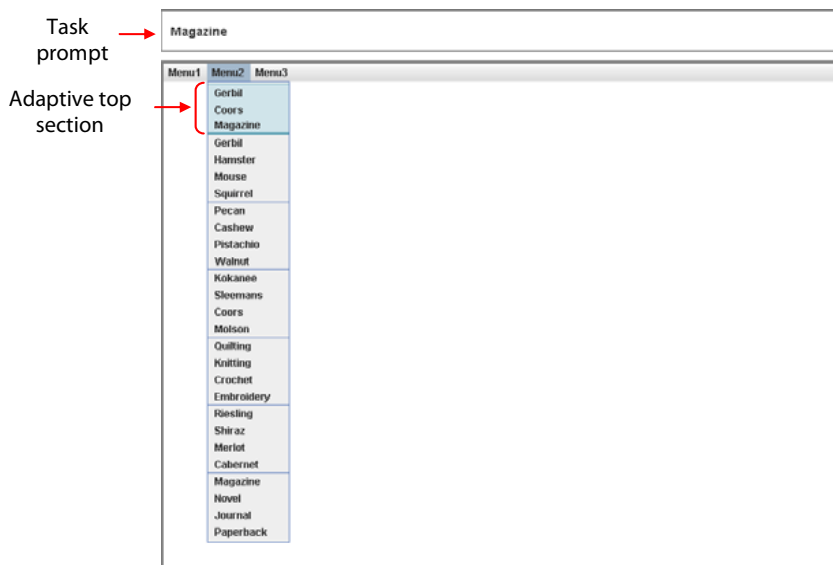


Figure 2.4: The large monitor with the adaptive toolbar used by Findlater and McGrenere [2008].

In a study by Paymans et al. [2004] participants viewed videos in a lounge and at a bus stop on an environment-aware mobile device. They found that assisting users in building a mental model of how the adaptive user interface worked improved ease-of-use, but decreased learnability.

2.4 Eye tracking

To conduct a proper eye tracking experiment former eye tracking experiments and NN Group's eye tracking methodology [Pernice and Nielsen, 2009] were investigated.

Castagnos and Pu [2010] created an e-commerce website based on the norm at the time. Participants were asked to find three products for themselves and one product for a gift. An eye tracker was used to count number of fixations on predefined areas of interest (AOI) and to generate heat maps. Eye tracking data was used because it was assumed that what person looks at can be used as an indication of their cognitive process.

Eye tracking data assisted Pretorius et al. [2005] in locating problems in the design of a network management tool. It was found that participants spent too long searching for crucial functions.

Furthermore, the methodologies in the studies by Goldberg et al. [2002] and Conati et al. [2007] were investigated to gather experience on how to conduct a usability experiments when using eye tracking.

Theory of learning

Having reviewed a large portion of the work conducted in the field of AUIs, one approach was found largely untouched. When exploring the interaction between an adaptive system and a learning user, it seem logical to examine how people go about attaining new skills and knowledge. Several models of learning seek to explain just this. As such, the following chapter will seek to explore different theories on learning and seek to conclude on any useful information these areas may provide to the field of AUIs.

3.1 Dreyfus and Dreyfus - Five stages of skill acquisition

One important distinction of knowledge used by Dreyfus et al. [1988] is the difference between “know that” and “know how”. “Know how” is knowledge which a person may posses, but is unable to explain in terms of rules and facts. An example hereof used by Dreyfus et al. [1988] is that of riding a bike. This skill may be acquired by practice and experience. However, a person proficient at bike riding can not - using only words - teach another person the same set of skills. As such, the knowledge is “know how” as the person may be unable to explain his skills, but still master them at a very high level. If a person is indeed capable of teaching the specific skill to another individual using facts and rules, the knowledge would instead be “know that” [Dreyfus et al., 1988].

Dreyfus et al. [1988] observe that human beings acquiring a certain skill, does not simply leap from rule based “knowing that” to the more experience based “knowing how”. Instead, Dreyfus et al. [1988] identifies five stages of qualitatively different perception of a individual tasks and/or modes of decision-making. These five stages are: Novice, advanced beginner, competent, proficient, and expert. Dreyfus et al. [1988] notes, that not all people will achieve an expert level in their skills. This also largely depends on the skill in question. According to this model, all skill acquisition begins at the novice level and works towards the expert level - constantly building on previously attained knowledge.

3.1.1 Novice

On the novice level an individual has no previous experience to reflect upon and as such can not make decisions based on this. Instead, the novice is bound by rules put forth concerning the specific skill, which instead will be used in decision making. The novice is said to act on context-free rules and the specific situation. The novice will hence only have access to “knowing that” knowledge. On the novice level the first experiences with the specific skill is attained. This knowledge will be used on the following skill levels.

3.1.2 Advanced beginner

After the first experiences are made, the advanced beginner level is characterised by the initial use of previously attained knowledge to adapt action to circumstances. Through practical experience in real situations containing meaningful elements, which are immediately undefinable, the advanced beginner recognizes certain context-free rules to this knowledge. The individual attains the first “know how” to the specific skill.

3.1.3 Competent

As the amount of experiences increases, the “competent” has attained so many context-free and situational rules that they may appear overwhelming in a actual real life situation. The competent is somewhat still unable to priorities tasks as being of different importance. Instead, the competent will regard a problematic situation as a set of individual tasks, which must be overcome in order to proceed. An example hereof could be a new doctor at an emergency room, who is asked to examine and treat all patients. If the doctor is at the competent level, he will systematically treat all patients without regarding the severity of their condition, i.e. not prioritising critical cases. So even though the doctor has a lot of practical experience treating patients, he is still unable to see the whole picture.

3.1.4 Proficient

Up until the “proficient” level of the skill acquisition, those who have attained new knowledge consciously made their decisions - or reflected on alternatives to these decisions - based on previous experience. As a “proficient” these decisions are more often replaced by a routine-based - almost intuitive - basis for decision. Though, still based on previous knowledge. Individuals on this level have experienced almost all possible scenarios to such a degree and to such an extent, that it provides the possibility of simply relating to a previous experience, which sufficiently resemble the current problem, and use that solution without further consideration or reflection.

3.1.5 Expert

The highest level of skill “acquisition” is the expert-level, which can not necessarily be achieved with all skills and by all individuals. In general, the expert will in all situations have a suitable solution to any issue (regarding that skill), based on a mature and practical deliberation. When an expert is working his skill, problems are no longer recognised, possible future issues no longer considered, nor future actions planned. Instead - as Dreyfus et al. [1988] chose to put it:

An expert's skill has become so much a part of him that he need be no more aware of it than he is of his own body.

So assuming that things work normally, an expert is not working on problem solutions or decision-making. He simply does that, which usually works. One example hereof would be the difference between a novice pilot and an expert pilot. The novice pilot will claim that he is flying a plane, whereas the expert pilot will claim that he is the one flying. The difference being in the user's perception of the performed action. [Dreyfus et al., 1988]

3.2 David A. Kolb - Experiential learning

David A. Kolb devised a model to explain how experiential learning functions as a continuous process, which is highly dependable on the individual in question's personal experience and learning preference. Kolb et al. [1984] work was in particular based on the previous work by John Dewey, Kurt Lewin, and Jean Piaget. One aspect in particular one must be aware of when discussing Kolb's approach to learning, was that he considered it a lifelong process. I.e. Kolb considers learning not to be limited to the knowledge obtained in schools, universities and other learning institutions - but something which occurs across school, work, and personal development. As such, Kolb et al. [1984] defined a number of parameters/characteristics to define how he thinks learning should be interpreted as, in regards to his own theories:

Learning is best conceived as a process, not in terms of outcomes

Classic behaviourist approach to learning and idealist educational approaches considers there to be a set amount of elements of consciousness - also referred to as "mental atoms" or "simple ideas" - and it is the variations in the combinations and associations between these ideas, which supposedly form the different patterns of thoughts. In other words, classical behaviourist approach to learning considers there to be constant, fixed elements of thoughts. This approach has resulted mainly in the tendency of regarding learning in regards to its outcomes. If ideas are indeed fixed and immutable, Kolb notes, it would seem possible to measure the amount learned by a given individual, simply by measuring the amount of these fixed ideas that individual has accumulated. Kolb et al. [1984]

Experiential learning have different - almost opposing - assumptions, i.e. that ideas are not fixed and immutable elements of thought. Rather, in the field of experiential learning ideas are considered formed and re-formed through experience. Kolb et al. [1984]. Put in another way, no two thoughts are ever the same since experience is thought to always influence them.

Learning is a continuous process grounded in experience

Knowledge is continuously derived from and tested out in the experiences of the learner.

Kolb et al. [1984]

Kolb states that it is in the space between expectations and experience that learning takes place. As such, any experience which does not differ from one's expectation in even slightest manner can not teach the individual anything, and is as such useless as an experience in this regard. Instead, since learning is a continuous process, which occurs based on experience - it implies that all learning is in fact re-learning. Kolb considers all people to have articulated theories to all possible fields of study, the difference lies in the correctness, validity and crudeness of these theories. This validity of these theories, however, is not the point in this matter. The point is that until people were taught/gathered experience in a given field, these crude theories were all the "knowledge" which these individuals had and it was these theories they would draw upon, whenever they were required to deal with these particular fields.

The process of learning requires the resolution of conflicts between dialectically opposed modes of adaptation to the world

The models of the 3 main authors, on which Kolb based his work: John Dewey, Kurt Lewin, and Jean Piaget, all suggested that learning in itself is a conflict-filled process. As such, knowledge is attained through confrontation between the four "modes" of experiential learning seen in Figure 3.1 on the following page. The four modes being *concrete experience*, *reflective observation*, *abstract conceptualisation*, and *active experimentation*. Thus, learning requires an individual to use abilities, which are opposing, and the learner must continuously decide which of these learning abilities to use in any learning situation. The two primary positions a learner can choose are concrete experiencing and abstract conceptualising. Put simply, the learner can move from being an *actor* or an *observer* in various degrees, or in other words use a "specific involvement" or a "general analytic" approach. Kolb et al. [1984]

Learning is an holistic process of adaption to the world

Learning is not just concerned with attaining new knowledge and skills, it also includes the general adaption of humans. This adaption is holistic, in the sense that includes all the integrated functions of the organism, i.e. thinking, feeling, perceiving, and behaving. Kolb et al. [1984] As such, learning allows the human organism to adapt - socially and physically - to its specific (or a general) environment.

Learning involves transactions between the person and the environment

Kolb stated this, as he found this seemingly obvious proposition to be ignored in research on learning and practice in the line of education, and thus found it necessary to underline as an essential aspect of learning.

Learning is the process of creating knowledge

Knowledge is the result of the transaction between social knowledge and personal knowledge.

Kolb et al. [1984]

Social knowledge is thought as the accumulation of previous human cultural experience, and can as such be thought to be *objective*, whereas personal knowledge is the accumulation of personal experiences and is as such *subjective*. Knowledge, hence, is the transaction between *objective* and *subjective* knowledge via the process of learning. As such, it is essential to understand the psychology of the learning process, before we can begin to understand knowledge. Kolb et al. [1984]

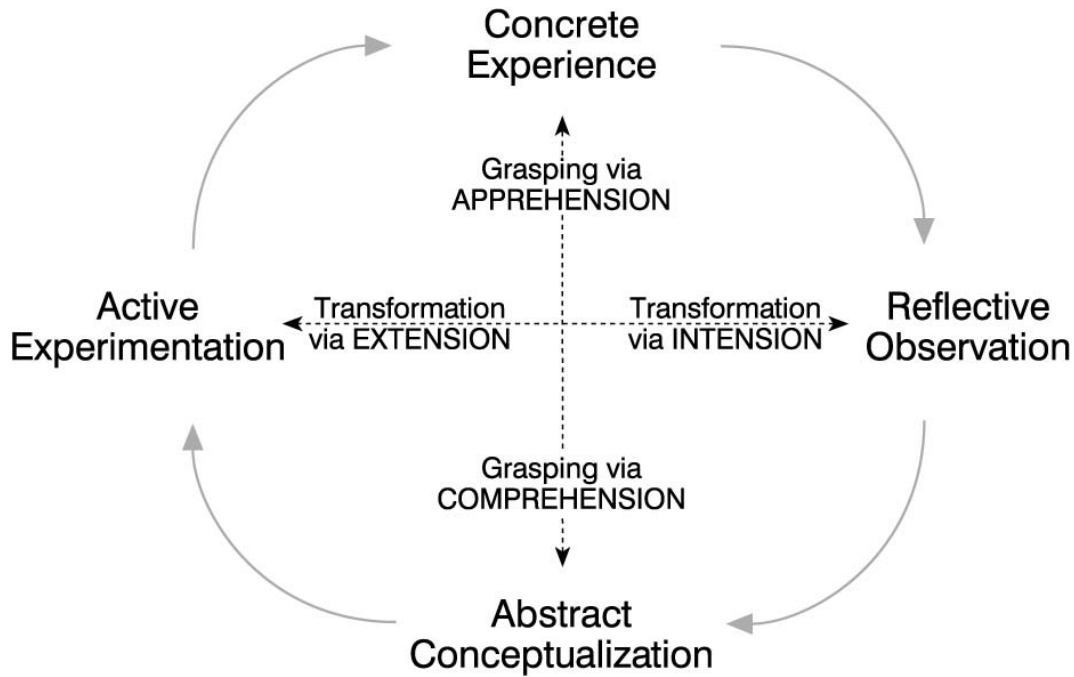


Figure 3.1: David A. Kolb's model for experiential learning

Having defined the parameters for learning, Kolb et al. [1984] continues to explain the process of obtaining new knowledge. This process is depicted in Figure 3.1 and is known as the *Experiential Learning Model* (ELM). The figure is relatively self-explanatory. At any time, an individual will have a certain set of concrete experiences. The individual is able to observe and reflect upon these experiences. Based on this reflection, the individual can form new abstract concepts/hypotheses. Having formed these new abstract concepts allows the individual to test these concepts in actual situations. This will provide the individual with new experiences, which can be reflected upon. I.e. the circle is complete and can be started anew.

3.3 Learning in Adaptive User Interfaces

Dreyfus & Dreyfus' model of different levels of skill acquisition could provide a clear distinction between users and the skills they possess at a specific time. Certainly, such a partitioning of users into different groups would provide an AUI algorithm with a basis on which to work upon. However, there are several problems in using this particular model.

First of all, the time it will take to reach the higher levels of Dreyfus' model makes it imprudent to have it as a set goal. One of the defining characteristics of the *expert-level* of the skill acquisition model, is that the user have tried every possible scenario in that skill to such an extent, that all decisions are made autonomously. However, if the AUI is constantly changing to best suit the user's need - based on the interaction between user and system - new scenarios will continue to arise; potentially even after many years of use. Put in another way: a user can never become an expert - as defined by Dreyfus et al. [1988] - with an AUI system. This would only be untrue if the user and system interaction reaches a state of total equilibrium. Even though it may not be necessary for a user to ever reach the expert-level with an AUI system, the general purpose of an AUI is to improve the interaction between user and system on a shorter-than-normal basis. Ideally, a user should be able to use the system to fulfill all needs that user may have, at any time during interaction. As such, the Dreyfus' model of skill acquisition does not provide any actual insight into how this may be achieved. It can instead be used for predicting how a user will adapt, as it provides some insight into decision making in a learning perspective. This, however, is not the primary goal of the present study.

Kolb's model of experiential learning offers an explanation as to how a user continuously attain new experiences during interaction, how these experiences are reflected up and formulated into hypotheses, which are then tested. This feedback model provides insight into how a new user will begin to understand a new software application; be it an AUI system or not. As users are thought to use their existing experience when approaching the system the very first time, it will likely be the experience of using similar systems, which will be drawn upon. This already provides an undesirable basis for an AUI system, when compared to a non-AUI system. As AUI systems are somewhat limited in their prevalence and implementation, most users will have more experiences with non-AUI system. As such, it is reasonable to assume that most new users will assume that a system is non-adaptive. This means that users must first be convinced that this is indeed an adaptive system - either by discovering it themselves or by having the system explain it. Thus, when using any AUI system the user must first have their assumptions corrected to fit the actual context.

One of the main issues with AUI system is that they change - possibly in so complex a manner that user are unable to derive the basis for the adaptation - in attempt to best fit the user. If one follows the experiential learning model put forth by Kolb et al. [1984], a user will test their hypothesis on how things work and reflect upon the results, i.e. work towards getting an understanding of the underlying model of the system. In an AUI system, the system itself will work towards modeling the user in order to change accordingly. In accordance with Kolb's model, this could easily result in the user making incorrect assumptions, simply because what may have been a correct assumption of the system at one point has changed later as the system adapts. Not only could this easily prolong the time it will take to learn/adapt the system, but also be the source of great frustration. To prevent this from happening, it must be made clear to the user why - or at the very least when - changes occur in the user interface.

3.4 Feedback

Feedback is used extensively in the field of HCI. “*Give Useful, Informative Feedback*” is one of Kortum [2008]’s eight guidelines for user interface designing with good usability in mind. Feedback can be defined as:

The return of information about the result of a process or activity.

[*The American Heritage® Stedman’s Medical Dictionary*, 2011]

Textual messages are not the only form of feedback employed in HCI. Windows 7 provides visual cues when an inactive application is attempting to acquire the attention of the user. The background of the application icon in the taskbar pulsates with a different colour. An example of this is provided in Figure 3.2.



Figure 3.2: *The pulsating background employed by Windows 7.*

3.4.1 Interaction frogger framework

Wensveen et al. [2004] addresses the issue of incoherence between user action and product function in electronic products. In mechanical products, it is more likely that the user understands the coupling between action and function, since these are naturally coupled. This coupling is not naturally inherent in electronic products, where the user requires information to properly guide his actions towards intended product function [Wensveen et al., 2004]. In an attempt to unify action and reaction and thus making the interaction seem more intuitive, Wensveen et al. [2004] put forth a framework consisting of the following six guidelines:

- Time - The product’s reaction and the user’s action coincide in time.

- Location - The reaction of the product and the action of the user occur in the same location.
- Direction - The direction or movement of the product's reaction (up/down, clockwise/counterclockwise, right/left and towards/away) is coupled to the direction or the movement of the user's action.
- Dynamics - The dynamics of reaction (position, speed, acceleration, force) is coupled to the dynamics of the action.
- Modality - The sensory modalities of the product's reaction are in correlation with the sensory modalities of the user's action.
- Expression - The expression of the reaction is a reflection of the expression of the action.

3.4.2 Feedback

Wensveen et al. [2004] also considers feedback an essential part in interaction design, and further defines three degrees of feedback, which each have their uses.

Functional feedback

Functional feedback is defined as the information generated by the system, when the system performs its function. An examples of functional feedback is music from a radio after having switched it on.

Augmented feedback

Augmented feedback uses a “not-directly-related-to-function” indication as an affirmation of having received input. This is commonly used when there is a delay in functional feedback, e.g. an LED light lights up when turning on a television.

Inherent feedback

Inherent feedback is the direct feedback - and thus often motor-perceptual - from performing an action possibility on a product. Inherent feedback could be that of feeling resistance from a button while pressing it down and also hearing it “click”. Often inherent feedback stems from the movement related to performing an action potential.

3.4.3 Feedforward

If a device is able to communicate information regarding the consequences of a function before use, it is called feedforward. Wensveen et al. [2004] also distinguishes feedforward as being one of three forms.

Functional feedforward

Functional feedforward provides more general information about the purpose and functions of a product. Ways of achieving functional feedforward is letting functional parts be visible to the user, or exploiting the user's semantic understanding.

Augmented feedforward

Augmented feedforward relates to the cognitive skills of a user, by feedforwarding information to the user about an action possibility - or the purpose of it. Augmented feedforward is achieved through written text, pictograms, spoken words, etc.

Inherent feedforward

Inherent feedforward relates to the action potential of the product and the user's motor-perception. It informs the user of what action can be carried out (pushing, sliding, rotating, etc.), and also which body part the user must use in order to use this function. Inherent feedback can be considered a form of limited affordance.

These guidelines are combined into the model put forth by Wensveen et al. [2004], which can be seen in Figure 3.3.

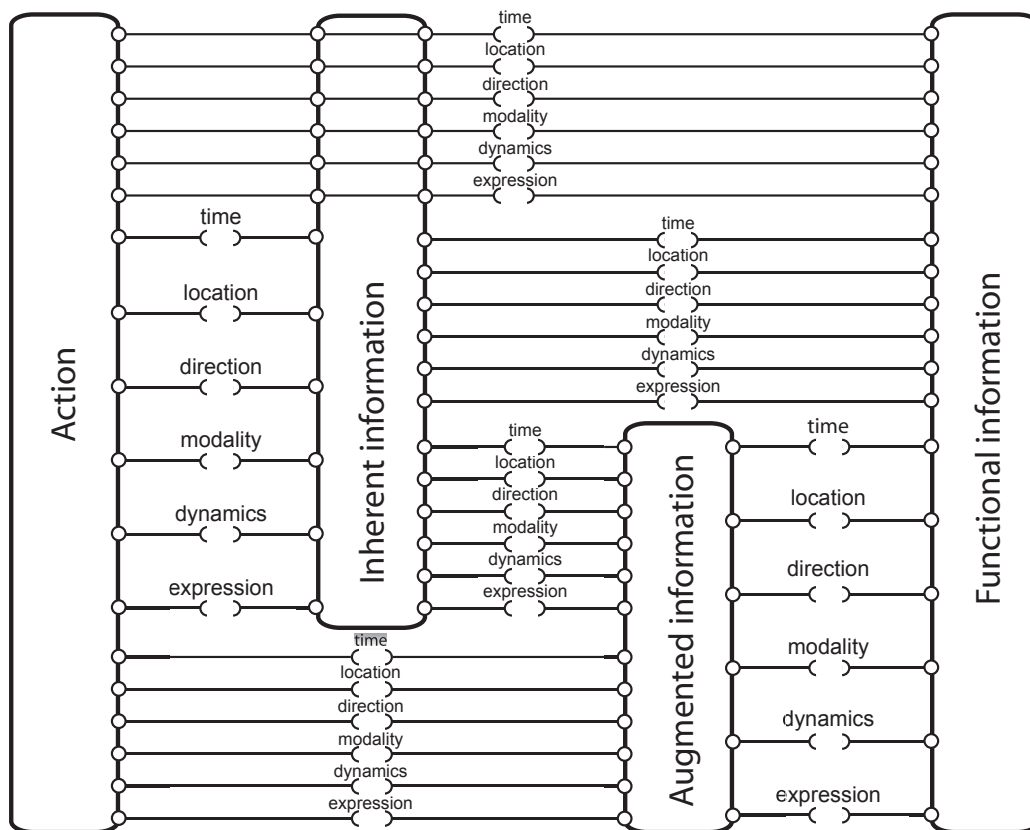


Figure 3.3: The model put forth by Wensveen et al. [2004] to couple action and function in interaction design. The model is designated The Interaction Frogger Framework.

The *Interaction Frogger Framework* provides an approach to analyse and designing products, irregardless of it being electronic or mechanical, to investigate how it provides information to a user.

Method and materials

4.1 Hypothesis

Adaptive toolbars has the potential to increase the efficiency of computer applications. Several studies have investigated various aspects of AUIs (as detailed in Chapter 2 on page 3). However, none of these studies attempted to make it apparent to the user when the adaptive toolbar updated itself. It is assumed, that if the user is made aware of changes in the adaptive toolbar, they will make use of the adaptive toolbar more often compared to a non-feedback version. Based on this, the test hypothesis was:

Providing feedback on changes in an adaptive toolbar will increase the use of the toolbar.

4.2 Purpose

The purpose of the experiment was to investigate feedback's influence on users' use of an adaptive component of a user interface. To this end, a user interface with an adaptive toolbar with feedback, which informed the user of changes in the toolbar, was tested. An identical user interface without feedback served as a baseline. Qualitative and quantitative data were used to evaluate any difference between the two user interfaces.

4.2.1 GUI design

In order to verify the hypothesis, two versions of a test application were required. One where feedback was implemented and one where it was not. Furthermore, an eye-gaze tracker was required to monitor which part of the user interface the participant was looking at. It is important to consider the layout of the user interface when including an eye-gaze tracker in an experiment, as the eye-gaze tracker is not 100 % accurate. The eye-gaze tracker available for the experiment was the Tobii X120. It is accurate down to 0.5° and functions optimally when the eyes are around 60 cm from the eye-gaze tracker.

The eye-gaze tracking commenced as soon as the experiment was started. Thus the eye-gaze tracker calibration did not interfere with the flow of the experiment. All tasks were presented in the same area of the screen, thus providing a common eye-gaze starting point. The shifts in scene occurred at the start of each task. The primary purpose of the eye-gaze tracker was to detect which toolbar the participant gazed on first.



Figure 4.1: *A sketch of the GUI for the test application.*

The adaptive toolbar was placed in the top, right hand corner, whereas the static toolbars were placed in the top left hand side. The purpose of this was to create a visual angle of no less than 20° [Gajos et al., 2008]. Thus, a gaze shift was required if the participant was to observe items in the adaptive toolbar if their focus was on the static toolbars. This also removes any considerations as to the accuracy of the tobii eye-gaze tracker, as the visual angle sought for this experiment renders the ± 0.5 degree accuracy level irrelevant. By performing triangle calculations it was found that the minimum distance between the static toolbars and the adaptive toolbar must be 21.16 cm (20° from 60 cm away). The only difference between the non-feedback and feedback version of the test user interface was in the adaptive toolbar. In the non-feedback version, the adaptive toolbar made no effort to draw attention towards itself. The feedback version provided the participant with a visual cue, indicating a change in the adaptive toolbar. A sketch of the GUI of the test application can be seen in Figure 4.1

Static toolbars

The static toolbars contained all possible function buttons required to complete the tasks and several others. The buttons not related to any task simulated the amount of available functions presented in modern text editing applications. Some buttons were contained within

drop-down menus, these required the participant to click an initial button to access the underlying buttons. Buttons, which gave access to a drop-down menu, indicated this by having a small triangle pointing downwards beneath the button graphics. Also, the presence of the drop-down menus was indicated in the introduction text as described in Section 4.5.1 on page 31. The static part of the interface contained two rows of ten buttons. The bottom row were all drop-down menus, whereas the top row were all normal buttons. Each drop-down menu button contained between four and twelve additional buttons. In total the static toolbars contained 110 buttons.

Adaptive toolbar

The adaptive toolbar contained a predefined set of buttons for each task. This, to ensure the accuracy rate of the simulated algorithm. The adaptive toolbar contained the requested button in 66.6 % of the tasks or 10 out of 15 tasks. This level was chosen based on research on AUI accuracy level and their influence on user acceptance and performance [Tsandilas and Schraefel, 2005; Gajos et al., 2008; Findlater and McGrenere, 2008]. A glow pulsated around the entire adaptive toolbar twice to attract the attention of the participant. The adaptive toolbar contained exactly five buttons at any time, i.e. to allow new buttons to be added, old ones would be removed.

4.3 Test application

The software for the conducted experiment was developed for the Adobe® AIR® Player, in order to be able to write to files on the computer. The software can be found in Appendix 1 on the attached CD. The readme.txt file explains how to run the software. Please note that this requires Adobe® Flash® CS4.

4.3.1 Program Structure

Figure 4.2 on page 22 shows the class diagram of the software. The software is compiled from the `Main.fla` file, which is an instance of the `Main` class. The `InfoViewer`, `QuestionPanel` and `TopBar` classes are all associated with a `MovieClip` - in the `Main.fla` file - of the same name as the corresponding class. These `MovieClips` holds the graphical elements of the GUI, which are controlled by the code in the class files. The `DataWriter` class contains no graphical elements and is implemented fully as `ActionScript`®. The following gives a short description of each class.

Main class: This is the main class of the software. An object of this class controls the overall flow of the software and the visual elements. Objects of the other classes communicates through an object of the `Main` class. Furthermore, the constructor of the `Main` class contains the task definitions for both tasks sets and event handlers which communicate the unique identifier of the solution to the current task, the current task number, and the current tasks set to the `TopBar` class. This is used to regulate the content of the

adaptive toolbar, which is also described in the **TopBar** class description. The **Main** class also contains the event listener, which matched the currently pressed element with the unique identifier of the current solution. Upon completion of all tasks in a given task set, the **Main** class also handles the creation of the data string, which is sent to the **DataWriter** class.

InfoViewer class: This class is used to display information to the participant. The **MovieClip** associated with this class has two frames. The first frame contains the information shown to the participant at the beginning of the experiment. In this frame the facilitator inputs a subject ID number, the starting tasks set, and whether feedback is enabled at the first or second task set. The second frame contains the information shown to the participant when the experiment is completed. The **InfoViewer** class contains methods for displaying either of the two frames, and event handlers for starting the experiment, when the participant clicks the button to start the experiment.

QuestionPanel class: This class is used to handle the questions shown to the participants after the completion of a task set. Once question four has been answered, the button, which allows the participants to continue to either the next task set or the outro text, appears. Upon pressing this appearing button, the questionnaire is reset and hidden - ready to be used again.

TopBar class: The **TopBar** class controls the bar in the topmost part of the screen and the content within, i.e. the static and the adaptive toolbars. The **TopBar** also handles the drop-down function of the static toolbars, which is all the buttons on the lowermost part of the static toolbar. To control these, the **MovieClip** associated with this class contains 11 frames to handle the 10 drop-down menus. The first frame is simply the menu with all drop-down parts closed. The second frame will then contain the drop-down buttons for the first drop-down button, the third frame contains the content for the second drop-down button and so on. Using a **MouseEvent** listener, the **TopBar** class will jump to the frame containing the content for any pressed drop-down button. To only show one drop-down menu at a time, and also automatically close the drop-down menus, an invisible frame was built around the buttons appearing in the drop-down menus. This frame functions as a button, and is activated when a mouse is placed over the frame. This will send the timeline of the **MovieClip** associated with the **TopBar** class back to frame 1, in which all the drop-down menus are hidden. The **TopBar** class also handles the content of the adaptive toolbar. It receives the solution for the current task, the current task number, and the current task set from the **Main** class, and from this it defines the content of the adaptive toolbar.

DataWriter class: This class is used to write test data to a CSV file. One line is written per task set. Each line in the CSV file is of the template:

- Subject ID
- Feedback enabled/disabled
- Answer in Question 1 on the questionnaire
- Answer in Question 2 on the questionnaire

- Answer in Question 3 on the questionnaire
- Answer in Question 4 on the questionnaire

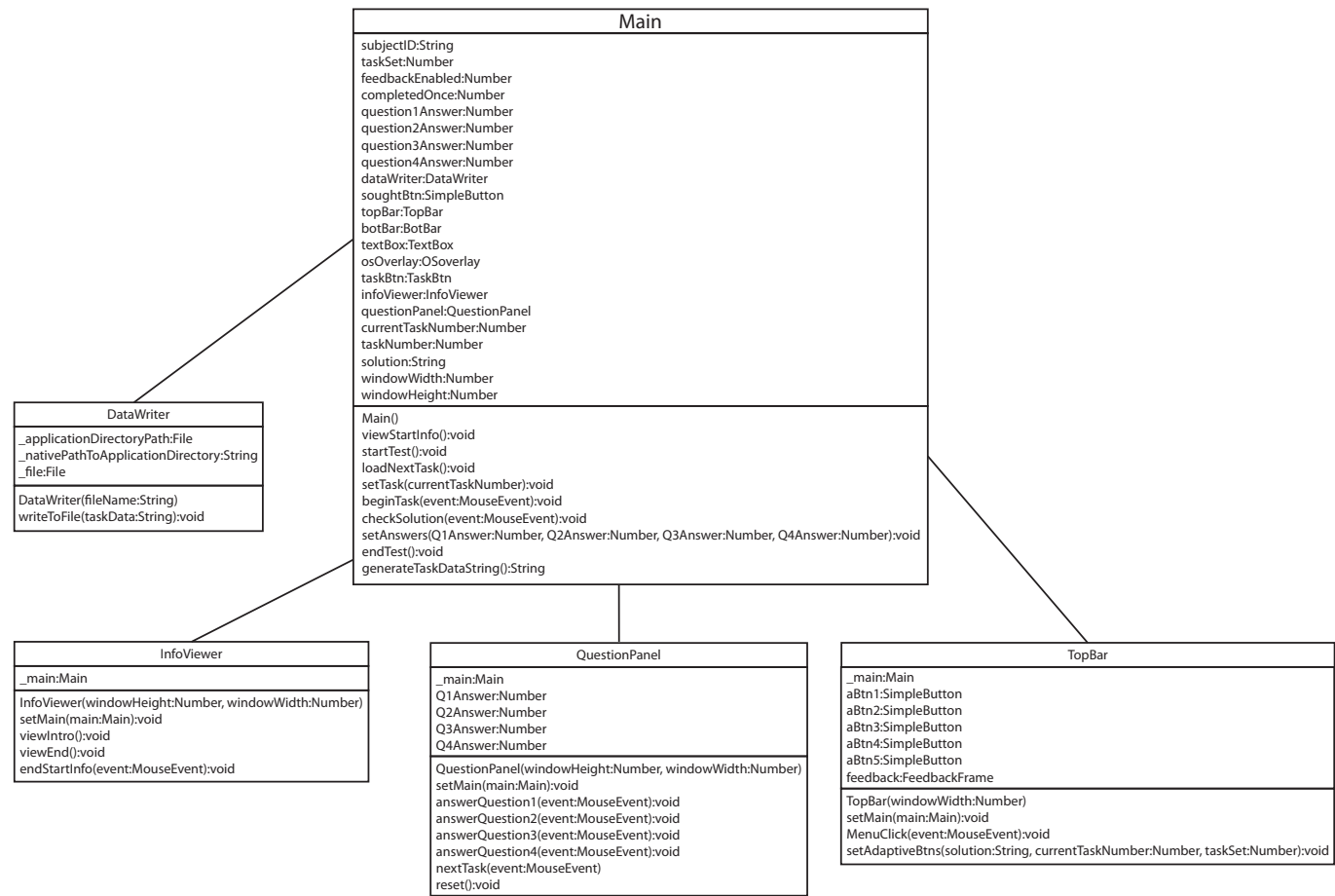


Figure 4.2: Class diagram of the ActionScript[®] 3.0 classes developed for the experiment.

4.4 Experiment software GUI

The software created for the experiment sought to simulate the visual aspects of some existing text editors, as also explained in Section 4.5 on page 29. The following presents a short description of each screen in the GUI.

Screen 1 - Introduction: This screen contains four elements, see Figure 4.3.

- Text box with the introduction to the experiment.
- Subject ID input box, the facilitator inputs a unique number for each test participant.
- Task set input boxes the facilitator defines the starting task set and whether feedback is enabled or not in these.
- Start button - this button links to Screen 2.

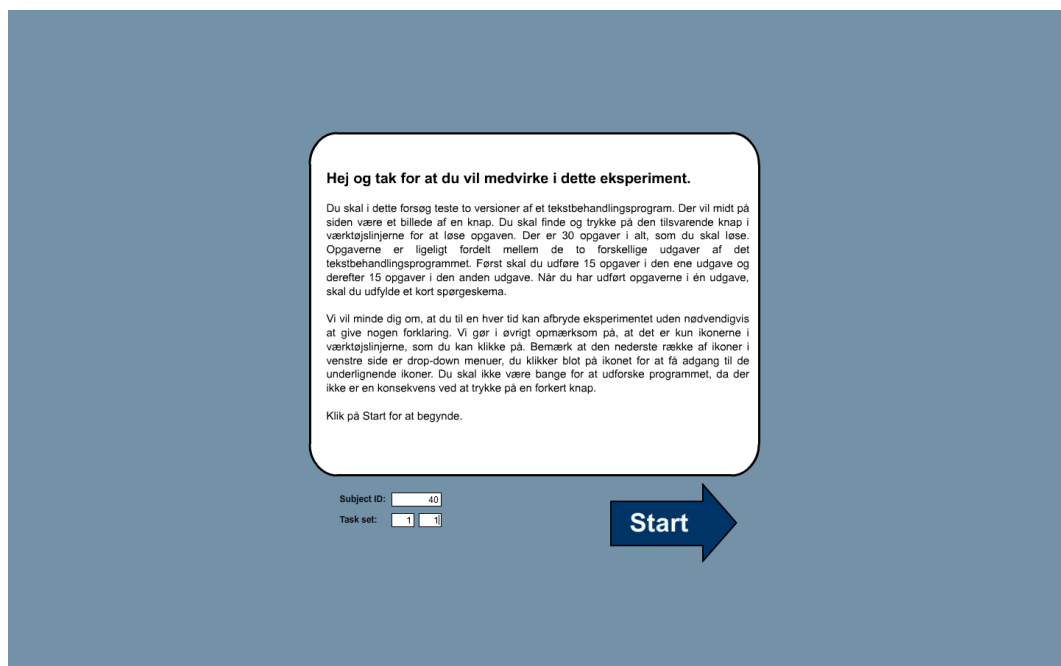


Figure 4.3: *Screen 1: Introduction.*

Screen 2 - Task definition This screen contains three elements, see Figure 4.4.

- The text defining the task.
- The button which the participants were tasked to find.
- Start button - this button links to Screen 3.

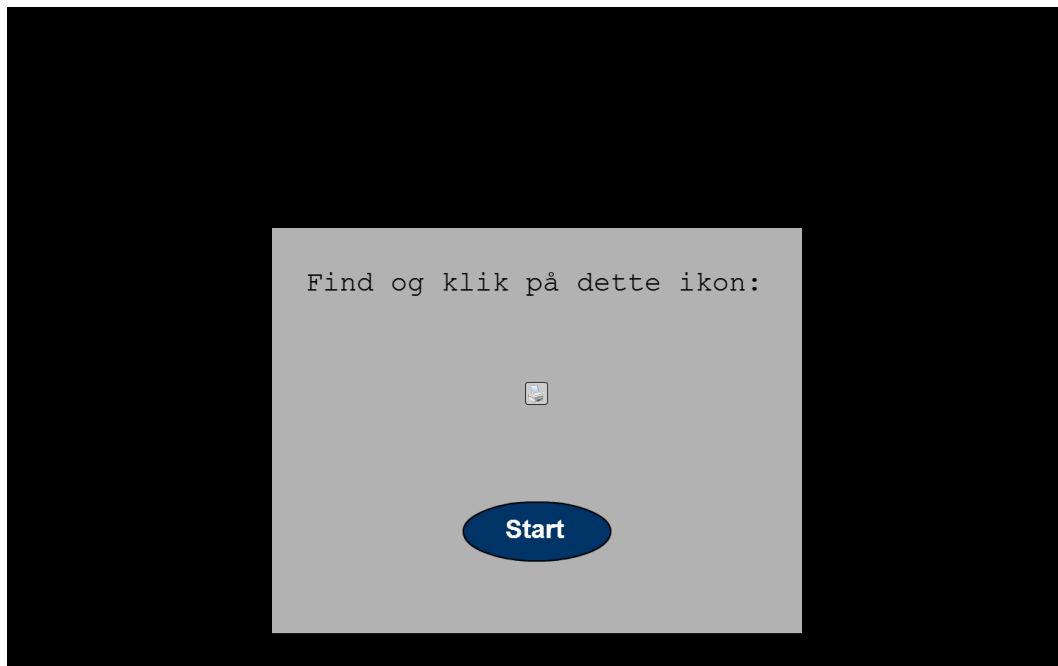


Figure 4.4: *Screen 2: Task definition.*

Screen 3 - Text editor: This screen contains three elements, see Figure 4.5.

- The button which the participants were tasked to find.
- The static toolbars in the top leftmost corner.
- The adaptive toolbar in the top rightmost corner.

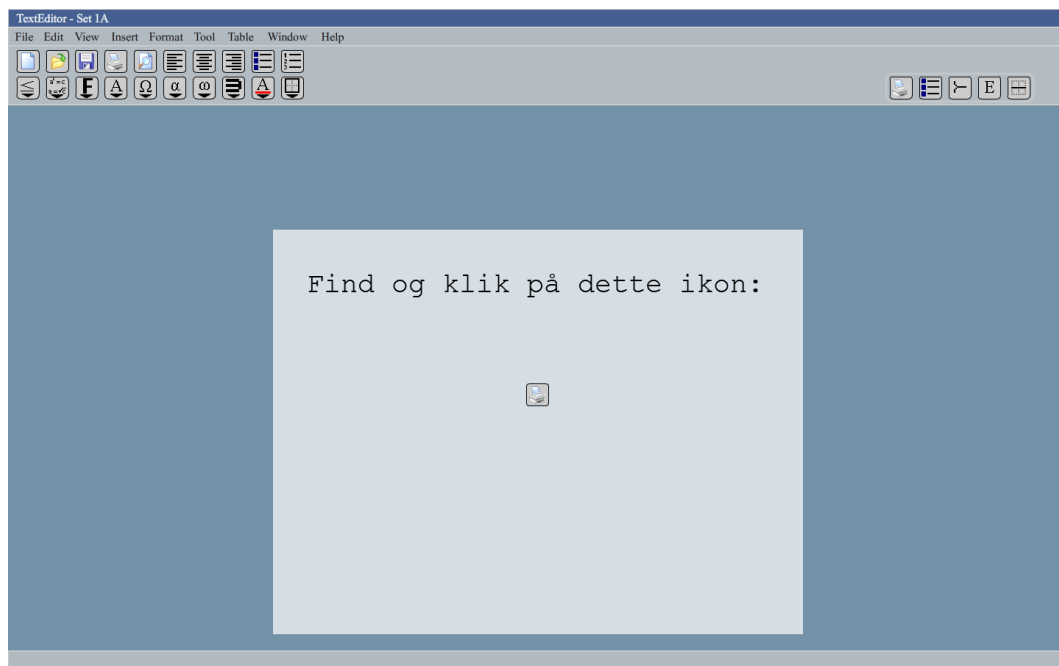


Figure 4.5: *Screen 3: Text editor.*

Screen 3 - Text editor continued: If feedback was enabled for the task in question, a pulsating glow would be present on the adaptive toolbar after a short delay, see Figure 4.6.

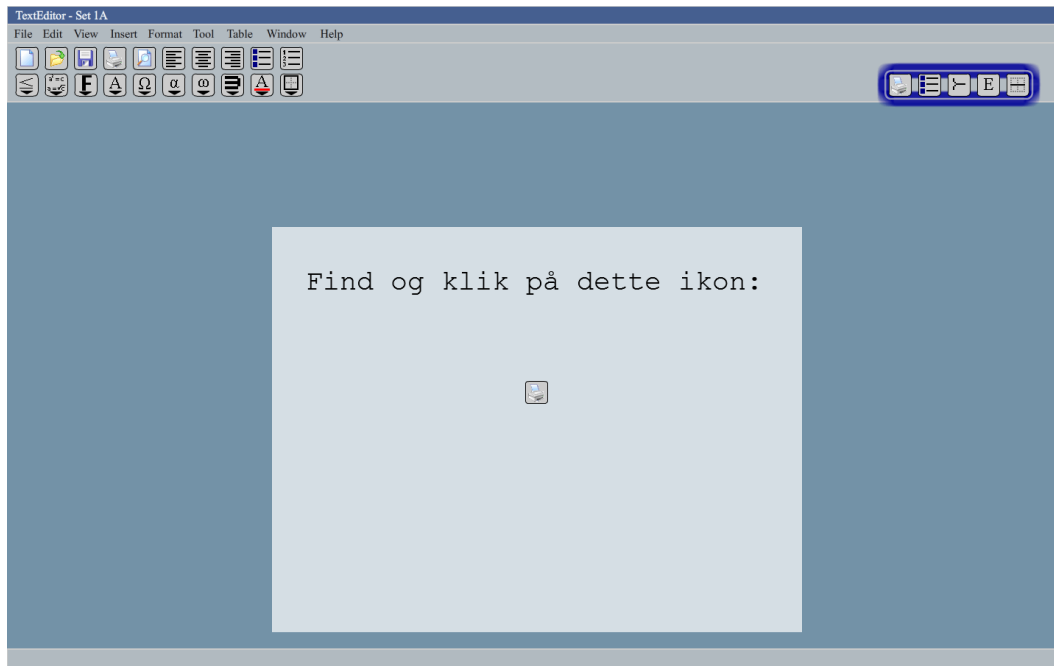


Figure 4.6: *Screen 3 continued: Text editor with feedback active.*

Screen 4 - Question panel screen: This screen contains one major element, see Figure 4.7.

- The question panel.

The screenshot shows a question panel with four questions, each followed by a seven-point Likert scale. The questions are in Danish and relate to the usability of an adaptive toolbar. The scales are labeled from 'Helt enig' (completely agree) to 'Helt uenig' (completely disagree), with intermediate points 'Enig', 'Lidt enig', 'Neutral', 'Lidt uenig', and 'Uenig'.

Question (Danish)	Helt enig	Enig	Lidt enig	Neutral	Lidt uenig	Uenig	Helt uenig
Den adaptive toolbar var nyttig. (Vælg ved at klikke på én af boksene.)	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
Jeg vidste på forhånd, når den søgte knap var i den adaptive toolbar. (Vælg ved at klikke på én af boksene.)	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
Den adaptive toolbar gjorde det mere effektivt at løse opgaven. (Vælg ved at klikke på én af boksene.)	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
Jeg følte, at jeg havde kontrol over det der skete på interface. (Vælg ved at klikke på én af boksene.)	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>

Figure 4.7: Screen 4: The question panel screen.

Screen 4 - Question panel screen continued: Upon answering the fourth question, a button would appear - allowing the participant to continue, see Figure 4.8.

The screenshot shows a questionnaire interface with four questions, each followed by a seven-point Likert scale. The scales are labeled from 'Helt enig' (completely agree) to 'Helt uenig' (completely disagree), with intermediate points 'Enig', 'Lidt enig', 'Neutral', 'Lidt uenig', and 'Uenig'. The first question is 'Den adaptive toolbar var nyttig.' with the first point selected. The second is 'Jeg vidste på forhånd, når den søgte knap var i den adaptive toolbar.' with the second point selected. The third is 'Den adaptive toolbar gjorde det mere effektivt at løse opgaven.' with the second point selected. The fourth is 'Jeg følte, at jeg havde kontrol over det der skete på interface.' with the first point selected. A blue arrow button labeled 'Fortsæt' is positioned below the fourth question.

Den adaptive toolbar var nyttig. (Vælg ved at klikke på én af boksene).	<input checked="" type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
	Helt enig	Enig	Lidt enig	Neutral	Lidt uenig	Uenig	Helt uenig
Jeg vidste på forhånd, når den søgte knap var i den adaptive toolbar. (Vælg ved at klikke på én af boksene).	<input type="checkbox"/>	<input checked="" type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
	Helt enig	Enig	Lidt enig	Neutral	Lidt uenig	Uenig	Helt uenig
Den adaptive toolbar gjorde det mere effektivt at løse opgaven. (Vælg ved at klikke på én af boksene).	<input type="checkbox"/>	<input checked="" type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
	Helt enig	Enig	Lidt enig	Neutral	Lidt uenig	Uenig	Helt uenig
Jeg følte, at jeg havde kontrol over det der skete på interface. (Vælg ved at klikke på én af boksene).	<input checked="" type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
	Helt enig	Enig	Lidt enig	Neutral	Lidt uenig	Uenig	Helt uenig

Fortsæt

Figure 4.8: Screen 4 continued: The question panel screen with the appearing button.

Screen 5 - Outro text: This screen contains one major element, see Figure 4.9.

- Outro text box.



Figure 4.9: Screen 6: The outro text box.

4.5 Method

A flow diagram for the second experiment can be seen in Figure 4.10 on the following page.

The participant was brought into a laboratory and placed at a computer. Then the required instructions were given and a declaration of consent was signed. The declaration of consent can be found in Figure 4.11 on page 31. Following this, further instructions were provided by the experiment software via the computer monitor. Having been properly introduced to the experiment the participant started the experiment. The text used to introduce the participant to the experiment can be found in Section 4.5.1 on page 31. After reading the introduction all odd numbered participants (as defined by their subject ID) completed the tasks with the user interface with feedback. After all tasks had been completed, the participant filled in a short questionnaire regarding the interface they had just used. The questionnaire can be found in Section 4.5.4 on page 33. Hereafter the tasks for the interface without feedback were completed and the same questionnaire was answered again. The tasks are described in Section 4.5.2 on page 32. For the even numbered participants the order was reversed. When all tasks in both user interfaces had been completed the participant was shown a debriefing text, thanking him for his participation and outlining the purpose of the

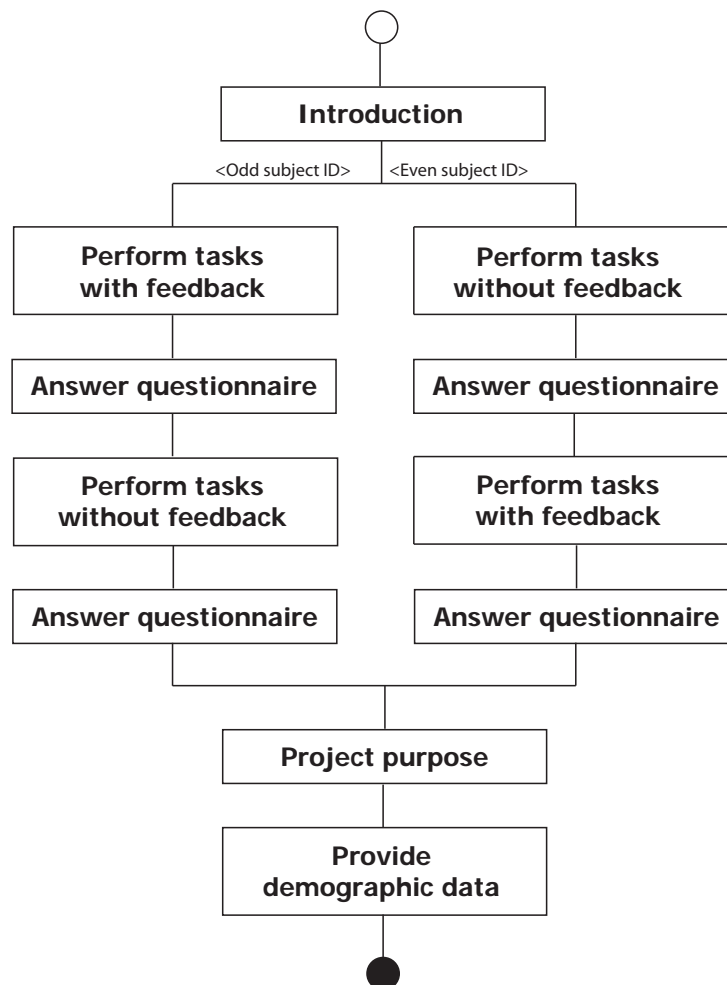


Figure 4.10: *The flow of the experiment.*

experiment. This text can be found in Section 4.5.5 on page 33. Lastly the participant was asked to fill in their name, age, and experience with Microsoft Word or similar applications on a piece of paper. The last question was answered via a five point scale ranging from novice to expert.

<p style="text-align: center;">Samtykkeerklæring Produkt- og designpsykologi Gruppe 1077</p> <p style="text-align: center;">Dato: ____/____-2011</p> <p>Jeg bekræfter hermed som forsøgsperson at:</p> <ul style="list-style-type: none">• Jeg har forstået den givne information og indvilliger i at medvirke i eksperimentet.• Jeg har forstået at jeg kan at forlade eksperimentet på hvilket som helst tidspunkt, uden yderligere forklaring.• En eye-gaze tracker bliver brugt til at observe hvilke områder af skærmen, jeg ønsker det.• Renskrevne citater og observationer fra eksperimentet og spørgeskemaer må blive brugt i den efterfølgende rapport.• Jeg giver tilladelse til at renskrevne citater, observationer og spørgeskemaer må videregives til vejleder og censorer i forbindelse med projektsamen og anvendes hertil. Ydermere at denne information må videregives til tredjepart efter skriftlig tilladelse fra projektgruppen. <p>Dit navn vil ikke kunne forbindes til specifikke resultater eller blive videregivet til hverken offentligheden, Aalborg Universitet og eventuel tredjeparts samarbejdspartnere.</p> <p style="text-align: center;">Forsøgspersonens navn: _____</p> <p style="text-align: center;">Underskrift: _____</p> <p style="text-align: center;">Forsøgsledernes navn: _____</p> <p style="text-align: center;">Underskrift: _____</p> <p style="text-align: center;"><i>Tak for din deltagelse!</i></p>
--

Figure 4.11: *The declaration of consent in Danish.*

4.5.1 Introduction text

Hej og tak for at du vil medvirke i dette eksperiment.

Du skal i dette forsøg teste to versioner af et tekstbehandlingsprogram. Der vil midt på siden være et billede af en knap. Du skal finde og trykke på den tilsvarende knap i værktøjslinjerne for at løse opgaven. Der er 30 opgaver i alt, som du skal

løse. Opgaverne er ligeligt fordelt mellem de to forskellige udgaver af det tekstbehandlingsprogrammet. Først skal du udføre 15 opgaver i den ene udgave og derefter 15 opgaver i den anden udgave. Når du har udført opgaverne i én udgave, skal du udfylde et kort spørgeskema.

Vi vil minde dig om, at du til en hver tid kan afbryde eksperimentet uden nødvendigvis at give nogen forklaring, som det også fremgår af sammentykkeerklæringen. Vi gør i øvrigt opmærksom på, at det er kun ikonerne i værktøjslinjerne, som du kan klikke på. Bemærk at den nederste række af ikoner i venstre side er drop-down menuer, du klikker blot på ikonet for at få adgang til de underlignende ikoner. Du skal ikke være bange for at udforske programmet, da der ikke er en konsekvens ved at trykke på en forkert knap.

Klik på Start for at begynde.

4.5.2 Tasks

The participant was asked to locate and click specific function buttons. These buttons were located on the adaptive toolbar in 66.6 % of the tasks (to simulate an algorithm with a 66.6 % accuracy). The buttons were always available in the static toolbars. An illustrations of the buttons in question were presented to the participant. Each task consisted of locating one function button. To simulate actual computer usage, the tasks varied in complexity. Some buttons would be directly available on the static toolbar, whereas others were placed in drop-down menus, which the participant had to navigate to, to complete the task. Two sets of tasks were created. Both sets followed a predefined pattern in regards to menu depth, drop-down menu size, and amount of tasks.

4.5.3 Measurements

To determine where the participant first sought the required function an eye-gaze tracker were employed. The adaptive and the static toolbars were both be defined as a Region of Interest (ROI), as were the task description area. The following measurements were performed:

ROI What area of the interface does the participant first seek the requested button. That is, does he look at the adaptive or static toolbars first.

Efficiency The time it took the participant to finish a task.

Clicks The total number of clicks that the participant performed for each task.

Toolbar used The toolbar used to solve the task.

4.5.4 Questionnaire

After all tasks in a test set for either the feedback or non-feedback version of the AUI had been completed, the participant was asked to rate the following statements on a seven point Likert scale.

1. The adaptive toolbar was useful.
2. I knew beforehand when the required button was in the adaptive toolbar.
3. The adaptive toolbar made solving the tasks more efficient.
4. I felt in control of what occurred on the interface.

The participants were asked to rate these statements to investigate the participants satisfaction ratings of an AUI with feedback compared to an AUI without feedback. The questions are excerpt from the satisfaction ratings defined by Gajos et al. [2008].

4.5.5 Outro text

Eksperimentet er nu færdigt.

Tak for din deltagelse.

Eksperimentets formål er at teste, hvorvidt det er muligt at forkorte tiden det tager, før en bruger begynder at anvende såkaldte “adaptive toolbars” - altså toolbars der tilpasser sig brugeren - ved hjælp af visuel feedback.

4.6 Materials

The experiment used a text editor application developed in Adobe® Flash® Creative Suite, 4th Edition specifically for this experiment. Instructions and data logging of user actions were handled by the test application. The application design was based on the graphical appearance of Microsoft® Word 2003, whereas the buttons and button grouping stem from TeXnicCenter. This was done to provide the participants with a familiar environment, but ensure that no participant could be an expert in the interface. The eye-gaze tracking was performed using a Tobii X120 eye-gaze tracker, running Tobii Studio version 1.7.3 software. All eye-gaze tracking data was handled by the Tobii software. All software was executed on a Fujitsu-Siemens Amilo Xi 2528 17” widescreen laptop (using a resolution of 1440x900 pixels). As a technician was required to monitor the state of the tracking on the laptop monitor, a Samsung SyncMaster 225BW 22” widescreen LCD monitor was connected to the laptop, and used to display the experiment software for the participant.

4.6.1 Environment

The experiment was conducted in a laboratory at Aalborg University. A sketch of the room can be seen on Figure 4.12. The participant only had a monitor and a mouse at his disposal. The eye-gaze tracker was placed directly beneath the monitor. The participant did not have a keyboard available as the task was to click a specific button with the mouse and not activate the function which the button represents, e.g. by using keyboard short-cuts. Having a keyboard would also make it more likely that the participant looked away from the monitor, thus causing the eye-gaze tracker to lose its tracking. The technician monitored the eye-gaze tracking status on the laptop monitor.

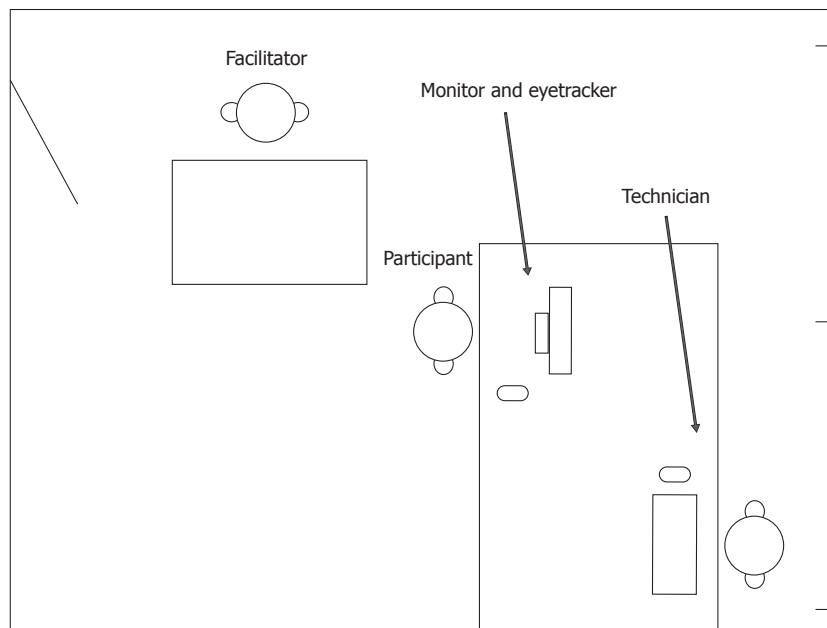


Figure 4.12: *A sketch of the test environment*

4.6.2 Pilot study

A pilot study was conducted prior to the actual experiment. The goal of the pilot study was two fold. Firstly to gather experience with the eye-gaze tracker as the facilitator and technician did not have any prior experience with this. Secondly to uncover any problems with the methodology of the experiment. Two participants partook in the pilot study. As a result of the pilot study, the second Likert statement was altered to clarify that the knowledge of whether the requested button was in the adaptive toolbar or not should be knowledge which the participant possessed prior to clicking the start button.

Results

All statistical tests in this analysis rejected the null-hypothesis at the 0.050 significance level. All data and script files can be found on Appendix 2 on the attached CD. A sample video of the eye-gaze data collected from one participant can be found on Appendix 3 on the attached CD. Only one video was included on the attached CD due to limited storage capacity.

5.1 Participants

Participants were recruited at the School of Information and Communication Technology at Aalborg University. All but one of the participants were students. A total of 20 participants (15 male, 5 female) participated in the experiment. The age span ranged from 20 to 33 years ($median = 24$, $sd = 0.55$). The participants were asked to rate their experience with text editors on a five point scale, where 1 represented novice user and 5 represented expert user. Unfortunately all participants rated themselves as three or above. The distribution can be seen in Table 5.1.

Experience level	1	2	3	4	5
Amount of participants	0	0	2	14	4

Table 5.1: *The participants' rating for their own text editor experience level. 1 = novice, 5 = expert.*

5.2 Toolbar used

FiXme Fatal: AG: Opdater alle prop.test og tilføj alternative For each task it was observed which toolbar the participant used to solve the task. Linear regression was employed to investigate which variables influenced what toolbar the participants used. As mentioned in Section 4.2.1 on page 19 the adaptive toolbar had a precision level of 66.6 %, thus in 10 out of the 30 tasks the requested button would not be available in the adaptive toolbar. These tasks were omitted in this analysis as it was only relevant to analyse the results where the participant actually had a choice as to which toolbar to use to solve the task. A general linear model was created, including all variables which it was assumed possibly had an influence on which toolbar the participant used to solve the task. These variables were:

FiXme Fatal: AG: O
alle prop.test og tilf
alternative

Feedback Whether or not feedback was provided during the specific task. Note that a special case exist for the 15th case where feedback actually was not provided even though it is analysed as such.

Menu depth Whether the required button was located in the a drop-down menu or in a top menu.

Task The task number.

Task set The task set which the current task was in.

Experience The participant's self evaluated experience with text editors.

Age The participant's age.

Gender The participant's gender.

ID A unique identifier for each participant.

Model reduction was performed to remove any variables which did not influence which toolbar the participant used to solve the task. The highest p-value variable was removed from the model and the model was re-evaluated until only variables with a significant influence on the toolbar which the participant used to solve the task remained. The resulting model included feedback, menu depth, task, task set and participant ID. The reduced model and the influencing variables can be seen in Table 5.2.

	Estimate	Standard error	p-value	
Feedback	1.135	0.023	0.002	**
Menu depth	2.119	0.274	< 0.001	***
Task	-0.065	0.030	0.028	*
Task set	1.890	0.563	< 0.001	***
ID	0.070	0.023	0.002	**

Table 5.2: *The variables which influenced what toolbar the participants used to solve the tasks.*

Linear regression does not provide any insight into how the variables influenced the response vector. Hence a more in-depth analysis was required. The ID vector was not investigated as it is of no interest to limit who uses adaptive interfaces. The exact effect of the task variable was not investigated either as task complexity varies in actual text editor usage just like it did in this experiment.

The mean frequency of using the adaptive toolbar based on the menu depth the button was located in can be seen in Figure 5.1 on the facing page. The plot indicated that the adaptive toolbar was used more often when the requested button was located in a drop-down menu, this trend is significant ($X^2 = 81.010$, $p < 0.001$).

The mean frequency of using the adaptive toolbar based on whether feedback was provided or not can be seen in Figure 5.6 on page 41. The plot indicated that the adaptive toolbar was used more often when feedback was provided, this trend is significant ($X^2 = 13.337$, $p < 0.001$).

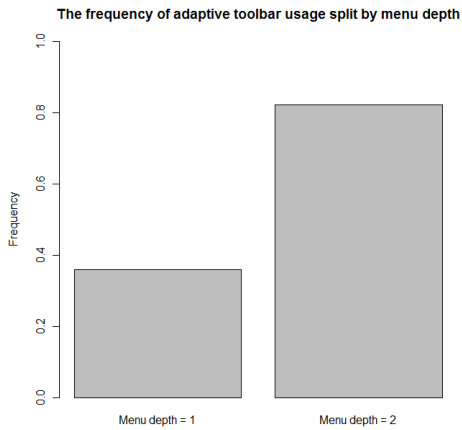


Figure 5.1: *Frequency of adaptive toolbar usage split by menu depth.*

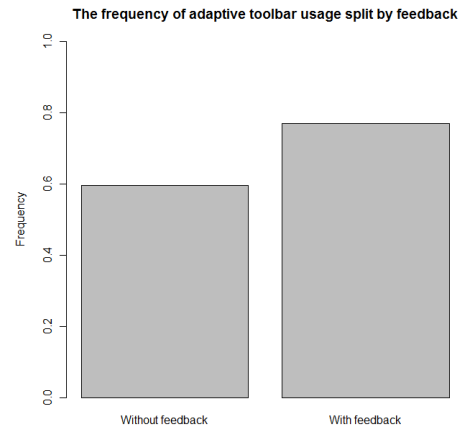


Figure 5.2: *Frequency of adaptive toolbar usage split by feedback.*

It has already been shown that users will use an adaptive adaptive toolbar if it is present[Gajos et al., 2006]. It is assumed that because no information was provided on how the five buttons (the adaptive toolbar) in the top right hand side of the interface, some learning was required before the participants understood the adaptive toolbar. To investigate feedback's effect on this learning curve the data was split into four groups. The four groups were:

Group 1: First ten tasks with feedback

Group 2: Last ten tasks without feedback

Group 3: First ten tasks without feedback

Group 4: Last ten tasks with feedback

The frequency of participants who used the adaptive toolbar was calculated for every tasks. Boxplots of the distribution of these frequencies for each group can be found in Figure 5.3 on the following page. The boxplots indicated that no difference exists between the groups, where participants were provided with feedback in the first ten tasks and were not provided with feedback in the last ten tasks. However, when participants were not provided with feedback in the first ten tasks, but were in the last ten tasks, a difference between the two conditions can be observed. To investigate this, the proportion of participants, which used the adaptive toolbar to solve the task for each group, was calculated. A within-subject test of the proportions equality was performed. The result of this test can be found in Table 5.3 on the next page. As it is seen, no difference exist between the feedback and no feedback condition, when participants were provided with feedback in the first half of the experiment. However, when participants were not provided with feedback in the first half of the experiment they used the adaptive toolbar more often in the last half of the experiment, when feedback was provided.

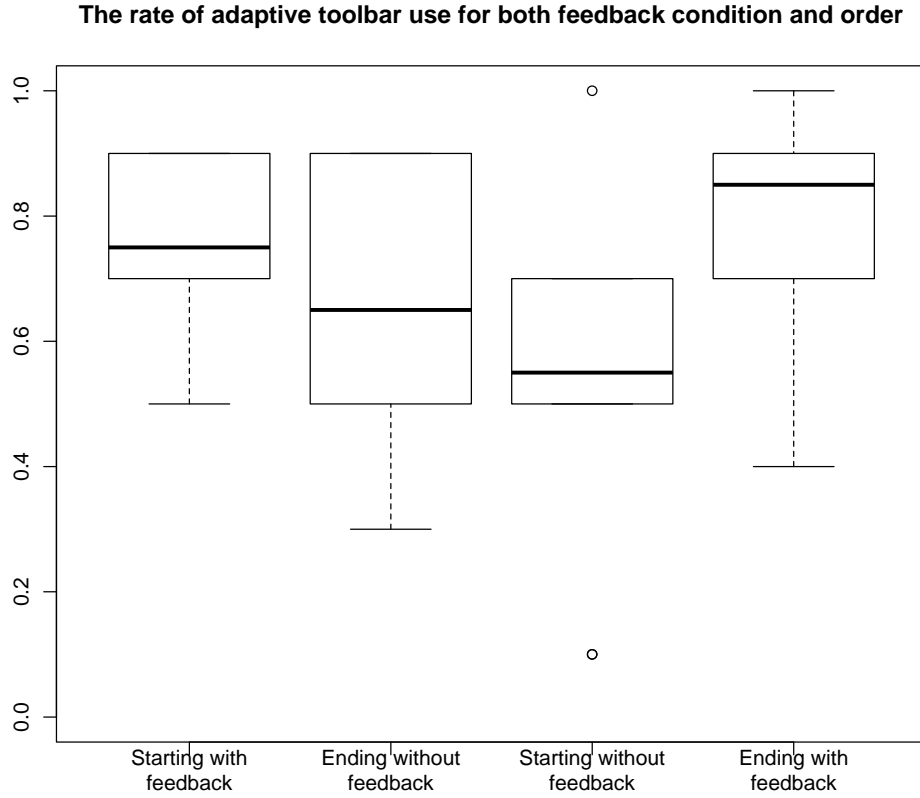


Figure 5.3: Frequency of adaptive toolbar used split by whether feedback was provided or not and what half of the experiment it was provided in.

	X^2	df	p-value
Starting with feedback	1.929	1	0.0825
Starting without feedback	12.928	1	< 0.001 ***

Table 5.3: The difference in proportions for usage of the adaptive toolbar between the first ten and last ten tasks, split between those that were provided with feedback in the first ten task and those that were provided with feedback in the last ten tasks.

5.3 Eye-gaze tracking results

The toolbar, where the participant first sought the requested button was observed for all tasks using an eye-gaze tracker. Unfortunately the eye-gaze tracker was not able to track the participants' eyes at all times. Thus, some data were non-existent. A total of 25 measurements out of 600 measurements were not available. All NAs (Not Availables) were omitted from the data prior to any statistical evaluation. A general linear model was defined, which included all variables which it was assumed possible had an influence on which part of the user interface the participant gazed upon first. These variables were:

Feedback Whether or not feedback was provided during the specific task. Note that a special case exist for the 15th case where feedback actually was not provided even though it is analysed as such.

Task The task number.

Task set The task set which the current task was in.

Experience The participant's self evaluated experience with text editors.

Age The participant's age.

Gender The participant's gender.

ID A unique identifier for each participant.

Model reduction was performed to remove any variables which did not influence which toolbar the participant used to solve the task. The highest p-value variable was removed from the model and the model was re-evaluated until only variables with a significant influence on the toolbar which the participant used to solve the task remained. The resulting model included feedback, task, and task set. The reduced model and the influencing variables can be seen in Table 5.4

	Estimate	Standard error	p-value	
Feedback	0.380	0.179	0.034	*
Task	0.049	0.011	< 0.001	***
Task set	0.756	0.179	< 0.001	***

Table 5.4: *A model of which variables influenced which toolbar participants gazed upon first.*

As the NAs were not distributed evenly among the tasks all data was normalised. For every task the amount of times that the participants fixated on the adaptive toolbar first was divided by the total amount of observation.

To investigate how feedback affected where the participants looked first further analysis was performed. The mean frequency of fixating on the adaptive toolbar first based on whether feedback was provided or not can be seen in Figure 5.5 on page 41. This difference of frequency was found to be significant ($X^2 = 3.532$, $p = 0.030$). Further analysis was performed to investigate if any change occurred when feedback was either added or removed

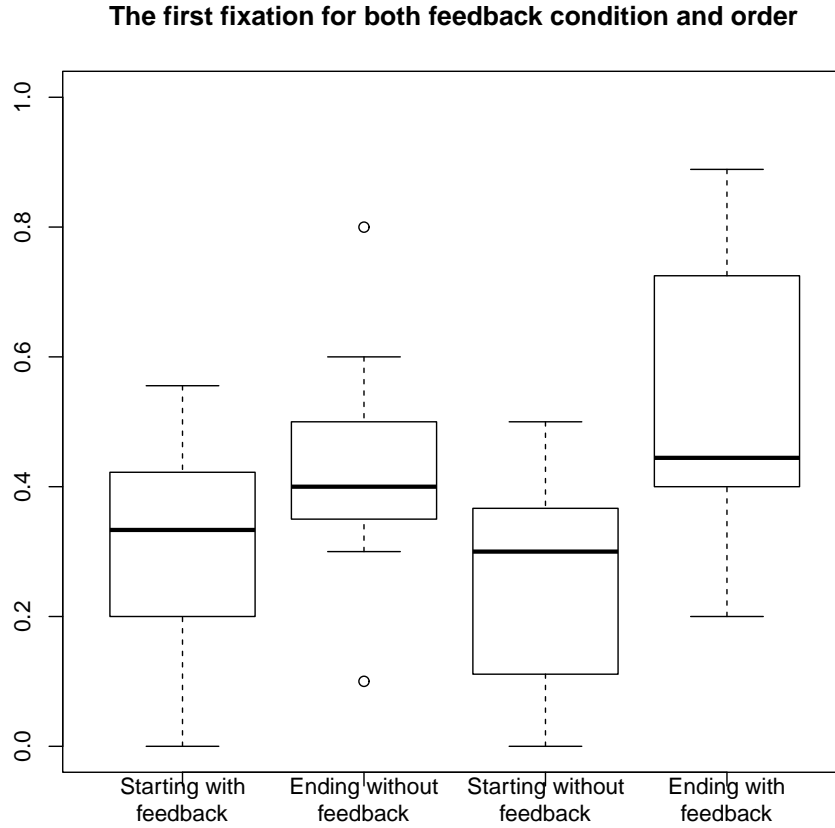


Figure 5.4: *Frequency of adaptive toolbar usage split by feedback.*

based on the starting condition. To perform this investigation the data was split into four groups. The four groups were:

Group 1: First fifteen tasks with feedback

Group 2: Last fifteen tasks without feedback

Group 3: First fifteen tasks without feedback

Group 4: Last fifteen tasks with feedback

The frequency of participants who looked at the adaptive toolbar first were calculated for each group. Boxplots of the distribution of these frequencies for each group can be found in Figure 5.4. The boxplots indicated that participants first sought the requested button in the adaptive toolbar more often in the latter half of the experiment. To investigate this observation, the proportion of participants which fixated on the adaptive toolbar first was calculated for each group. A within-subject test of the proportions equality was performed. The result of this test can be found in Table 5.5 on the next page. As it is seen, the proportion of tasks where the participants first sought the requested button in the adaptive toolbar was large in the latter half of the experiment regardless of the feedback order.

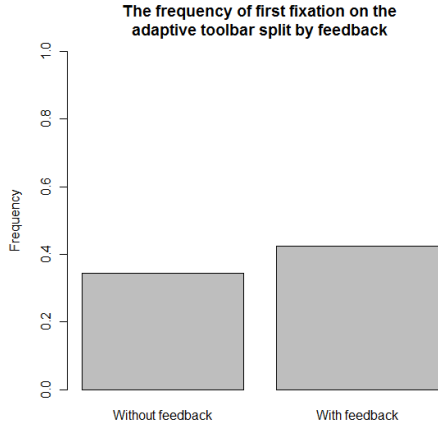


Figure 5.5: Frequency of participants' who's first fixation was on the adaptive toolbar split by feedback.

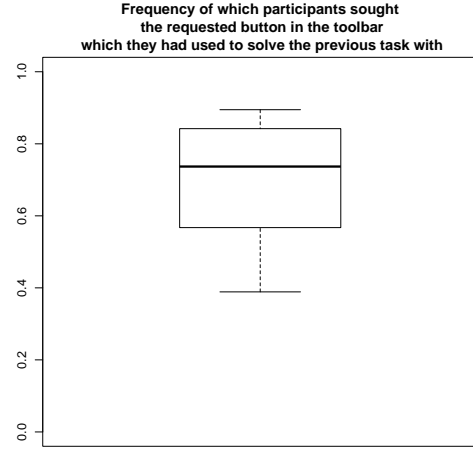


Figure 5.6: Frequency of participants who's first fixation was on the toolbar they used to solve the previous task.

	X^2	df	p-value	
Starting with feedback	5.372	1	0.010	*
Starting without feedback	24.324	1	< 0.001	***

Table 5.5: The difference in proportions for first seeking the requested button in the adaptive toolbar based on feedback condition and feedback order.

It was investigated whether the toolbar, which the participant had just used to solve the previous task, was the one they would look at first. Task 1 was removed from this analysis as there were no previous task to compare it to. The result can be found in Figure 5.6. In 70.1 % of the trials the toolbar they had used to solved the previous task was the one that the participant gazed on first. This is significantly above random ($X^2 = 77.864$, $p < 0.001$).

5.4 Task Duration

The duration of each individual task for all participants was recorded. Linear regression was employed to investigate what variables influenced the task duration. A linear model of task duration as the response was defined, which included all variables which were thought possibly influenced the response. These variables were:

Menu depth Whether the required button was located in the a drop-down menu or in a top menu.

Feedback Whether or not feedback were provided during the specific task. Note that a special case exist for the 15th case where feedback actually was not provided even though it is analysed as such.

Fixation Count The amount of fixations the participant had for that particular task.

Task The task number.

Task set The task set which the current task was in.

Experience The participant's self evaluated experience with text editors.

Age The participant's age.

Gender The participant's gender.

ID A unique identifier for each participant.

Model reduction was conducted to identify the variables influencing the model's response - the task duration. For each iteration, the highest p-value variable was removed from the model and the model was evaluated again. This was repeated until only variables with a significant influence on the response remained. The final model included menu depth, fixation count, task, age, gender and experience. The final model and the degree of influence of the different variables can be seen in Table 5.6.

	Estimate	Standard error	p-value	
Menu depth	0.488	0.861	0.007	**
Fixation count	0.297	0.004	< 0.001	***
Task	-0.0197	0.009	0.029	*
Age	0.072	0.023	0.002	**
Gender	-0.983	0.190	< 0.000	***
Experience	0.537	0.153	< 0.000	***

Table 5.6: *A model of which variables influenced the task duration.*

That menu depth had an influence on task duration was not surprising, as it requires more time to access items in a drop-down menu compared to locating it in the top menu. Fixation count also has a significant influence on task duration. This is not surprising either, as it requires more time to have more fixations - and a person will naturally have more fixations the longer they spent solving a task. Task is also shown to have an influence. As the tasks can be of either menu depth, this is most likely just a result hereof. Age and gender appears also to have an influence on the task duration. This possible correlation is of no immediate interest for this study. The participant's experience also influenced the task duration. This is not unexpected, as the test application was based on the looks of a typical text editor program. This also included the position of some of the function buttons, which they participants had to locate to solve the tasks. As a result hereof, previous experience is transferable for some tasks, which would decrease the time it took to solve that task.

5.5 Questionnaire analysis

After the participants had completed a task set, they were required to answer four questions on 7-step likert scales. The questions were (freely translated from danish):

1. The adaptive toolbar was useful.
2. I knew beforehand when the required button was in the adaptive toolbar.
3. The adaptive toolbar made solving the tasks more efficient.

4. I felt in control of what occurred in the interface.

The results of the these question - divided into groups based on whether they had feedback or not for the relevant task set - can be seen in Figure 5.7, 5.8, 5.9 and Figure 5.10 on page 45.

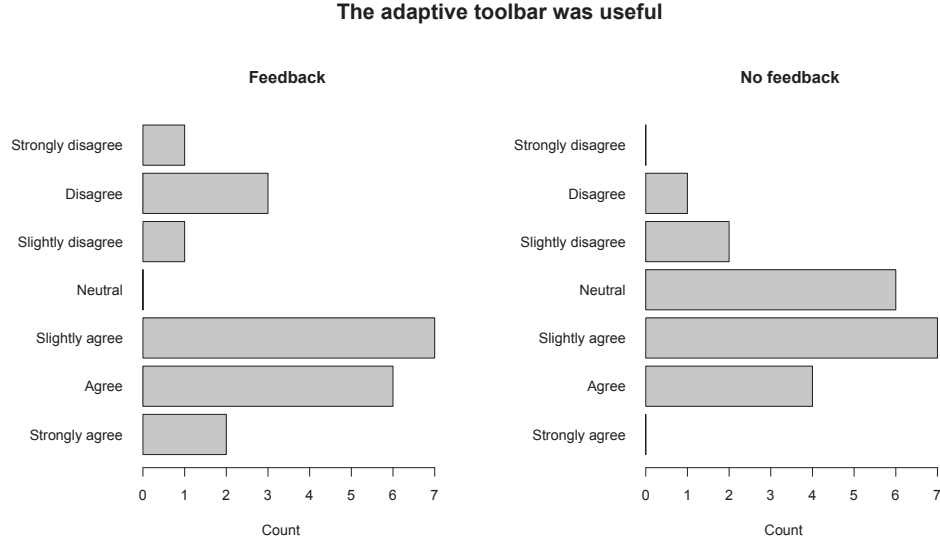


Figure 5.7: *Likert rating for Question 1.*

The mean, standard deviation, and median for each question can be seen in Table 5.7. The answers have been grouped by a feedback/no feedback condition, i.e. whether the participant had solved the task set with or without feedback enabled.

	Mean	Standard deviation	Median
Question 1 with feedback	3.250	1.773	3
Question 1 without feedback	3.450	1.099	3
Question 2 with feedback	4.100	1.714	4.5
Question 2 without feedback	4.350	1.531	4.5
Question 3 with feedback	3.050	1.538	2.5
Question 3 without feedback	3.600	1.188	4
Question 4 with feedback	3.350	1.599	4
Question 4 without feedback	3.150	1.461	3

Table 5.7: *The responses to the questions asked in the experiment, grouped according to the feedback condition. (Note that some medians are odd numbers, which is caused by an even amount of responses).*

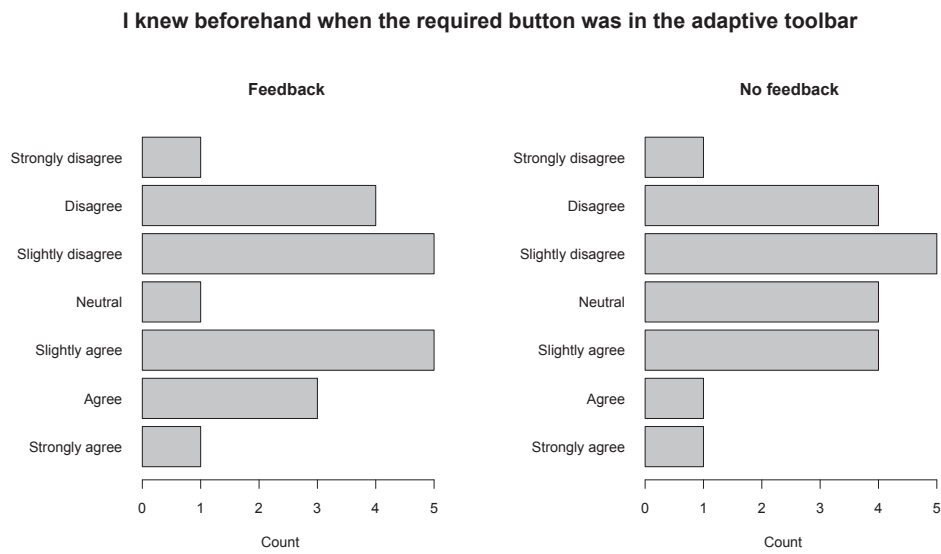


Figure 5.8: *Likert rating for Question 2.*

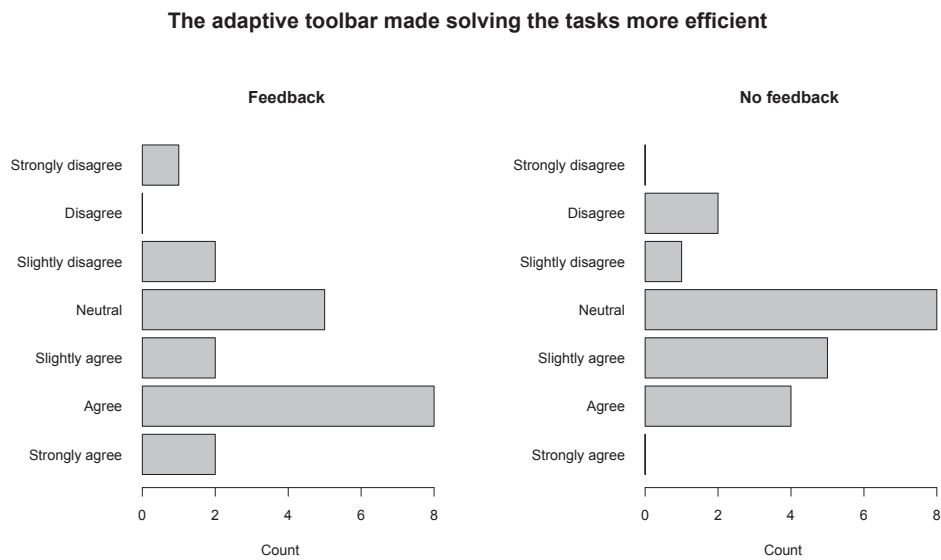


Figure 5.9: *Likert rating for Question 3.*

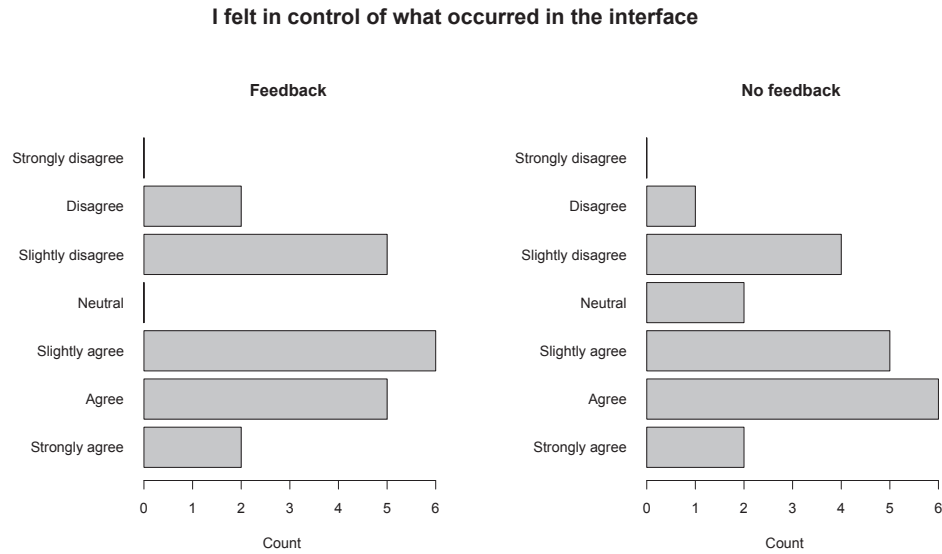


Figure 5.10: *Likert rating for Question 4.*

5.6 Summary

This section will provide a quick overview of the results of the experiment.

Toolbar used

A linear model showed that feedback, menu depth, task, task set, and ID all had significant influence on whether the participant used the adaptive toolbar to solve the task or not. Further analysis showed that the participants used the adaptive toolbar significantly more when locating an icon in a drop-down menu compared to icons located in the top level of the interface. It was also found that the participants used the adaptive toolbar significantly more when provided with feedback compared to the no feedback condition. Upon dividing the participants into four groups - based on feedback/no-feedback and the order in which this was given - a within-subject test of the proportions equality was performed. This showed that participants did not use the adaptive toolbar significantly more when provided with feedback in the beginning. However, the participants whom started without feedback used the adaptive toolbar significantly more when provided with feedback.

Eye-gaze tracking

A linear model showed that feedback, task, and task set all had significant influence on what toolbar the participant first sought the requested button. The mean frequency of fixating on the adaptive toolbar first was investigated in response to the feedback/no-feedback condition.

This difference was found to not significant. Upon dividing the participants into four groups (starting with feedback, ending without feedback, starting without feedback, and ending with feedback) a within-subject test of the proportions equality was performed. The test of the proportions equality showed that participants did not have their first gaze significantly more on the adaptive toolbar when provided with feedback in the first half of the experiment. Instead, participants were found to have their first gaze significantly more on the adaptive toolbar when the first half of the experiment was without feedback. Also, it was investigated whether the toolbar, which the participant used to solve the previous task, had any influence on which toolbar they first sought the next requested button in. This showed that in 71.2 % of all trials, participant would have their first fixation in the toolbar they had just used to solve a task. This frequency was significantly above random.

Task Duration

A linear model showed that menu depth, fixation counts, task, age, gender, and experience set all had significant influence on the duration of a task. None of these correlations were explored further, as they were of no immediate interest to this study.

Questionnaire analysis

The answers to the 7-step Likert questions in the questionnaire were divided into groups based on a feedback/no-feedback condition. Plots revealed no distinct trends in the ratings as a response to feedback. The mean, standard deviation, and median were also observed for all four question based on the feedback/no-feedback condition. This did not reveal any observable trend either. As the questionnaires were primarily included to possibly provide hindsight to questionable observations or results, no further analysis was performed on these.

Discussion

The discussion can be found in the article.

Appendices on the attached CD

- Appendix 1 on the attached CD contains the Flash Software and an animation viewer.
- Appendix 2 on the attached CD contains the data from the experiment and the R script used to analyse the data.
- Appendix 3 on the attached CD contains a video sample of eye-gaze tracking of a participant.
- Appendix 4 on the attached CD contains a digital copy of the article.
- Appendix 5 on the attached CD contains a digital copy of these worksheets.

Bibliography

- Benyon, D. [1993a], ‘Accommodating individual differences through an adaptive user interface’, *HUMAN FACTORS IN INFORMATION TECHNOLOGY* **10**, 149–149.
- Benyon, D. [1993b], ‘Adaptive systems: a solution to usability problems’, *User Modeling and User-Adapted Interaction* **3**(1), 65–87.
- Castagnos, S. and Pu, P. [2010], Consumer decision patterns through eye gaze analysis, in ‘Workshop on Eye Gaze in Intelligent Human Machine Interaction (IUI 2010), Hong Kong, China’.
- Conati, C., Merten, C., Amershi, S. and Muldner, K. [2007], Using eye-tracking data for high-level user modeling in adaptive interfaces, in ‘PROCEEDINGS OF THE NATIONAL CONFERENCE ON ARTIFICIAL INTELLIGENCE’, Vol. 22, Menlo Park, CA; Cambridge, MA; London; AAAI Press; MIT Press; 1999, p. 1614.
- Dreyfus, H., Dreyfus, S. and Athanasiou, T. [1988], *Mind over machine*, Thinking Allowed Productions.
- Findlater, L. and McGrenere, J. [2008], Impact of screen size on performance, awareness, and user satisfaction with adaptive graphical user interfaces, in ‘Proceeding of the twenty-sixth annual SIGCHI conference on Human factors in computing systems’, ACM, pp. 1247–1256.
- Gajos, K., Czerwinski, M., Tan, D. and Weld, D. [2006], Exploring the design space for adaptive graphical user interfaces, in ‘Proceedings of the working conference on Advanced visual interfaces’, ACM, pp. 201–208.
- Gajos, K., Everitt, K., Tan, D., Czerwinski, M. and Weld, D. [2008], Predictability and accuracy in adaptive user interfaces, in ‘Proceeding of the twenty-sixth annual SIGCHI conference on Human factors in computing systems’, ACM, pp. 1271–1274.
- Goldberg, J., Stimson, M., Lewenstein, M., Scott, N. and Wichansky, A. [2002], Eye tracking in web search tasks: design implications, in ‘Proceedings of the 2002 symposium on Eye tracking research & applications’, ACM, pp. 51–58.
- Kolb, D. et al. [1984], *Experiential learning: Experience as the source of learning and development*, Prentice-Hall Englewood Cliffs, NJ.
- Kortum, P. [2008], *HCI beyond the GUI: design for haptic, speech, olfactory and other nontraditional interfaces*, Morgan Kaufmann.
- Liu, J., Wong, C. and Hui, K. [2003], ‘An adaptive user interface based on personalized learning’, *IEEE Intelligent Systems*, pp. 52–57.
- Miller, C. and Hannen, M. [1998], User acceptance of an intelligent user interface: a Ro-

- torcraft Pilot’s Associate example, *in* ‘Proceedings of the 4th international conference on Intelligent user interfaces’, ACM, pp. 109–116.
- Nazemi, K., Stab, C. and Fellner, D. [2010], ‘Interaction Analysis for Adaptive User Interfaces’, *Advanced Intelligent Computing Theories and Applications* pp. 362–371.
- Paymans, T., Lindenberg, J. and Neerincx, M. [2004], Usability trade-offs for adaptive user interfaces: ease of use and learnability, *in* ‘Proceedings of the 9th international conference on Intelligent user interfaces’, ACM, pp. 301–303.
- Pernice, K. and Nielsen, J. [2009], ‘Eyetracking Methodology—How to Conduct and Evaluate Usability Studies Using Eye Tracking’.
- Pretorius, M., Calitz, A. and van Greunen, D. [2005], The added value of eye tracking in the usability evaluation of a network management tool, *in* ‘Proceedings of the 2005 annual research conference of the South African institute of computer scientists and information technologists on IT research in developing countries’, South African Institute for Computer Scientists and Information Technologists, pp. 1–10.
- The American Heritage® Stedman’s Medical Dictionary* [2011].
URL: <http://dictionary.reference.com/browse/feedback>
- Thomas, C. [1993], ‘Design, implementation and evaluation of an adaptive user interface’, *Knowledge-Based Systems* **6**(4), 230–238.
- Tsandilas, T. and Schraefel, M. [2005], An empirical assessment of adaptation techniques, *in* ‘CHI’05 extended abstracts on Human factors in computing systems’, ACM, pp. 2009–2012.
- University of Strathclyde [2011], *Computing Terms*, <http://www.strath.ac.uk/its/helpdesk/glossary/>. Last seen: April 14th, 2011.
- Wensveen, S., Djajadiningrat, J. and Overbeeke, C. [2004], Interaction frogger: a design framework to couple action and function through feedback and feedforward, *in* ‘Proceedings of the 5th conference on Designing interactive systems: processes, practices, methods, and techniques’, ACM, pp. 177–184.

Part I

Appendices

Literary review notes

Read by: Brian

Article: Gajos 2008 - Predictability and Accuracy in Adaptive User Interfaces

Relevance: High

Resume: An experiment seek to test which has the highest impact on user likeness and precision: accuracy or predictability. Accuracy is shown to have the highest impact on user performance, utilization and some satisfaction ratings than the improvement in predictability

Read by: Brian

Article: Paymans 2004 - Usability Trade-offs for Adaptive User Interfaces: Ease of Use and Learnability

Relevance: Medium-High

Resume: Paymans stipulate that unpredictable behavior from an adaptive interface reduce a systems usability. To compensate for this, Paymans seek to develop user's mental model of how the adaption functions. They find that the user support increases ease of use, but unexpectedly also reduces learnability. As such, the increase in ease of use can be achieved without improving user's mental model.

Read by: Brian

Article: Kühme - A user-centered approach to adaptive interfaces

Relevance: Low-Medium

Resume: Kühme suggests a CAA (Computer-Aided Adaption) system, which enables users to express their needs and preferences when adapting a interface - i.e., it is a tool for controlling automatic adaption. The paper describes a CAA environment, which is currently being developed.

Read by: Brian

Article: Nazemi 2010 - Interaction analyze for adaptive user interfaces

Relevance: High

Resume: Nazemi develops a prediction algorithm to predict user interaction, recognises user activities and detect user preferences on several abstract levels. The developed algorithm appears to outperform exiting algorithms with this same purpose in mind.

Read by: Brian

Article: Billsus 2002 - Adaptive interfaces for ubiquitous web access

Relevance: None

Resume: The article discusses the possibilities provided by having mobile devices access any information at any time from any location.

Read by: Brian

Article: Blumendorf 2010 - Bridging Models and Systems at Runtime to Build Adaptive User Interfaces

Relevance: Medium - depends on focus

Resume: Blumendorf describes several system, which creates a feedback loop between the runtime model and the system. This makes it possible to build systems that are aware of context information as well as of their own state by building internal models that the systems can reason on. The relevance of this article depends highly on the focus of the project.

Read by: Brian

Article: Findlater 2008 - Impact of Screen Size on Performance, Awareness, and User Satisfaction With Adaptive Graphical User Interfaces

Relevance: Medium - primarily the reference list

Resume: Findlater seeks to test previous finding in the field of adaptive user interface on small screen devices.

Read by: Brian

Article: Gajos 2006 - Exploring the design space for adaptive graphical user interfaces

Relevance: High

Resume: Gajos designs and implements 3 adaptive user interfaces and tests these against a “baseline” UI without adaption in two experiments. In conclusion, they synthesize the results with previous work and discuss how different design choices and interactions affect the success of adaptive graphical user interfaces.

Read by: Brian

Article: Grundy 2002 - An environment for developing adaptive, multi-device user interfaces

Relevance: Low.

Resume: Grundy seeks to design and implement adaptive user interfaces which may run on and adapt to the characteristics of multiple display devices and networks - as well as different users and tasks.

Read by: Brian

Article: Jason 2010 - The Evaluation of an Adaptive User Interface Model

Relevance: Medium

Resume: This article seeks to develop and test a adaptive user interface model in a specific case study, to test any improvement in performance as a response to the individual adaption from the user interface.

Read by: Brian

Article: Kumar 2006 - An Approach to Adaptive User Interfaces using Interactive Media Systems

Relevance: Low

Resume: Kumar suggests 3 different adaptive systems - a system for creating interactive media, adapter chain for bringing about a user interface depending on user preference, terminal capabilities, and network constraints and a Interactive Player. No user tests are conducted on any of the suggested systems.

Read by: Anders

Article: Liu 2003 - An Adaptive User Interface Based on Personalized Learning

Relevance: Medium

Resume: A description of the development and implementation of an episodes identification and association method for AUIs. A pilot study indicates higher productivity during typing tasks.

Experiment:

1.
 - Typing task
 - Participants provided priors (old documents)
 - Qualitative evaluation (questionnaire)
 -
 2.
 - 2 groups of 7 one with, one without phrase completion
 - Probably with same priors as the first experiment
-

Read by: Anders

Article: Miller 1999 - User Acceptance of an Intelligent User Interface A Rotorcraft Pilot's Associate Example

Relevance: Low

Resume: Investigates pilots acceptances of an intelligent Cockpit Information Manager. Pilot statements show that the pilots finds the intelligent system more usable, but the data has not been analysed.

Experiment:

- Full/part mission simulations
- Defined objectives, tactical decisions by participants (pilots)
- Simulation and classroom training
- 14 part missions (20-50 minutes)
- 4 full missions (1-1 $\frac{1}{2}$ hours)

- Within subject
 - Quantitative and qualitative data collected
-

Read by: Anders

Article: Benyon 1993 - Adaptive systems a solution to usability problems

Relevance: Medium

Resume: A frame work for adaptive user interfaces grounded in usability. Has some good considerations on how to analyse an existing system in order to make a more user friendly system (section 6). Includes an extensive example of how the model can be used (section 7).

Experiment: None

Read by: Anders

Article: Benyon 1993 - Accommodating individual differences through an adaptive user interface

Relevance: Low to medium

Resume: A lot of background information on different psychological evaluations. Usage requires us to somehow evaluate the psychological traits described below of our subjects.

Experiment: Then reports on an experiment where people were tested for six various psychological traits, e.g. spatial ability and field dependence. The subjects were then assigned into two groups depending on the score. They were asked to solve 12 tasks in 6 different interfaces. Time measured as a measure of performance.

Read by: Anders

Article: Eisenstein 2001 - Applying model-based techniques to the development of UIs for mobile computers

Relevance: Low

Resume: Discusses and presents ideas for a model to be used for all mobile devices user interface. Might have some ideas for a specific user interface if we elect to design from scratch instead of building AUI on top of something known.

Experiment: None

Read by: Anders

Article: Paramythis 2001 - A modular approach to the evaluation of adaptive user interfaces

Relevance: N/A

Resume: Present some method to evaluate AUIs. Will read when we get closer to experiment

Read by: Anders

Article: Thomas 1993 - Design, implementation and evaluation of an adaptive user interface

Relevance: Medium

Resume: Creates an adaptive interface in Excel, calls it Flexcel. Though outdated it holds some valid arguments regarding how an adaptive system should behave.

Experiment: Method not reported, can be found in: Oppermann 1994 - Adaptively supported adaptability

Read by: Anders

Article: Vainio 2003 - A critical approach to an adaptive user interface design

Relevance: Medium

Resume: Holds several usable definitions and a list of good considerations regarding the design of AUIs.

Experiment: None

Read by: Anders

Article: Cha 2006 - Learning styles diagnosis based on user interface behaviors for the customization of learning interfaces in an intelligent tutoring system

Relevance: None

Resume: Designs an AUI algorithm that learns the user's preferred learning style for the purpose of improving learning efficiency.

Experiment: Tests their algorithm, no usability evaluation is performed.

Read by: Anders

Article: Pretorius 2005 - The added value of eye tracking in the usability evaluation of a network management tool

Relevance: Medium (Method)

Resume: Creates a non-adaptive user interface (network management (NM) tool). The tool was evaluated in a user study with an eye-tracker. Holds information on user studies with eye-trackers.

Experiment: Participants solved normal NM tasks. During the evaluation the following measurements were collected:

1. Usability
 - (a) Effectiveness
 - Task completion rate
 - Task completed with/without assistance
 - Error rate recovery
 - (b) Efficiency
 - Task completion time
 - Real-time events (Mouse clicks etc.)
 - (c) Satisfaction (Questionnaire)
2. Eyetracking
 - Number of fixation
 - Fixation duration
 - Number of fixations on each ROI
 - Number of gaze on each ROI
 - Scanpath
 - Time to the 1st fixation on target ROI

Read by: Anders

Article: Castagnos 2010 - Consumer decision patterns through eye gaze analysis

Relevance: Medium (Method)

Resume: Builds a generic perfume e-shop. Then conducts user experiments with an eye-tracker with the purpose of defining customer decision patterns

Experiment: Within subject. Participants asked to find 3 to them unknown perfumes which they would buy for themselves and to find a perfume for a gift (2 separate tasks). Measured fixations on ROIs and reading heat maps. Also counted amount of mouse clicks and where the participants clicked. Furthermore an interview was conducted.

Read by: Anders

Article: Callahan 1988 - An emperical comparison of pie vs linear menus

Relevance: High

Resume: Conducts a within-subject study of the efficiency of linear menus vs. radial menus. 33 participants each solved 60. They find radial menus to be more efficient, but participant preferences to be split. It is also noted that radial and linear menues are only practical within a limited menu size. Furthermore they reference that the efficiency of all menu styles seem to converge after practise.

Gaze plot/Scanpath

Goldberg 2002 - Eye tracking in web search tasks design implications

Castagnos 2010 - Consumer decision patterns through eye gaze analysis

Pretorius 2005 - The added value of eye tracking in the usability evaluation of a network management tool

AOI

Castagnos 2010 - Consumer decision patterns through eye gaze analysis

Dwell times

Goldberg 2002 - Eye tracking in web search tasks design implications

Castagnos 2010 - Consumer decision patterns through eye gaze analysis

Transition matrix (number of transitions from and to each AOI)

Goldberg 2002 - Eye tracking in web search tasks design implications

Conati 2007 - Using eye-tracking data for high-level user modeling in adaptive interfaces

Fixations

Castagnos 2010 - Consumer decision patterns through eye gaze analysis

Pretorius 2005 - The added value of eye tracking in the usability evaluation of a network management tool

Gaze

Pretorius 2005 - The added value of eye tracking in the usability evaluation of a network management tool

Time till 1st fixation

Pretorius 2005 - The added value of eye tracking in the usability evaluation of a network management tool

Saccade duration

Goldberg 2002 - Eye tracking in web search tasks design implications

Heat map

Goldberg 2002 - Eye tracking in web search tasks design implications

Castagnos 2010 - Consumer decision patterns through eye gaze analysis

Pretorius 2005 - The added value of eye tracking in the usability evaluation of a network management tool

Amount of fixations on heatmap

Goldberg 2002 - Eye tracking in web search tasks design implications

Pretorius 2005 - The added value of eye tracking in the usability evaluation of a network management tool