Universidad Autónoma de Baja California

School of Chemical Science and Engineering

"Interaction Rules for Graphical User Interfaces based on Executive Functions."

THESIS

For the degree of PhD

Presents:

Andrés Mejía Figueroa

Dr. J. Reyes Juárez Ramírez
Thesis director
Universidad Autónoma de Baja California

Dr. David Eduardo Saldaña Sage
Thesis Codirector
Universidad de Sevilla

Tijuana, Baja California, Mexico

January 2018

Acknowledgements

To my parents for being helped and supported in fulfilling my goals, for having sacrificed for my wellbeing so that I could have a good future.

To my dear fiancé, Nancy, for having patience and understanding to withstand my so many absences and carelessness during my graduate studies. For having been with me in the highs and lows and support me in moving forward with my goals.

To my groups of students of Requirements Engineering and Interaction Design, having served as guinea pigs and for their support in various projects and research.

To all the special education schools in Tijuana (CADEE, ACI, Polaris, CeVi, Fane, Eduke, Pro Autismo…sorry of I missed anyone), parents, and children that participated in this research and lets us experiment with them for the making of this thesis.

To professor Martha Lilia Lopez, which introduced me to the main concepts of psychology that guided this research, and for training psychology students for the applications of the neurological tests.

To the National Council of Science and Technology (CONACYT), for the financial support provided during the duration of the 4 years of my PhD studies.

To the Autonomous University of Baja California, School of Chemical Sciences and Engineering and the Computer Engineering Coordination, for all the support they gave me to carry out research and participation in conferences.

To the University of Seville, Department of Developmental and Educational Psychology, for giving such a warm welcome, for all their invaluable help and guidance with the research needed, helping mesh together different disciplines.

To Coca-Cola, Jose's Colombian Coffee, and Cune wine… for giving me the energy to withstand all those sleepless nights doing research, crunching numbers, and writing papers (no, I'm not sponsored by them, I promise).

Table of Contents

Abstract 1

Resumen 2

Introduction 3

CHAPTER I: Theoretical Foundations 11

User Modelling in HCI 11

Executive Functions 14

Conceptual models of Executive Functions 16

Development of Executive Functions in a Lifetime 20

Executive Functions and Interaction Design 22

The problem of cognitive load 23

CHAPTER II:Related work 30

Cognitive Modelling 30

Human Factors and Adaptive User Interfaces 33

Adaptive User Interfaces applied for special need users 38

User Interface Adaptation 39

CHAPTER III: Proposed Solution: Interaction Rules based on Executive Function measurements 42
CHAPTER IV: Experimental Design 48
Research questions 48
Hypothesis 48
Experimental Design 49
Objectives 49
Sample 49
Structure of study 49
Variables 50
Instruments 52
CHAPTER V: Experimental Results 58
Data analysis 58
Usability Testing Results 58
Binary Success 58
Levels of Success 62
Time on task 66
Error rate 69
Cognitive load 70
Interaction Rules 72

Analysis of Variance of Cognitive Load and Executive Functions 78

Regression Analysis 79

Discussion 83

CHAPTER VI: Conclusion 86

References 89

Table Index

Table 1. Conceptual models of executive functions 19

Table 2. Keystroke-level Model operators. 31

Table 3. Interface parameters and correlated human factors. 34

Table 4. Results of binary success for the group with no ASD. 59

Table 5. Success rate for group with ASD level 1. 60

Table 6. Success rate for group with ASD level 2. 61

Table 7. Number of cases for each level of success. 62

Table 8. Percentage of each level of success per prototype. 62

Table 9. Count of level of success for group of ASD level 1. 63

Table 10. Percentage of level of success for the group of ASD level 1. 64

Table 11. Level of success count of the group with Autism level 2. 65

Table 12. Level of success percentage of the group of ASD level 2. 65

Table 13. Time on task for all interface patterns for group with no ASD. 67

Table 14. Time on task for all interface patterns for group with ASD level 1. 67

Table 15. Time on task for all interface patterns for group with ASD level 2. 68

Table 16. Error for all interface patterns for group with no ASD. 69

Table 17. Error for all interface patterns for group with ASD level 1. 70

Table 18. Error for all interface patterns for group with ASD level 2. 70

Table 19. Cognitive load test results for all interface patterns for group with ASD level 0. 71

Table 20. Cognitive load test results for all interface patterns for group with ASD level 1. 71

Table 21. Regression analysis of the accordion pattern. 79

Table 22. Regression analysis of the accordion pattern. 79

Table 23. Regression analysis of the accordion pattern. 80

Table 24. Regression analysis of the accordion pattern. 80

Table 25. Regression analysis of the accordion pattern. 80

Table 26. Regression analysis of the accordion pattern. 81

Table 27. Regression analysis of the accordion pattern. 81

Table 28. Regression analysis of the accordion pattern. 81

Figure List

Figure 1. Example of several interface design patterns. Credit goes to: https://theresaneil.wordpress.com/page/6/ 23

Figure 2. Ruler used for the mental workload measurements, figure extracted from (Laurie-Rose et al., 2014). 54

Figure 3. Success rate for group with no ASD........ 59

Figure 4. Success rate for group of ASD level 1........ 60

Figure 5. Success rate for group with ASD level 2. 61

Figure 6. Level of success of group with no ASD. 63

Figure 7. Level of success for group with Autism level 1 64

Figure 8. Level of success of group of ASD level 2. 66

Figure 9. Decision tree for the best performing patterns based on usability testing scores. 73

Figure 10. Decision tree for the second best-performing patterns based on usability testing scores........ 75

Figure 11. Decision tree for the eight best performing patterns based on usability testing scores. 77

Figure 12. Basic layered architecture of the presentation layer of an app with the integrated interaction rules........ 84

Abstract

Human Factors have been for several decades one of the main factors of contention when trying to develop a usable system. From physical to cognitive characteristics and everything in between, the attributes of a user can impact on several aspects of the design of said software. Cognitive aspects, like executive functions are considered to some extent, mainly working memory, but several others are not considered, like cognitive flexibility and planning, which can have a direct impact on how we can take on a task and follow through it, affecting how we interact with an interface. Although there are guidelines for some characteristics, there is no definitive model with what characteristics to consider, what metrics to use and their effect on the interface. In this doctoral thesis, we propose a set of interaction rules that obtained by measurements of some of our executive functions of working memory and cognitive flexibility with different interface design patterns and analyze the relationship with the cognitive load produced by the design pattern, with the end goal of obtaining a set of guidelines more precise than the currently available.

Keywords: executive functions, interface design, cognitive load, working memory, cognitive flexibility.

Resumen

Los Factores Humanos han sido por varias décadas uno de los mayores factores de contención cuando se intenta desarrollar un sistema usable. Desde característica físicas a cognitivas y todo de por medio, los atributos de un usuario impactan varios aspectos del diseño del software. Algunos aspectos cognitivos se consideran actualmente, como la función ejecutiva de memoria de trabajo, pero muchos otros no, como es el caso de la flexibilidad cognitiva y la planeación. Aunque existen algunas guías de diseño para algunas características, no existe un modelo definitivo sobre qué características considerar, que métricas usar y sus efectos en la interface. En esta tesis doctoral, proponemos una serie de reglas de interacción obtenidas por medio de mediciones de nuestras funciones ejecutivas de memoria de trabajo y flexibilidad cognitiva dependiendo del patrón de diseño de interfaz utilizado y analizamos la relación de la carga cognitiva generada en el usuario por los patrones de diseño, con el fin de obtener guías de diseño más precisas.

Keywords: funciones ejecutivas, diseño de interface, carga cognitiva, memoria de trabajo, flexibilidad cognitiva.

Introduction

Computers today have changed the way we live, in some way or another ingrained in our everyday activities, from work to leisure, having access to news and information in an instant in the palm of our hand. But it wasn't always that way. In the early days of computing, computers were huge, extremely expensive, and needed specialized training to use them. Along the way, computers were used in more and more situations and applications, from military to commercial, but with a lot of challenges. One being the Software Crisis, coined in 1968, which refers to the rising complexity of producing software, giving birth to the discipline of Software Engineering (Bauer, Bolliet, & Helms, 1968).

Another problem which slowly emerged, but was noted, was the problem of usability of software and physical systems. Bad design choices caused several fatal errors and accidents because of user error that could have been prevented with a good design (Jacko, 2012). As such the discipline of Human-Computer Interaction (HCI) emerged, formed by several other disciplines, such as human factors, cognitive science, and psychology.

HCI first appeared as the application of cognitive science in the information technology scene. Incidentally, as personal computing took off in the late 70's, so did cognitive science as a way to model the mind and its processes (Jacko, 2012). Thus, laying the groundwork of HCI by presenting the necessary tools needed to tackle the slew of new challenges presented with universal access to personal computing.

One of the biggest contributors to HCI was the discipline of Human Factors and Ergonomics (HFE), which focuses on human characteristics and its implication on the design of any device and system. Many of the HFE contributions in industry and ergonomics made its way into HCI; such is the case with cognitive ergonomics (or cognitive engineering), which studies cognition in a work environment; this includes things such as perception, mental workload, decision-making, stress, among others (Jacko, 2012).

Some examples of cognitive engineering include the Human-Processor Model and the GOMS family of models proposed by Stuart K. Card, Thomas P. Moran & Allan Newell in 1983. Described in their seminal work: The Psychology of Human-Computer Interaction (Card, Moran, & Newell, 1983), in which they propose and describe, a cognitive model used to assess the usability of a software system and the time to complete a task in that software. They divided the human factors into three subsystems: perceptual, cognitive, and motor, all which are needed to interact with any system properly. In each subsystem, they outlined time frames for each process to calculate the estimated time of completion of a specific task. The completion time consisted of the sum of the time of all the processes involved in each subsystem in the task. This model served as a base for future cognitive architectures and subsequently modified versions of the GOMS model.

The perceptual subsystem consisted of the senses, the perceptual processor, and the visual image storage, and auditory image storage, which makes up the working memory. The cognitive subsystem, which is formed from the working memory, long-term memory and the cognitive processor. And the motor subsystem, consisting of the motor processor and the movement response.

Each part of each subsystem has a set of values depending on their function and nature. Each processor has a cycle time, and each memory has a decay time depending on the type of information stored. For example, the values for the mean of pure working memory capacity is 3 chunks, with a range of 2.5-4.2 chunks, decay half-life of working memory mean is of 7 seconds, with a range of 5-225 seconds.

So, for HCI to achieve its goal of usable interfaces, it must consider not only computer factors (software and hardware) but also the task and all the human factors that could affect achieving the completion of the task at hand (Kim, 2015).

As a result, there exist several guidelines for the design of interfaces for users with some form of special needs of varying characteristics, such as the Web Accessibility Initiative by the W3C (Caldwell, Cooper, Reid, & Vanderheiden, 2008) to make web content more accessible to people with some form of disability.

Some example guidelines from the W3C include: provide text alternatives for non-text content, provide an alternative to video only and audio only content, provide captions for videos with audio, follow a logical structure, don't use presentation that relies solely on color, accessible by keyboard only, among others.

Most guidelines are very similar or are an offshoot of the golden rules of HCI, first proposed by Shneiderman in 1987 (2010). These principles laid the foundation for most guidelines, and if followed, provide a framework in which one could design a usable interface (most of the time at least). The golden rules of HCI are: Strive for consistency, enable frequent users to use shortcuts, offer informative feedback, design dialog to yield closure, offer simple error handling, permit easy reversal of actions, support internal locus of control, and reduce short-term memory load.

There are other more specific guidelines for certain ailments, such as for learning disabilities for example. In (Freyhoff, Hess, Kerr, Tronbacke, & Van Der Veken, 1998) a series of guidelines and recommendations are outlined for the access of information by people with a language disability. Some of the guidelines are:

- Never use a picture as a background for the text.
- Try to put one sentence on one line.
- Keep sentences together on one page.
- Do not fill with too much information.
- Use maximum two typefaces.
- Use a large font size.
- Use a clear typeface.
- Careful on how to emphasize text. Do not use block capitals and italics in the text. Use bold or underline for emphasis.
- Illustrations must be in sharp focus.
- Don’t use inverted printing (light text on a dark background), use dark print (dark print on light background).
- Use colors for pictures, boxes, etc.
- Use headings and other navigational aids.

In similar fashion to the recommendations and guidelines for interface design, there are also some rules of thumb for usability evaluation of the interfaces similar to the golden rules of

HCI. Proposed by Jacob Nielsen (Nielsen, 1994, 1995) where he compared sets of usability heuristics with a database of existing usability problems to determine what heuristics best explain actual usability problems. From the analysis he derived the heuristics that had best coverage of the usability problems, the heuristics being:

- Visibility of system status. The system should always keep users informed about what is going on, through appropriate feedback within reasonable time.
- Match between system and the real world. The system should use concepts and language familiar to the user’s.
- User control and freedom. Give users an “emergency exit” when they stumble upon a part of the system they weren’t meant to or make a mistake. Support undo and redo.
- Consistency and standards. Follow platform conventions of design so that users must not be surprised and learn a different interface for each part of the system.
- Error prevention. Design good error messages and take steps to prevent user errors.
- Recognition rather than recall. Minimize the user's memory load by making objects, actions, and options visible. Make an effort for that the user should not have to remember information, make the software save pertinent variables, instructions, etc., so the user won’t have to.
- Flexibility and efficiency of use. Give options for experienced users, such as shortcuts and access to frequent actions.
- Aesthetic and minimalist design. Present only relevant information, too much information can cause a memory overload and diminish visibility.

- Help users recognize, diagnose, and recover from errors. Simple and explained in plain language, indicating the problem, and suggesting a solution.
- Help and documentation. Include help guides and instructions in the software and outside the software, such as user manuals.

There are a lot of factor to consider when designing a Graphical User Interface (GUI). There are cognitive models which help predict user behavior when using a GUI, but there is still much to do since there are other factors that affect usability. One of which is the cognitive load produced by the design of the interface.

One of the more important aspects of consideration when designing a usable GUI is one of the golden rules of HCI: Reduce cognitive load (Jacko, 2012; Shneiderman, 2010). The way we normally try to achieve this is by simplifying the number of items on the interface to the bare minimum, so we don't overload the user with information and choice.

Cognitive models, such as the Human-Processor model, don't directly have a component of cognitive load, but have parts that have an impact or can be considered to be affected by cognitive load. For example, in the Human-Processor model we have memory components (short and long term) with a certain capacity (chunks) and decay rate; cognitive load can be considered the number of chunks used up by information in the short-term memory.

In a similar vein, the Keystroke Level Model (Card et al., 1983) has an operator (M) which represents the mental preparedness for the task; things like planning, analyzing, decision making, etc. Denotes some high level cognitive functions, although they aren't outlined which ones or how they could have an impact. This task, by its own nature, produces some load on the mental resources since they are needed to analyze what is needed to be done and act accordingly.

Much of the nature of cognitive load is determined by many factors, one of which are our cognitive capabilities, including some of our executive functions. There is currently no metric for the cognitive load produced by an interface design, although it can be inferred through other means (Antonenko, Paas, Grabner, & Van Gog, 2010; F. Paas, Ayres, & Pachman, 2008; Van Gerven, Paas, Van Merriënboer, & Schmidt, 2004). Every aspect of the interface design consumes mental resources, thus causing a certain degree of cognitive load, thus the choice of interface design patterns, depending on the task to be done, can affect the level of cognitive load induced and in turn affect the usability of the software.

Even with guidelines and models, it is still hard to properly design an interface that can be usable by all. Using the guidelines can only take you so far, since even if you consider most of them, there will always be users that will still have a hard time using the software, especially if they have some form of disability. The guidelines are too general, encompassing a wide variety of users with a wide range of characteristics. There are efforts to specify much more specific user models with their metrics, but there is still much work to be done.

In this dissertation, we present a set of interaction rules based on our executive functions of working memory and cognitive flexibility and their relationship with interface design and common interface design patterns and analyzed the influence of the cognitive load produced by the interface design patterns , in an effort to determine much more specific guidelines and be able to select better suited design patterns depending on the potential users of the software.

Chapter I:

Theoretical Foundations

"If we want users to like our software we should design it to behave like a likable person: respectful, generous and helpful."

- Alan Cooper

Theoretical Foundations

User Modelling in HCI

Modelling the user has become one of the more important parts of the software development life-cycle. With the current trend of User-Centered Design and User Experience Design, developing software is not only reserved to software engineers, now it is needed team members with training in other disciplines and other expertise to analyze and develop a more usable and friendly system.

As mentioned earlier HCI developed from a combination of disciplines, from human factors to cognitive science. One of the first contributions of HCI was an attempt to model the way a person thinks and behaves in order predict how a person would use a system. As a result, there are several cognitive architectures in HCI as a means to model user behavior and interface evaluation. Byrne (2001) defines a cognitive architecture as a hypothesis about aspects of human cognition that remain constant over time and are independent of the task. As such we can make such those approximations of the real human cognition to have a better understanding of the proper design and analysis of software interfaces.

One of the milestones of cognitive modeling in HCI was the Model Human Processor (Card et al., 1983), where it was the first that introduces the notion on information processing from the user and qualitative and quantitative predictions of user actions based on psychological studies. It also gave rise to the GOMS family of cognitive architectures which are some of the most widely used today.

GOMS (Goals, Operators, Methods, and Selection Rules) (Card et al., 1983) is a derivation of the Model Human Processor. The GOMS model functions in a similar way to a task analysis by analyzing the actions needed to perform a task efficiently. The GOMS model has several components, them being:

- Goals. What the user wants to accomplish.
- Operators. Actions needed to perform each goal, such as keyboard strokes, button presses, etc.
- Methods. Sequences of operators needed to accomplish said goal
- Selection Rules. Conditions where the user might select one method over another.

With GOMS, much like the Human Model Processor, we can calculate an average time of completion of a task based on the average timing of basic actions, such as a button press or move the cursor, obtained from previously done research.

Subsequent variations of the GOMS model were presented, such as the KLM (Keystroke-level Model) (Card et al., 1983), where the task analysis reaches keystroke level operators, such as key presses. Execution time is the emphasis ignoring all else, such as goals, methods and selection rules. It is considered a simplified version of GOMS. And NGOMSL proposed by Kieras (1994), where a formalized version of GOMS is presented with structured rules and notations.

Another popular architecture, EPIC (Executive-Process Interactive Control) (D. E. Kieras & Meyer, 1997), where executive processes are represented and control other processes during a multitasking performance in a similar way to executive functions. Also, perception and motor behaviors are given their module. Procedures and production rules specify modules, thus requiring knowledge about the task.

The ACT-R (Adaptive Control of Thought-Rational) architecture, proposed by Anderson (1996), was first developed as a theory of higher cognition and learning but subsequently adapted for HCI. It assumes that knowledge can be classified into two kinds: declarative and procedural. Declarative knowledge is represented in the form of chunks and are saved and made accessible through buffers. The most recent version, ACT-R/PM, proposed by Byrne (2001), incorporates perceptual motor capabilities similar to the EPIC architecture.

Another type of area where human factors take center stage in the development of the software is in Adaptive Interfaces. Adaptive Interfaces, as the name implies, adapt to the user's characteristics and the context of use of software, providing, in theory, a more personalized and usable experience (Dieterich, Malinowski, Kühme, & Schneider-Hufschmidt, 1993; Lavie & Meyer, 2010).

One of the core parts of the architecture of an Adaptive Interface is the User Model (Lavie & Meyer, 2010). To develop an adaptive system, we must first determine what will be the user characteristics that will be considered thus composing the user model. With this model, the user interface can be adapted based on the values, rules, and guidelines that resulted of user research. Most software applications that use user models often just consider some aspects of the user that they deem relevant to the application, there is no generic solution to be used, although there is research towards achieving that goal.

One example is the AVANTI project (Stephanidis et al., 1998) in which an AUI was developed for web documents, adapting itself to the users to some extent based on some information submitted by the user, such as language, eyesight, motor abilities, language and experience with similar applications. It also adapts itself in a dynamic way as they interact with the system based on user familiarity with specific tasks, ability to navigate, error rate, disorientation, user idle time, and repetition of interaction patterns. The adaptation also supports multi-modal systems by featuring integrated support to various input and output devices with appropriate interaction techniques to accommodate users with disabilities. The UI supports users with light to severe motor disabilities and blindness.

Although cognitive characteristics are being considered, there are still several that there is little to no research on the impact they could have on interface design, some of those being executive functions. One of the most commonly considered is working memory since it serves as a component of several other cognitive functions and is very used for academic purposes, but others such as planning and cognitive flexibility have had drawn very little attention.

Executive Functions

Executive function is a term used to describe cognitive processes necessary for the completion of goal-oriented tasks, usually carried out in large part at least, in the frontal lobes.

As many terms and definitions in science and engineering, executive function defies a formal definition for the reason that as more research is done on the nature of executive functioning, sometimes contradictory results are presented generating controversy and discourse. Even with similar yet sometimes different definitions, practically all agree that executive functions are the set of cognitive processes where cognitive abilities are used in goal-oriented tasks.

Delis (2012) defines executive functions as the ability to manage and regulate one’s behavior to achieve the desired goal. Similar to Delis, Miller and Cohen (2001) suggest that executive control involves the cognitive abilities needed to perform goal-oriented tasks. Lezak (2004) describes executive functioning as a collection of interrelated cognitive and behavioral skills that are responsible for the goal-directed activity, includes intellect, thought, self-control, and social interaction.

Notions of executive functioning were first observed in studies related to patients with abnormalities or injuries in the prefrontal cortex such as in Pribram (1973) and Luria (1973), which noted impaired ability to evaluate and regulate their behaviors and goal-oriented mindset.

Some of the elements of executive functioning include short and sustained attention, task initiation, emotional regulation, working memory, cognitive flexibility or shifting, planning and problem-solving ability. Encompassing not only cognitive processes but also emotional responses and behaviors (Hongwanishkul, Happaney, Lee, & Zelazo, 2005).

For this reason, executive functions are classified in “hot” or “cool” processes, where “hot” processes are the ones that have an affective component, such as emotional regulation, and “cool” processes which involve purely cognitive functions, such as working memory (Hongwanishkul et al., 2005).

Further research pertaining the prefrontal cortex and executive functions were done achieving similar observations, but subsequent studies also reproduced contradicting results. Such is the case in (Godefroy, Cabaret, Petit-Chenal, Pruvo, & Rousseaux, 1999) where the notion that all the processes for executive functions were in the frontal lobes was put in doubt. A study showed that patients with lesions of the prefrontal or posterior cortices were subjected to a series of conflicting and combined tasks. The results showed the expected prominent role of the frontal lobe but also that executive functions depend on multiple, separate, and modular control processes because certain patients with frontal lobe injury performed well on tests designed to assess executive functioning while others did not.

These revelations make sense since the prefrontal cortex is shown to be dependent on outgoing and ingoing connections to all other brain regions, such as the occipital, temporal and parietal lobes. Damage or deficits in any related region would result in some executive function impairment. The pre-frontal cortex works as a hub of communications between all regions of the brain that pertain to executive functioning, if communication between those regions is lost, or the region is damage, impairment in executive function occurs.

Symptoms of executive dysfunction include the inability to maintain attention, lack of impulse control, low working memory capacity, inability to plan future activities, inability to shift attention between different tasks or stimuli, difficulty generating new knowledge, among others (Chan et al., 2009; Hill, 2004a, 2004c).

As we can see, although there is some controversy on the nature of executive functioning and the great number of definitions proposed, there is the consensus that it involves the cognitive processes that manage goal-directed behavior.

Conceptual models of Executive Functions

To explain the patterns of impairments, and relationships of the executive functions, several models were proposed. For the time being, none have been formally adopted since the rationale behind each model differs depending on the field of study. On some of the models, several of the executive functions overlap or have different names for the same function, causing, even more, confusion among practitioners.

One of the first notions of executive functions was put forth by Baddeley (A. D. Baddeley & Hitch, 1974) when working memory takes the central role pertaining something he called the central executive. Baddeley (A. D. Baddeley & Hitch, 1974) defines working memory as "a limited capacity system for the necessary storage and manipulation of information necessary for such complex tasks as comprehension, learning, and reasoning."

Originally Baddeley proposed the first version of the working memory model (A. D. Baddeley & Hitch, 1974) which expanded upon the notion of short-term memory, formed by three components: Visuospatial, phonological loop and central executive (which regulates the other components).

Based on this first model, Baddeley later proposed a modified version of the model in which he expanded it to include an episodic buffer as a fourth component (A. Baddeley, 2000), where working memory of a central attentional system, the central executive, which manages selective attention and exchanges of information with the long-term memory, it also regulates other subsystems (visuospatial and phonological loop). These subsystems serve a temporal memory of visual and audial information separately. The episodic buffer integrates information from the subsystems and also from the long-term memory to create a single integral representation of an event.

Luria (1973) documented behaviors of patients with frontal lobe trauma. When confronted with a problem, they could not plan ahead, could not analyze the limits and constraints of a problem, and also control their impulses. With these observations, Luria concluded that the frontal lobes had an enormous impact on problem-solving behavior and regulation, later called executive functions.

She described the executive functions as being anticipation (set a goal), planning (plan ahead of time the necessary steps to achieve said goal), execution (execute said plan and adapt to the constraints of the problem or task, or to new unforeseen challenges that arise, and stay on point when carrying out the necessary steps), and self-monitoring (evaluate each step in order to be sure that they are correctly done).

In similar fashion to Luria, Lezak (2004) conceptualized Executive Functioning in four domains: Volition, purposive action, effective performance, and planning. Each consisting of their characteristics and behaviors. Volition is defined as the will or conscious decision to initiate and perform an action. It requires a capacity to plan and set goals, motivations, and self-awareness. Once a goal is defined, and there is volition to carry it out, then it is necessary to plan the actions needed to achieve the desired goal. Planning is the ability to identify the actions and consequences of said actions that needs to be done. Once a plan has formulated the sequence of actions are needed to be performed, in this part the executive domain of purposive action, which consists of the initiation of the required steps of the plan, as well as the capacity to monitor, evaluate and modify the plan as needed. The final part of the framework is effective performance, in which consists of the capacity to monitor, regulate and self-correct behavior.

Some defined conceptual models as a way to try to explain a particular affliction or disorder. Such is the case of Barkley (1997). The Self-Regulatory Model proposed as a way to explain the symptoms of Attention Deficit Disorder defined self-regulation as a way to modify the subsequent response of individuals to an event. According to Barkley, self-regulation comprises most of the executive functions from other models, such as planning and goal-directed behavior and emotional regulation. He considers that the first necessary part for proper executive functioning is behavioral inhibition, in charge of providing the necessary delay for executive processes to occur, managing four executive domains: working memory, self-regulation of affect/motivation/arousal, internalization of speech, and reconstitution.

These are just a few of the conceptual model proposed, as mentioned previously, each varies a bit and sometimes overlap depending on the area of expertise of each researcher. They all coincide with the functions that regulate goal-oriented behavior, being the most commonly mentioned: planning, cognitive flexibility or shifting, working memory, response inhibition, generalization, and self-monitoring (Table 1).

Table 1. Conceptual models of executive functions

Model	Executive Functions
Working memory model (A. Baddeley, 2000; A. D. Baddeley, Baddeley, & Braddlely, 1986; A. D. Baddeley & Hitch, 1974)	Working memory (Represented by the central executive)
Luria (Luria, 1973)	Anticipation, planning, execution, self-monitoring
Lezak (Lezak, 2004)	Volition, purposive action, effective performance, planning
Self-Regulatory Model (Barkley, 1997)	Working memory, self-regulation of affect/motivation/arousal (comprising planning and goal-directed behavior and emotional regulation), internalization of speech, and reconstitution.

Development of Executive Functions in a Lifetime

Our Executive Functions mature and decline during the progression of our lifetime, developing first in the formation of our nervous system in the womb, to our peak in our adulthood, to the slow decline of our cognitive capabilities as we age into our later years.

Anderson *et al.* (2010) present a timeline on how our executive functions develop and decline over our lifetime, from kindergarten to adulthood, to our elderly years.

During the first 12 weeks of life, infants can detect a structure of events with a defined goal. During around the 7 to 8 months, the firsts signs of inhibition and working memory can be observed and the ability to generate knowledge, such as the ability to distinguish animate and inanimate objects. At the 1-year mark, children show the ability for joint attention, which refers to the ability to shift attention to another stimulus appointed by another person, for example, the parents pointing attention to something in a coloring book. By the 14 months of age, children exhibit the ability for social referencing, meaning that they can recognize emotions from facial expressions from other people and use that information in decision making. By two years of age, shows improvements in working memory and inhibition, and an understanding of pretense, this being shown most commonly in the form of imaginary play.

During the preschool and early year, major strides of the development of executive functions can be observed. During the third year, major improvements in inhibitory control and sustained attention are seen, as well as affective decision making. By age 4, cognitive flexibility and better success at the false-belief task, meaning they can better infer emotions and intentions from other persons by “being in their shoes.” At age five significant gains in working memory are observed and beginnings of proper planning strategy formation and goal-directed behavior, also an awareness that a belief can be held about another's belief. At the age of 6 years, the theory of mind is more developed to near adult-like status; and at the age of 7, a better understanding of conflicting mental states are developed.

In late childhood, some executive functions reach relative maturity while other suffer significant improvements. At age eight cognitive flexibility matures; inhibition, awareness and sustained attention improve in a span till age 11; and improve understanding of social deception and metaphors. At age nine gains in planning and working memory are seen. Also, an understanding of social rules develops until the age of 11. At the age of 12, the goal-directed behavior increases.

Adolescence is characterized mainly by significant advances in Executive Functions related to the problem solving and decision-making domain. By the age of 14, there are significant affective decision-making improvements until the age of 17. By the age of 15, there is improved attention, processing speed and inhibition. In the range of 16-19 years of age, there are gains of working memory, planning and problem-solving.

In adulthood, we see the full development of the Executive Functions with a decline in performance as we mature in age. From the age of 20-29, working memory, planning, affective decision making, are all mature. From the age of 50-64, we start seeing a decrease in the formation of new knowledge, planning, working memory, shifting, processing speed and goal setting. From 65-74 years there is a significantly reduced performance in affective decision making. At 75 years and more, usually, there is a severe theory of mind deficit, which looks like a lack of empathy.

Executive Functions and Interaction Design

A User Interface (UI) is defined as the point where information exchange between the user and the system occur (Dix, 2009). Today, with most personal computers and mobile devices such as tablets and smartphones, GUI's are still the preferred type of interface for the average user.

GUI's normally consist of several interface components, such as buttons, labels, text areas, for example. Depending on the type of task, there can be a wide selection of interface components on which the user can accomplish the task. The recurring use of a set of interface components for a particular task is called user interface design patterns (Dix, 2009). For example, one of the most common patterns is the login screen. Usually, it consists of two textboxes, one for your login and other for your password, and two buttons, one to enter the system and other to cancel. Other examples of design patterns include tabs, wizard, accordion, and navigational drawer just to name a few (Figure 1).

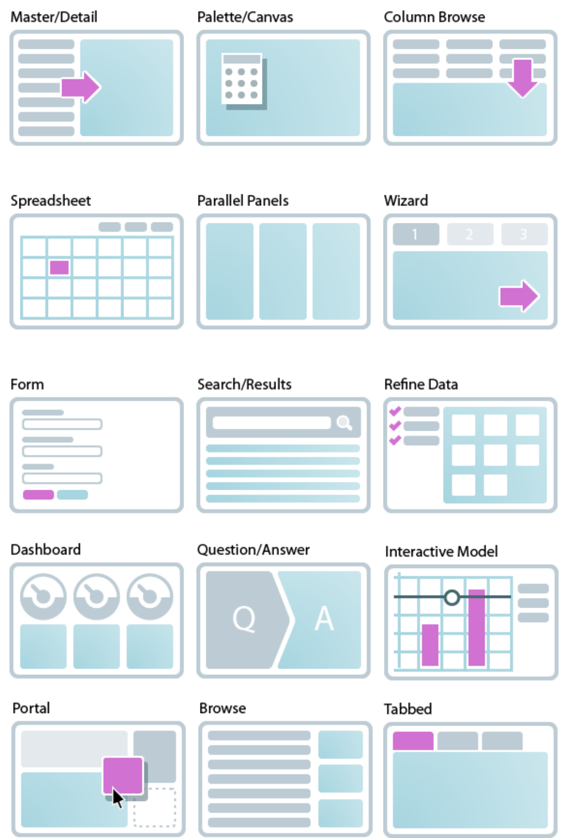


Figure 1. Example of several interface design patterns. Credit goes to: https://theresaneil.wordpress.com/page/6/

Often during software development, depending on the results of the user and task analysis, certain usability patterns and interface design decisions are made in the hopes that the interface will be usable for most of its intended user audience, nevertheless often it is not the case.

Some patterns might be useful for certain tasks, but require certain user characteristics to be properly used, limiting the usability for some users that might have an impairment or notably different characteristics.

The problem of cognitive load

One of the main problems that HCI tries to tackle in interface design is to minimize as much as possible the cognitive load upon the user caused by the design decisions of the user interface.

Cognitive load is defined as the mental resources available for the task at hand (Mayer & Moreno, 2003). The capacity of mental resources varies from person to person and has several factors that contribute to it, but is more tied to working memory and long-term memory. This makes working memory one of the most important executive functions since is used as a component of other more complex executive functions, such as planning and cognitive flexibility (Miyake *et al*., 2000).

Cognitive load can is classified into three categories depending on the nature of the load (Mayer & Moreno, 2003). Some load is imposed on the working memory by the nature of the complexity of the task, called intrinsic load. Another category of cognitive load takes place from the way the information or task is presented to the user. In the case of interface design, that means that the design decisions of the interface have a direct impact on this type of cognitive load, which is called extraneous cognitive load. Both types of the cognitive load must be dealt with by the resources available to the working memory allocated to both cognitive loads, this type of load is called germaine load, which are the resources needed for learning and storing information in schemas in long-term memory.

Cognitive load can never be fully eliminated from any task since we always need mental resources for any action, be it a movement, a selection or simply searching for an item on the screen; but there are some steps to minimize it.

In the case of interface design, we can attack the problem by minimizing the extraneous load caused by the difficulty of use of the interface and quantity of information shown at once, and in some cases the intrinsic load, by simplifying and automating some steps of the task.

Can we measure cognitive load?

One of the leading researchers in cognitive load theory, John Sweller in his work (Sweller, Ayres, & Kalyuga, 2011) lists several ways in which we can measure cognitive load.

Although there is no exact metric for cognitive load, there is a way we can infer the effects of cognitive load and have an approximation of a measurement.

Indirect measures

The first approximations of a measurement for cognitive load were developed by indirectly by experimentation that examining the relationship between problem-solving and learning. One of the approximations was with the use of computer models. It started with research focused on the inefficiency of problem-solving ability as a learning strategy where it demonstrated that learning strategies that used a lot of problems solving search led to worse outcomes than strategies with the less problem-solving search. They argued in (Sweller, 1988) that the problem-solving searches cause high extraneous cognitive load, impeding proper schema creation and acquisition.

With a production system model, Ayres and Sweller (2014) demonstrated that higher problem-solving search required a much more complex model to be simulated, thus giving credence that it would be a much higher burden to the working memory.

Another approximation to infer cognitive load was the use of performance indicators during the acquisition or learning phase. Chandler and Sweller (1992) showed that students using a learning technique that increased cognitive load impacted performance during the learning and acquisition phases. In later research, it also showed that it also increased error rates were higher during the acquisition phase where high cognitive load learning techniques were used. Subsequent research gave credence to the idea that error rates might also be used as an indirect way of measuring cognitive load (P. L. Ayres, 2001) since results showed that students made the most errors in mathematical tasks that required high decision-making skills with many variables to be considered.

Subjective measures

Other ways to get an approximation of cognitive load is using subjective measures. Previous research (Bratfisch, 1972) indicated that user introspection of mental effort could be used as an index of cognitive load. Mental effort is similar to cognitive load that refers to the cognitive capacity that is allocated to accommodate demands imposed by the task, thus can be considered a reflection of cognitive load.

In (F. G. Paas, 1992) a 9 point Linkert Scale was used, which ranged from 1, which represented low mental effort, and nine which represented high mental effort. The results showed a correlation between students that used hypothesized lower cognitive load instructional design and the lower mental effort rating that they gave; the same with the students that used a higher hypothesized cognitive load instructional design, which yields higher mental effort scores.

Measurement through a secondary task

The traditional method for the measurement of cognitive load is through the use of a secondary task (Britton & Tesser, 1982) in what is called a dual-task methodology. It requires that the subject engages in a subsequent task after a primary one. Depending on the cognitive load that the subject suffers from the first task, his performance will suffer in the second one because of it.

Physiological measures

One of the first physiological measures of the cognitive load was of cognitive pupillary response (Van Gerven et al., 2004). By testing several tasks with a very varied cognitive load to working memory, they found that there was a relationship between pupil dilation and perceived memory load. Pupil dilation increased with rising cognitive load measures, although there are some age limitations, the dilations decreases with age. In the study, elderly participants did not show pupil dilation on cognitive tasks.

Another successful way of measuring cognitive load in a more accurate and precise way than with subjective measurements is with the use of Functional Magnetic Resonance Imaging (fMRI) and Electroencephalography (EEG) (Antonenko et al., 2010; F. Paas et al., 2008). During an experiment where a subjective measure of mental effort was used in a task, an EEG captured alpha, beta and theta brain wakes. It showed that even when in the subjective scores there was no discernable difference between tasks, the EEG measurements were sensitive enough to show a sizeable difference in perceived mental effort.

Other forms of measuring cognitive load are the use of eye tracking in multimedia environments and applications (Palinko, Kun, Shyrokov, & Heeman, 2010; Van Gog & Scheiter, 2010). It found that different combinations of text and images required a different level and cognitive processes, thus more cognitive load; which correlated with varying degrees of eye fixation.

Although there are some ways to get an approximation of the cognitive load that a user is currently having, there is no real way to measure the induced cognitive load of a user interface on the user. Also, no precise design guidelines for different levels of executive functions, such is the case of working memory, which is greatly affected by cognitive load.

Chapter II:

Related works

"If I have seen further than others, is by standing in the shoulders of giants"

- Isaac Newton

Related work

Cognitive Modelling

Predicting how a user will perform when using a software system can helps us determine when and where usability problems can occur in a given system. Understanding how a user thinks can help us model a better way of interaction. HCI started out by modeling the user with the Human-Processor Model (Card et al., 1983), based on current cognitive models of the time of certain functions, outlined values for actions that the user must take to interact with a system tan approximation of time of completion of a given task. This model served as a basis for subsequent models, such as GOMS and KLM, each one with a different emphasis.

GOMS (Goals, Operators, Methods, and Selection Rules) functions as a task analysis by dividing each task into different components (Goals, Operators, Methods, Selection Rules), while KLM (Keystroke-level Model) (Card et al., 1983) focuses on execution time consisting of six operator each with a set value, thus making it simple to calculate an execution time depending on the task (Table 2).

Table 2. Keystroke-level Model operators.

Operator	Time (sec)
K	Keystrokes Guidelines: .08 (135 wpm: best typist) .12 (90 wpm: good typist) .20 (55 wpm: average skilled typist) .28 (40 wpm: average typist) .50 (typing random letters) .75 (typing complex codes) 1.20 (worst typist and unfamiliar with the keyboard)
P	1.1
H	0.4
D	.9nD +. 16 lD
M	1.35
R	system dependent
B (mouse button press or release)	0.1
Click a Link/ Button	3.73
Pull-Down List (No Page Load)	3.04
Pull-Down List (Page Load)	3.96
Date-Picker	6.81
Cut & Paste (Keyboard)	4.51

Operator	Time (sec)
Typing Text in a Text Field	2.32
Scrolling	3.96

New adaptations of the KLM model emerge with the rise in popularity of touch screen devices, such as tablets and smartphones. The Touch-Level Model (Rice & Lartigue, 2014) takes some operators of the KLM and adapts them to new devices while also proposing some new operators. Currently it is being expanded upon as more operators are added as research progresses. Some new operators are: Distraction level, pinch, zoom, tap, swipe, tilt, rotate, and drag.

EPIC (Executive-Process Interactive Control) (D. E. Kieras & Meyer, 1997), was designed to be a cognitive architecture suited for multimodal and multiple task interaction. In some ways similar to the Human-Processor Model, it includes sensory-motor processors, while adding a production-rule cognitive processor.

The ACT-R (Adaptive Control of Thought-Rational) architecture, proposed by Anderson(1996), consist of two modules: perceptual-motor modules and memory modules. Memory modules are classified into declarative and procedural. Declarative knowledge consists of facts, for example: 2+3=5, while procedural is made of productions or knowledge of how things are done. Declarative knowledge is represented in the form of chunks that are saved and made accessible through buffers and a pattern matcher searches production that matches the state of the buffers. Depending on the state of the buffers, the production can modify the buffers and change the state of the system. The most recent version, ACT-R/PM, proposed by Byrne (Byrne, 2001), incorporates perceptual motor capabilities similar to the EPIC architecture.

4CAPS (Cortical Capacity-Constrained Concurrent Activation-based Production System, is a cognitive architecture, built as a production system, where computations are distributed among different processing centers, mimicking the distribution of cerebral activity in the different cortical regions of the brain, with certain assumptions: each cortical can perform multiple cognitive functions, each cortical has limited resources, the cortical network changes as the regions resources as saturated, and communication between cortical regions are subject to constraints, and the activation of a cortical area as measured by imaging techniques varies as a function of its cognitive workload (Varma & Just, 2006).

As mentioned in (Bratfisch, 1972), one of the principal problems with cognitive architectures is the knowledge engineering problem, since all architectures need information about how the interface is used, what the tasks are, and how to do said tasks.

Human Factors and Adaptive User Interfaces

One of the core parts of the architecture of an Adaptive Interface is the User Model. To develop an adaptive system, we must first determine what will be the user characteristics that will compose the user model. With this model, the user interface can be adapted based on the values of the characteristics that are being considered of the user.

Most software applications that use user models often just consider some aspects of the user that they deem relevant to the application, there is no generic solution to be used, although there is research towards achieving that goal.

Some works center on what user characteristics must be considered for interactive systems, such as Zhang *et al* (Zhang, Carey, Te'eni, & Tremaine, 2005) where it is proposed a methodology to integrate Human-Computer Interaction practices in the software development life cycle. They describe some user characteristics that they consider important to take into consideration when designing software.

Zudilova-Seinstra (2007) which notes what human factors can be taken into consideration based on the Wagner's Ergonomical Model (Wagner, 1992) when designing scientific problem-solving software to improve usability. Seven human factors were evaluated regarding their influence in manual adjustments by the users by using Yule's coefficient of colligation. The human factor analyzed were: gender, age, learning abilities, verbal and nonverbal IQs, locus of control, attention focus and cognitive strategy. There were five interface parameters that could be adapted based on the user characteristics considered: Access to working processes, type of dialogue, level of help instructions, data representation form and color palette. The results were as followed (Table 3):

Table 3. Interface parameters and correlated human factors.

Interface Parameter	Correlated Human Factor
Access to working processes	Learning abilities, Attention focus
Type of dialogue	Learning abilities, Locus of control, Attention focus
Level of help instructions	All except gender
Data representation form	Verbal IQ, Nonverbal IQ, Gender
Color palette	Locus of control, nonverbal IQ

In similar fashion Biswas *et al* (Biswas, Bhattacharya, & Samanta, 2005) proposed a user model to be used in the design of personalized interfaces for motor impaired users taking into consideration certain related characteristics, such as demographic and cultural characteristics, experience, language and education levels, keeping with the objective of the model being application independent and clustering users according to their characteristics using the individual and clusters of users as a basis for adaptation.

In (Biswas, Robinson, & Langdon, 2012) to develop more inclusive interfaces, they simulated how different users with different impairments interact with the interface. That way developers can have a better understanding of what the users see and feel when using an application. They created a simulator with the following components:

- The application model. Represents the task at hand, broken up using the KLM model (Card et al., 1983).
- The interface model. Decides the input and output devices by the user and sets the parameters for an interface.
- The user model. Simulates the interaction of the users for the task configured in the interaction model. It uses the sequence of phases of the Human-Interaction Processor (Card et al., 1983).

The simulation works based on their previous work of a perception model (Biswas & Robinson, 2009a), by simulating the phenomenon of visual perception of users with different visual acuity. The model also simulates the effects of several visual impairments, such as color blindness, macular degeneration, diabetic retinopathy, among others. The cognitive model (Biswas & Robinson, 2008) is used to break up the task into a set of atomic tasks that the application will perform. And lastly, the motor behavior model (Biswas & Robinson, 2009b) in which they use to predict pointing time, obtained by a statistical analysis of cursor traces from motor-impaired users. With all these models, a simulation of several users with varying characteristics can be done so that one can directly perceive the HCI as well as they do.

Tarpin-Bernard & Habieb-Mammar (2005) developed the Cognitive User Modelling for Adaptive Presentation Of Hyper-Documents (CUMAPH) environments, which aims to adapt web pages by adapting the presentation of said web pages to the user's cognitive profile. The profile is formed by four sectors (attention, memory, language and visuospatial) involving cognitive indicators to form a cognitive profile:

- Verbal/Visual/Musical working memory
- Verbal/Musical short-term memory
- Verbal/Visual/Musical long-term memory
- Form recognition
- Categorization
- Understanding
- Glossary
- Spatial Exploration

To adapt web pages to the cognitive profile of the user, they divide each page into different components that depending on the values of the indicators will adapt their content. The perceptual/cognitive components are:

- Visual component.
- Audio/Musical component.
- Kinesthetic component.

- Language component.

With the cognitive profile of the user at hand, an algorithm is applied where the optimal multimedia elements are selected for the web page by calculating averages of the cognitive indicators and a compatibility factor, obtaining as a result an adapted web page to the user's cognitive profile. The results for the experiment showed better user performance for those with a lower cognitive profile, since users with a high cognitive profile can better adapt to a non-adapted interface. Even with positive results to the adaptation, the how and why the decisions for the adaptations are not clear, there is no mention on how they determined the adaptations needed for the different measurements in the cognitive profile of the user.

Kaklanis *et al* (2014) presents one of the most recent advances towards a standardization of a user model to be used across different platforms for simulation and adaptation purposes. The VUMS (Virtual User Modelling and Simulation Standardization) cluster of projects that works toward the development of an interoperable user model as a generic solution for the modeling for non-disabled users and users with disabilities. Currently, the model includes a myriad of user attributes but for the moment no actual adaptation rules based on the attributes. The VUMS cluster aims to develop:

- A standard user model able to describe older people and people with disabilities.
- Common data storage format for user profiles.
- Common calibration and validation techniques.
- Collaboration on ethical issues between researchers.
- Ensuring sustainability by making the user attributes available within a standard.

Adaptive User Interfaces applied for special need users

Because Adaptive User Interfaces (AUI's) can adapt to the user characteristics, they can be used in software systems designed for special needs users, but for some reason, AUI's are not widely adopted for this use.

One example is the AVANTI project (Stephanidis et al., 1998) in which an AUI was developed for web documents, adapting itself to the user to some extent based on some information submitted by the user, such as language, eyesight, motor abilities, language and experience with similar applications. It also adapts itself as they interact with the system based on user familiarity with specific tasks, ability to navigate, error rate, disorientation, user idle time, and repetition of interaction patterns. The adaptation also supports multi-modal systems by featuring integrated support to various input and output devices with appropriate interaction techniques to accommodate users with disabilities. The UI supports users with light to severe motor disabilities and blindness.

Also research is being made toward improving accessibility for all users, especially older users, using adaptive and adaptable interfaces and multi-modal interaction. Although it is still considered that there is much work to be done before a definitive methodology for the development of said systems and for the different measures to improve accessibility to be adopted (Jorge, 2001).

User Interface Adaptation

In (Gajos & Weld, 2004) is presented an automatic personalized user interface generator called SUPPLE. With SUPPLE it is possible to generate interfaces adapted to the user's devices, tasks, and preferences. Also, there is case study presented where it also considers some user characteristics, in this case, motor skills, by developing a usage model based on users with motor deficits. They define interface generation as an optimization problem and demonstrate that is computationally feasible. Adaptation is based on a device model for the limitations of the device for which an interface is generated and a usage model to represent user preferences, in conjunction with a cost function that is used as a way to calculate the current effort of use.

In (Thevenin & Coutaz, 1999) a new property of interactive systems was introduced for user interface adaptation. Plasticity is defined as the capacity of a user interface to withstand variations of both the physical system characteristics and the environment while preserving usability. A model-based framework is proposed to develop plastic interfaces mostly based on the context of use, meaning the user task model and the system task model. In later work (Calvary et al., 2002) a unifying framework is proposed where they decompose the context of use into the end users, the hardware and software, and the physical environment where they are working. Although the authors consider the user as part on the context of use, there is no real mention of what user characteristics are to be considered and what kind of adaptation is to be expected, as of now it mostly centers on the device and the task that the user must accomplish.

As we can see, there has been work on work on improving usability by integrating HCI and usability engineering practices in the software development lifecycle. Frameworks were proposed for the generation of usable user interfaces, mostly focusing on the tasks that the user must accomplish with some integrating some user modeling aspects. There is significant advancement towards a standardized user model, but there is still much work to be done before we can have a generic solution with a standard rule set for each user characteristic.

Chapter III:

Proposed Solution

Interaction Rules based on Executive Function measurements

"You can't control what you can't measure."

- Tom DeMarco

Proposed Solution: Interaction Rules based on Executive Function measurements

Different combination of GUI components and interface design patterns can cause varying levels of cognitive load depending on the user's mental capacities, which can vary greatly. Adapting to these very different and varying capacities can be a daunting task, making designing a usable software across a varied user base extremely difficult to achieve, such is the case with specialized software for users with special needs.

There is no consensus in to what interface elements can be adapted, what design patterns or what user characteristics to consider about the user and how it affects the adaptivity of the interface. Each developer stipulates his own rules. This is particularly true when designing software for users with some form of disability.

Also, cognitive load continues to be one of the main problems when designing a usable interface, without an easy to use metric or a cognitive load score to interface design patterns it is hard to determine what level of cognitive load it produces to the user, since cognitive load depends on several factors including the user's cognitive capabilities.

Based on a previous work (Figueroa, Juárez-Ramírez, Inzunza, & Valenzuela, 2014; Mejía, Juárez-Ramírez, Inzunza, & Valenzuela, 2012), where we established a relation of the tasks to be done in a software and user characteristics, we also need to establish a relationship with user interface elements to develop better design decision for the interface.

As a first step, we must establish the characteristics that the user needs to use properly a software system. Based on several of the cognitive architectures and models in HCI (Card et al., 1983; D. E. Kieras & Meyer, 1997) we can establish that a user (*U*) has a set of characteristics needed for software interaction:

$U = \{Sn, Cg, Mp\}$

$Sn = \{Sn_1, \ldots, Sn_n\}$

$Cg = \{Cg_1, \ldots, Cg_n\}$

$Mp = \{Mp_1, \ldots, Mp_n\}$

Where *Sn* represents the set of the senses, such as eyesight, hearing, and touch; *Cg* represents cognitive functions, such as memory and attention; and *Mp*, which represents motor function for different parts of the body.

In our previous work, we defined software (*S*) as a set of functionalities (*F*) operated by the user, which also includes a user interface. This can be expressed as follows:

$S = \{\{F_1, \ldots, F2\}, I\}$

Which can in turn functionality can also be considered tasks (*T*) that are possible to do with the given software.

$S = \{\{T_1, \ldots, T_2\}, I\}$

A software also includes a graphical user interface (*I*) for which the user interacts with the system, which can be defined as follows:

$I = \{Uip_1, \ldots, Uip_n\}$

$Uip = \{\{Cp\},\{C\}\}$

Where an Interface (*I*), is a set of user interface patterns. A user interface pattern can be defined as a set of GUI components and/or interaction styles (*Uip*) for addressing a particular interaction problem, which in turn also need certain user characteristics (*C*) to operate said pattern. With this said, for each task there are user interface patterns that can achieve the task:

$\forall T : Uip$

The task (*T*) involves a set of actions executed by the user. This is expressed as follows:

$T = \{Ac_1, \ldots ,Ac_n \}$

Each action (*Ac*) can be of one of the following types: input, indication or interpretation. An input action involves operating an input device, such as mouse, keyboard or microphone; with the appropriate user attribute, such as hands and voice. An indication action also involves using the same user attributes as input, but also involves perception, attention, and information processing. A perception action involves employing the senses such as eyesight, hearing or touch. An information processing action involves employing cognitive function such as working memory and cognitive processor. Considering these assumptions, an action can be expressed as follows:

$Ac = \{t,\{C\}\}$

Where the type of action (t), and a set of user attributes (C) employed to interact with the software application and needed to accomplish the action. Examples of user attributes are: Vision, hearing, motor function, working memory, etc. Based on this, the task or functionality T can be expressed as follows:

$$T = \{ \{t,\{C\}\}_1, \ldots ,\{ t,\{C\}\}n \}$$

As we previously defined, executive function (Ef) is a term used to describe cognitive processes necessary for the completion of goal-oriented behavior. As such we can establish:

$$Ef = \{ A,\{Cg\}\}$$

Where A is the ability needed, such as planning, working memory, cognitive flexibility, and Cg the set of cognitive processes needed for that particular ability.

In this case, we are analyzing executive functions and the effect on interaction. As such we can express as task as follows:

$$T = \{ \{t,\{A \{Cg_1, \ldots ,Cg_n\} \}\}_1, \ldots , \{ t,\{A \{Cg_1, \ldots ,Cg_n\} \}\}_n \}$$

So, in order for the user to be able to use the software, he must have certain characteristics needed to use the software's user interface. So, we can assume:

$$(\forall U \cap A) \exists Uip$$

Which denotes that for each action, there is at least one user interface pattern for the particular user with particular characteristics.

To better define what to do with different user characteristics we conducted a study where we obtained interaction rules based on measurements of executive functions to choose better interface design patterns based on the executive function measurements and usability testing metrics with a wide variety of users. While also analyzing the data to see the relationship between the cognitive load produced by each interface pattern and the executive functions.

The expected rules should conform to the following structure. The idea is to have a more precise rule or guideline for the selection of interface design patterns to be used in a given software. The structure of the rules should be similar to the one that follows:

IF <working memory meaurement> IS HIGHER | LOWER THAN <wmvalue>& <cognitive flexibility measurement> IS HIGHER | LOWER THAN <cfvalue> THEN USE <interface design pattern>

Where *working memory measurement* represents one of the working memory neuropsychological tests and *wmvalue* the score of said test. Similarly, *cognitive flexibility measurement* represents one of the cognitive flexibility neurophychological tests and *cfvalue* the score of said test. Depending on the values of the tests scores the rule returns *interface design pattern*, which is the interface pattern with the best usability scores for that case.

Chapter IV:

Experimental design

"I love fools' experiments. I am always making them"

- Charles Darwin

Experimental Design

To determine the rules or guidelines we set off to gather data from children with different cognitive characteristics, and extensive usability tests of several interface patterns. The research questions that motivated the research were:

Research questions

We have two main research questions:

1. What interface design patterns are better suited for different levels of executive functions (In this case working memory and cognitive flexibility)?
2. What is the relationship between cognitive load caused by user interface design patterns and the different levels of working memory and cognitive flexibility?

Hypothesis

With interaction rules, based on measurements of executive functions (cognitive flexibility and working memory) we can improve usability for users with deficiencies in executive functions by determining the relationship between executive function measurements, usability patterns, and cognitive load, and adapt the interface by presenting a more usable interface design pattern to the user.

Experimental Design

Objectives

The purpose of the experiment is to determine if there is a relationship between executive functions measurements (working memory and cognitive flexibility), GUI design patterns and cognitive load with the use of neuropsychological tests, while also determine possible interaction rules based on the measurements of executive functions and usability testing results.

Sample

Group of 105 children, ages 7-12, composed of three subgroups: 35 children with Autism Spectrum Disorder (ASD) level and 35 children with ASD level 2 according to the DSM-5 (Association, 2013); and 35 typical children. Children with ASD were chosen because of the impairments in the key executive functions (working memory and cognitive flexibility) as a symptom of Autism (Hill, 2004b).

Structure of study

The experiment is divided into three main phases:

- Phase 1: Application of all neuropsychological tests to the children needed to evaluate the executive functions of working memory and cognitive flexibility.
- Phase 2: Usability testing of eight prototypes of an Augmentative and Aumentative Communication (AAC) app using the Picture Exchange Communication System (PECS), each one using a different interface design pattern. The interface design

patterns selected were: Accordion, carousel, scrolling menu, navigational drawer, tabs, one-window drilldown, two-panel select, and wizard. The task was the same for each one of the prototypes, form a three-part sentence, for example, I want cookie.

- Phase 3: Analysis of the data, with different objectives. The first objective is to obtain a ruleset using Classification Trees with the target being the user interface design pattern based on the executive function measurements; usability testing results are used here. The second objective is to find out if there is a direct influence of the executive functions on the cognitive load produced when using each of the applications, to see if there is a meaningful statistical significance using regression analysis.

Variables

Measurements of the following were taken of all participants:

- Working memory. A temporary system where we can store and manipulate information in the short-term memory. The WISC-IV intelligence test was used for measurement.
- Cognitive flexibility. Ability to shift to a different thought or action in response to a situation change. The neuropsychological battery NEPSY II was used to measure cognitive flexibility.
- Cognitive load. Mental resources used by the task at hand. Measured by using a modified version of the NASA-TLX for children.

Usability testing metrics considered are:

- Binary success. Metric with two possible values (1 or 0) which denotes if the user completed the task or not.
- Error rate. Number of errors per task.
- Time on task. Time needed to complete the task, measured in seconds.
- Level of success. Subjective value based on the observations of the usability testing evaluators. It has a range of 1 to 4, where 1 represents no problems to complete the task, 2 represents minor problems, 3 represents major problems, and 4 represents that the user was not able to complete the task.

Interface design patterns:

- Pattern 1: Accordion
- Pattern 2: Carrousel
- Pattern 3: Scrolling menu
- Pattern 4: Navigational Drawer
- Pattern 5: Tabs
- Pattern 6: One-Window Drilldown
- Pattern 7: Two-panel select
- Pattern 8: Wizard

Instruments

NEPSY II

NEPSY II (Korkman, Kirk, & Kemp, 2007) is a set of neuropsychological test that is used in different combinations to evaluate the neurological development of children from the ages of 3 to 16 years in six domains. In this case, we are only going to center on the executive function domain.

The attention and executive function domain is composed of the tests of animal sorting, auditory attention, response set, clocks, design fluency, inhibition, and statue. These subcomponents evaluate several executive functions, such as sustained and selective attention, set shifting or cognitive flexibility, planning, inhibition and auto-regulation and monitoring.

The tests used from the NEPSY II attention and executive function domain are the ones related to cognitive flexibility or set shifting:

- Animal Sorting (AS). This test is designed to evaluate the ability to formulate basic concepts, transfer them into actions and change the focus of his attention. The participant classifies cards into two groups, each one with different criteria defined by the child. The test has no maximum score value, but a test score of 8 is considered very high.
- Auditory Attention (AA) and Response Set (SS). Auditory attention evaluates selective auditory attention and to sustain it. Response set assesses the ability to set shift and maintain new conditions by inhibiting previously learned responses and

correctly responding to new stimuli. The child listens to a list of colors and touches the appropriate circle of color when he or she hears the target word. Both tests have a maximum score of 30.

WISC-IV

WISC-IV (Wechlser, 2003) is an intelligence test administered to children between the ages of 6 and 16 years of age. It generates a general intelligence coefficient (IQ) and five primary IQs, of which for this study we will be focusing on the tests needed for the working memory index.

The working memory index is the one in charge of evaluating the storage and retention of information and its manipulation. It is composed of two sub-tests:

- Digit Span (DS). Analyzes short-term memory, which gives us an idea of his sequencing abilities, planning ability, and cognitive flexibility.
- Letter-Number Sequencing (LN). Analyzes memory retention capacity and the ability to combine different types of information, organize said information and elaboration of organized groups based on the same information. Both digit span and letter and number sequencing have a maximum score value of 16.

Modified NASA-TLX test for children

An adapted version of the NASA Task Load Index for children (Laurie-Rose, Frey, Ennis, & Zamary, 2014) for measuring mental workload, also referred as cognitive load. NASA-TLX is a subjective assessment of mental workload of a perceived task divided into six sub-scales: mental demand, physical demand, temporal demand, performance, effort, and frustrations.

Usually, the original NASA-TLX used Linkert scales for each of the sub-scales, ranging from very low to very high, with a range of values from 0 to 100 with increments of five. In the modified version, a physical ruler was used with a cartoon on each side of the scale representing both extremes of the subscale being measured. For example, for the mental demand subscale, on the least demanding side, a cartoon of a relaxed child can be seen doing homework, while on the other extreme a very tired cartoon of a student doing homework can be seen (Figure 2).

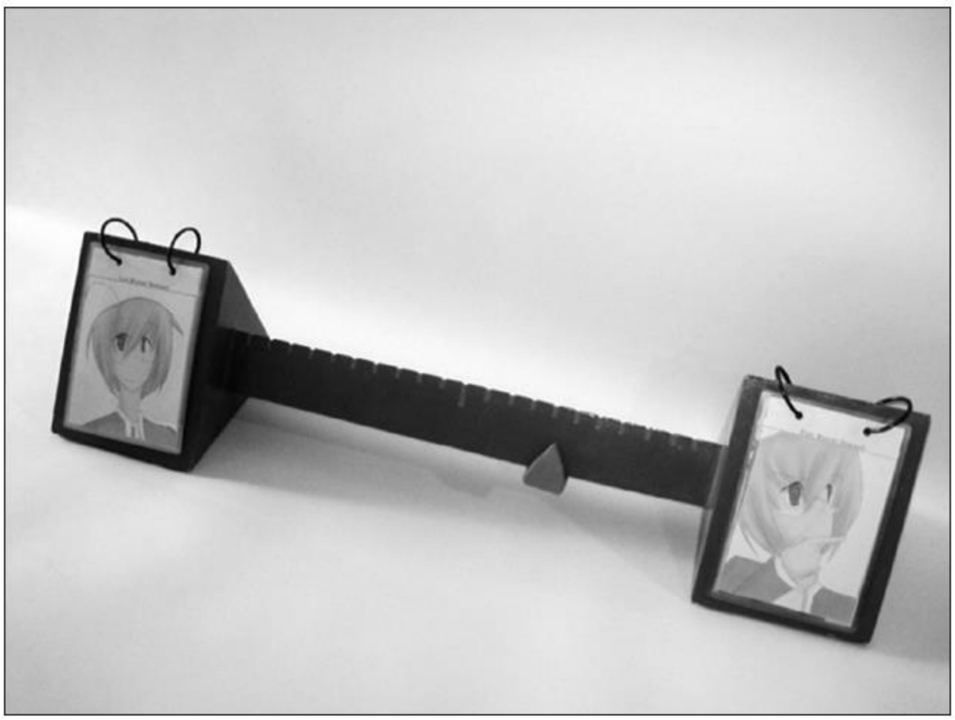


Figure 2. Ruler used for the mental workload measurements, figure extracted from (Laurie-Rose et al., 2014).

It is also accompanied by an extended dialog and questions that the examiner must explain before each of the scales is responded by the child. Both tests return a single value in the 0-100 range for mental workload of the given task.

Methodology

Measurements of working memory and cognitive load are taken using the WISC and NEPSY tests. The values obtained are not transformed and standardized for their age according to the normal procedures of evaluation using those tests. The values used are the number of right answers of each test to properly observe the variance of the results between the subjects. If the results were standardized, it would have been a result based on the age of each of the participants. For example, it is not the same a score of 12 in a test for a participant of an age of 7 years than that same score for a participant of 12 years of age.

After measurements of the executive functions are taken, the participants proceed to usability testing of eight prototypes of an AAC application using the PECS system on an Android tablet. The task is to form a simple three-part sentence in each prototype. Each session is videotaped for analysis, and the usability metrics are obtained. The order of the testing was counterbalanced and spreaded out in different days depending on the availability of the participants. Each usability test took about around 5 minutes per prototype.

After each usability test, the modified NASA-TLX test was applied, and a cognitive load score was obtained for each prototype tested for each participant.

The neuropsychological tests and modified NASA-TLX test were applied by psychology students and trained personnel from the schools, and the usability testing carried out by interaction design students.

For the research question of which interface design patterns are better suited for different levels of working memory and cognitive flexibility, we used classification trees to obtain rules with the target on the best interface design pattern based on the results of the executive function testing.

The main dataset was divided into eight minor datasets based on the results of the usability testing, for example, one dataset consists of the best performing registry of every participant, another one of the second best performing registry, and so on until we reach the eight datasets consisting of the worst performing registry of each participant. The main usability metric used to classify the minor datasets was the level of success, and a mixture of error rate and time on task.

Classification trees were used on each dataset, and a ruleset was obtained, which was subsequently tested with a newer and smaller dataset consisting of 30 randomized children which did not participate in the composition of the original dataset.

To find the relationship between cognitive load and the executive functions we transformed the data and performed a repeated measures ANOVA and a regression analysis for each design pattern used.

Chapter V:

Experimental Results and Discussion

"Science, my boy, is made up of mistakes, but they are mistakes which it is useful to make because they lead little by little to the truth."

- Jules Verne, Journey to the center of the Earth

Experimental Results

Data analysis

The dataset obtained consisted of eight registries per participant, each registry made of the five results of the executive functions measurements, the cognitive load score, the interface design pattern, and the usability testing metrics, for a total of 840 registries.

All executive function measurements showed signs of high collinearity between them (Correlation of minimum .716 and a maximum of .971), and VIF values ranging from 2 to 5 depending on the regression being performed.

For the repeated measure ANOVA and the regression analysis, it was necessary to perform data transformations to comply with the normality of residuals and the independence assumptions since the data failed the Shapiro-Wilks test and showed a high correlation between variables. To deal with collinearity, we used Principal Component Analysis on the executive function measurements which yielded one component for the five variables that explained 70.325% of the variance and for the normality of residuals we transformed the data by using Box-Cox transformation.

Usability Testing Results

Binary Success

The results of the binary success metric for the different prototypes for the three groups are the following:

For the group consisting of typical children (Table 4 and Figure 2):

Table 4. Results of binary success for the group with no ASD.

Interface Design Pattern	Success rate	Confidence interval
Pattern 1: Accordion	82.86%	27.00%
Pattern 2: Carrousel	94.29%	18.00%
Pattern 3: Scrolling menu	85.71%	25.00%
Pattern 4: Navigational drawer	82.86%	27.00%
Pattern 5: Tabs	94.29%	18.00%
Pattern 6: One - window drilldown	82.86%	27.00%
Pattern 7: Two - panel select	91.43%	21.00%
Pattern 8: Wizard	88.57%	24.00%

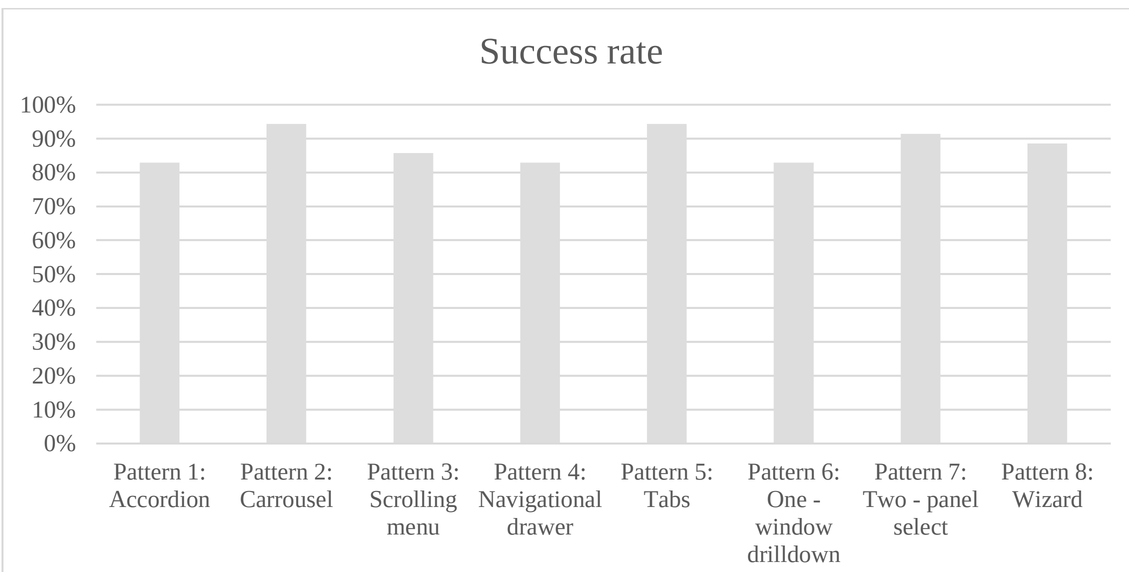


Figure 3. Success rate for group with no ASD.

The results show a mostly even success rate among the interface patterns, with an average of 87.85% completion rate. The lowest success rate of the patterns was the Wizard pattern.

For the group consisting of children with ASD Level 1 (Table 5 and Figure 3):

Table 5. Success rate for group with ASD level 1.

Interface Design Pattern	Success rate	Confidence Interval
Pattern 1: Accordion	100.00%	10.00%
Pattern 2: Carrousel	100.00%	10.00%
Pattern 3: Scrolling menu	100.00%	10.00%
Pattern 4: Navigational drawer	94.29%	18.00%
Pattern 5: Tabs	100.00%	10.00%
Pattern 6: One - window drilldown	100.00%	10.00%
Pattern 7: Two - panel select	100.00%	10.00%
Pattern 8: Wizard	88.57%	24.00%

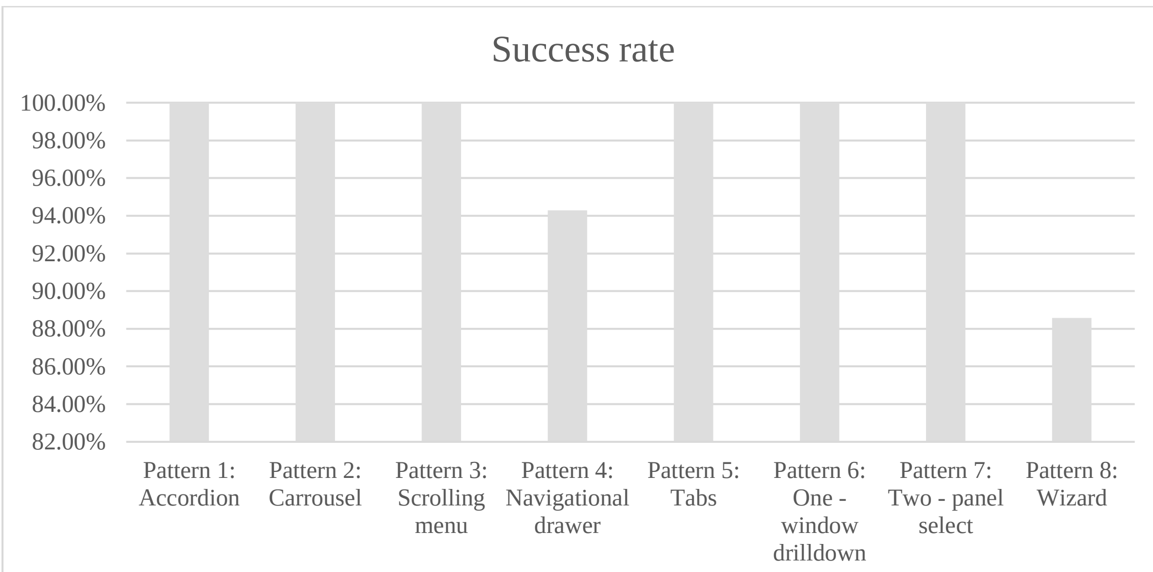


Figure 4. Success rate for group of ASD level 1.

In the ASD level 1 group the results showed a 97.85% average completion rate among the design patterns, even better than the group with no ASD, with the most trouble being the Wizard pattern also.

The results for the group consisting of children with ASD level 2 are (Table 6 and Figure 4):

Table 6. Success rate for group with ASD level 2.

Interface Design Pattern	Success rate	Confidence Interval
Pattern 1: Accordion	88.57%	24.00%
Pattern 2: Carrousel	34.29%	33.00%
Pattern 3: Scrolling menu	34.29%	33.00%
Pattern 4: Navigational drawer	31.43%	32.00%
Pattern 5: Tabs	34.29%	33.00%
Pattern 6: One - window drilldown	31.43%	32.00%
Pattern 7: Two - panel select	34.29%	33.00%
Pattern 8: Wizard	20.00%	28.00%

Figure 5. Success rate for group with ASD level 2.

Here we see a much more noticeable difference in the success rate between groups, since with a more severe level of Autism, there are more severe impairments in their executive functions.

The group with level 2 ASD performed much worse in most of the patterns compared to the other two groups, with an average of 38.94%. Still the worst performing pattern being the Wizard.

There was a statistically significant difference between the results of the binary success metric determined by one-way ANOVA with a result of $F(2,21) = 50.934$, $p < 0.05$.

Levels of Success

In the case of the group of children with no Autism, the results for level of success for each successful attempt were as follows (Table 7, table 8 and figure 5):

Table 7. Number of cases for each level of success.

Interface Design Pattern	Level of success			
	1	**2**	**3**	**4**
Pattern 1: Accordion	16	11	3	5
Pattern 2: Carrousel	25	8	0	2
Pattern 3: Scrolling menu	16	12	2	5
Pattern 4: Navigational drawer	13	13	3	6
Pattern 5: Tabs	27	4	2	2
Pattern 6: One - window drilldown	15	14	0	6
Pattern 7: Two - panel select	30	2	0	3
Pattern 8: Wizard	14	14	3	4

Table 8. Percentage of each level of success per prototype.

Interface Design Patterns	Level of success			
	1	**2**	**3**	**4**
Pattern 1: Accordion	45.71%	31.43%	8.57%	14.29%
Pattern 2: Carrousel	71.43%	22.86%	0.00%	5.71%
Pattern 3: Scrolling menu	45.71%	34.29%	5.71%	14.29%
Pattern 4: Navigational drawer	37.14%	37.14%	8.57%	17.14%

Interface Design Patterns	Level of success			
Pattern 5: Tabs	77.14%	11.43%	5.71%	5.71%
Pattern 6: One - window drilldown	42.86%	40.00%	0.00%	17.14%
Pattern 7: Two - panel select	85.71%	5.71%	0.00%	8.57%
Pattern 8: Wizard	40.00%	40.00%	8.57%	11.43%

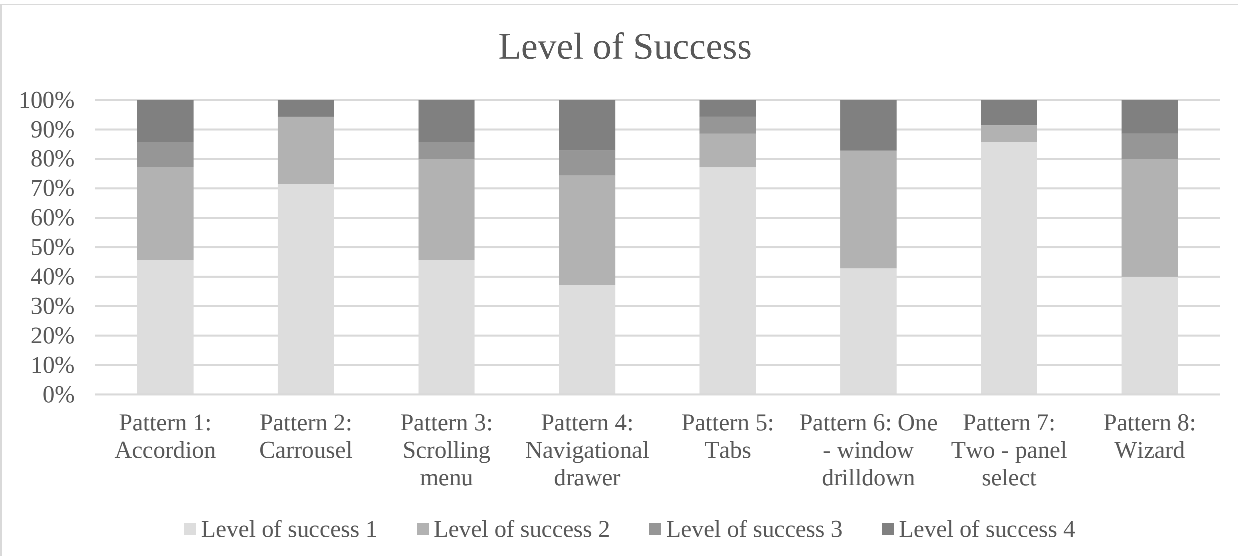


Figure 6. Level of success of group with no ASD.

As can be seen, the patterns with the best level of success were two-panel select and tabs, followed by carrousel.

For the group with Autism level 1, the results were the following (Table 9, 10, and figure 6):

Table 9. Count of level of success for group of ASD level 1.

Interface Design Pattern	Level of success			
	1	**2**	**3**	**4**
Pattern 1: Accordion	12	13	10	0
Pattern 2: Carrousel	28	7	0	0
Pattern 3: Scrolling menu	17	10	8	0
Pattern 4: Navigational drawer	5	19	9	2
Pattern 5: Tabs	27	8	0	0

Interface Design Pattern	Level of success			
Pattern 6: One - window drilldown	14	17	4	0
Pattern 7: Two - panel select	34	1	0	0
Pattern 8: Wizard	3	21	9	2

Table 10. Percentage of level of success for the group of ASD level 1.

Interface Design Pattern	Level of success			
	1	**2**	**3**	**4**
Pattern 1: Accordion	34.29%	37.14%	28.57%	0.00%
Pattern 2: Carrousel	80.00%	20.00%	0.00%	0.00%
Pattern 3: Scrolling menu	48.57%	2.99%	22.86%	0.00%
Pattern 4: Navigational drawer	14.29%	54.29%	25.71%	5.71%
Pattern 5: Tabs	77.14%	22.86%	0.00%	0.00%
Pattern 6: One - window drilldown	40.00%	48.57%	11.43%	0.00%
Pattern 7: Two - panel select	97.14%	2.86%	0.00%	0.00%
Pattern 8: Wizard	8.57%	60.00%	25.71%	5.71%

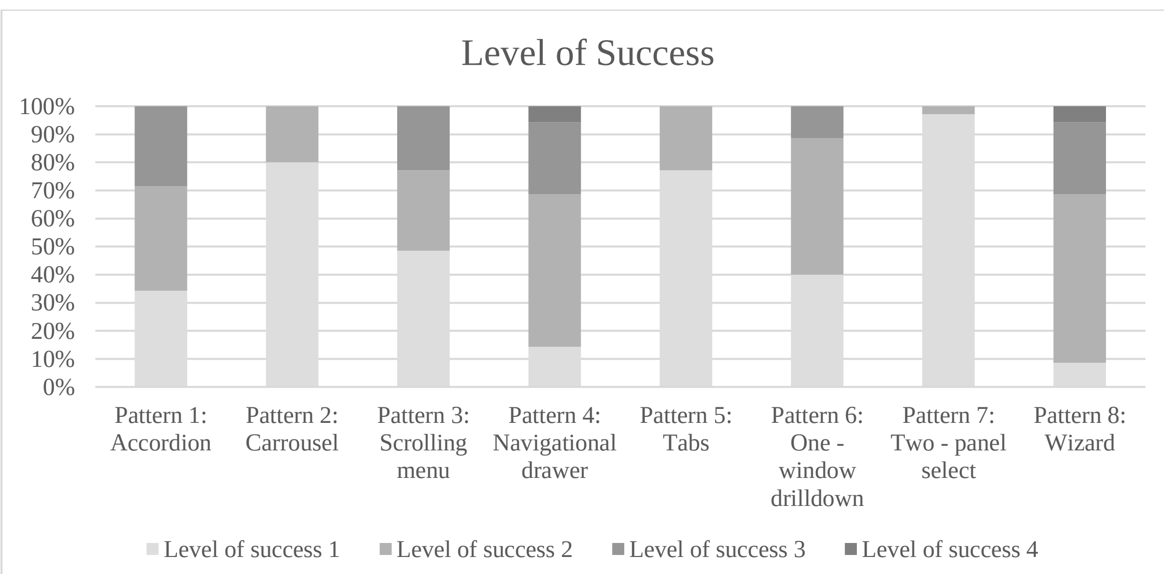


Figure 7. Level of success for group with Autism level 1

Here we see a similar pattern of the levels of success between the group with ASD level 1 and the group with no ASD, being the patterns with the best level of success the two-panel select, followed by tabs and carrousel.

And for the group with Autism level 2 (Table 11, 12, and figure 7):

Table 11. Level of success count of the group with Autism level 2.

Interface Design Pattern	Level of success			
	1	**2**	**3**	**4**
Pattern 1: Accordion	0	4	8	23
Pattern 2: Carrousel	5	5	2	23
Pattern 3: Scrolling menu	2	2	7	24
Pattern 4: Navigational drawer	0	1	8	26
Pattern 5: Tabs	8	3	1	23
Pattern 6: One - window drilldown	0	4	7	24
Pattern 7: Two - panel select	11	1	0	23
Pattern 8: Wizard	0	2	5	28

Table 12. Level of success percentage of the group of ASD level 2.

	Level of success			
	1	2	3	4
Pattern 1: Accordion	0.00%	11.43%	22.86%	65.71%
Pattern 2: Carrousel	14.29%	14.29%	5.71%	65.71%
Pattern 3: Scrolling menu	5.71%	5.71%	20.00%	68.57%
Pattern 4: Navigational drawer	0.00%	2.86%	22.86%	74.29%
Pattern 5: Tabs	22.86%	8.57%	2.86%	65.71%
Pattern 6: One - window drilldown	0.00%	11.43%	20.00%	68.57%
Pattern 7: Two - panel select	31.43%	2.86%	0.00%	65.71%
Pattern 8: Wizard	0.00%	5.71%	14.29%	80.00%

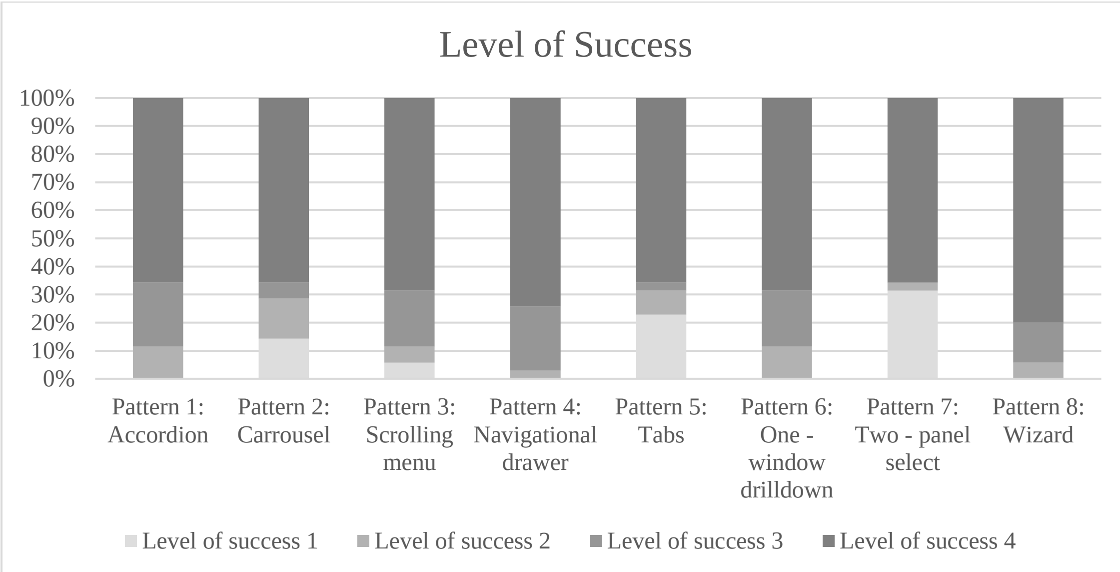


Figure 8. Level of success of group of ASD level 2.

Here we see an overwhelming level of success score of four in all of the patterns, bit even with that overwhelming bad score we see that the few that managed to complete the task still present a better result on the two-panel select, tabs, followed by carrousel.

Time on task

For each group, the time on task for each prototype was analyzed. The time considered is only in the cases where the users completed the task (N).

In the case of the group of typical children with no autism, the results are as follows (Table 13):

Table 13. Time on task for all interface patterns for group with no ASD.

User Interface Pattern	Average	Median	Geometric mean	Upper bound	Std. Deviation	Lower bound	Confidence Interval	N
Pattern 1: Accordion	73.06	75	60.21	87.1	38.55	59.03	14.03	29
Pattern 2: Carrousel	53.36	43	39.17	66.98	39.92	39.74	13.62	33
Pattern 3: Scrolling menu	69.56	58	57.07	84.3	41.18	54.82	14.73	30
Pattern 4: Navigational drawer	82.83	90	69.36	99.54	45.94	66.1	16.72	29
Pattern 5: Tabs	45.42	32	32.5	57.64	35.81	33.2	12.21	33
Pattern 6: One - window drilldown	59.34	57	50.93	71.29	32.84	47.39	11.95	29
Pattern 7: Two - panel select	30.06	18.5	21.86	39.15	26.24	20.96	9.09	32
Pattern 8: Wizard	58.7	55	48.91	70.22	32.7	47.19	11.51	31

Here the pattern repeats itself, the best performing patterns with the less time to complete the task were two-panel select and tabs.

Time on task for the group with Autism level 1 was (Table 14):

Table 14. Time on task for all interface patterns for group with ASD level 1.

User Interface Pattern	Average	Median	Geometric mean	Upper bound	Std. Deviation	Lower bound	Confidence Interval	N
Pattern 1: Accordion	85.74	87	74.36	99.87	42.66	71.6	14.13	35
Pattern 2: Carrousel	43.47	45	35.74	52.42	26.2	35.06	8.68	35

User Interface Pattern	Average	Median	Geometric mean	Upper bound	Std. Deviation	Lower bound	Confidence Interval	N
Pattern 3: Scrolling menu	70.74	62	62.48	81.94	33.81	59.54	11.2	35
Pattern 4: Navigational drawer	69.6	62	62.2	81.54	34.99	57.66	11.93	33
Pattern 5: Tabs	36.11	36	31.08	42.13	18.18	30.09	6.02	35
Pattern 6: One - window drilldown	59.14	57	54.94	67.36	24.81	50.92	8.21	35
Pattern 7: Two - panel select	26.48	28	23.1	30.54	12.25	22.42	4.05	35
Pattern 8: Wizard	59.84	58	53.47	69.35	27.86	50.34	9.5	33

For the group of Autism level 2 is as follows (Table 15):

Table 15. Time on task for all interface patterns for group with ASD level 2.

User Interface Pattern	Average	Median	Geometric mean	Upper bound	Std. Deviation	Lower bound	Confidence Interval	N
Pattern 1: Accordion	103.91	97.5	98.07	115.34	34.49	92.48	11.42	12
Pattern 2: Carrousel	73	48	53.69	94.07	37.25	51.92	21.07	12
Pattern 3: Scrolling menu	72.45	66	64.19	98.02	43.27	46.88	25.27	11
Pattern 4: Navigational drawer	80.88	70	71.91	109.61	43.96	52.16	28.72	9
Pattern 5: Tabs	43.33	37	36.44	58.29	26.43	28.37	14.95	12
Pattern 6: One - window drilldown	89.8	85	79.98	117.86	47.48	61.73	28.06	11

User Interface Pattern	Average	Median	Geometric mean	Upper bound	Std. Deviation	Lower bound	Confidence Interval	N
Pattern 7: Two - panel select	29.41	24.5	25.83	39.67	18.11	19.16	10.25	12
Pattern 8: Wizard	46.28	36	42.23	63.15	22.77	29.41	16.86	7

Error rate

The results of the error rate metric are as follows. For the group of typical children with no Autism the results were. Only cases where the task was successfully completed were considered (N) (Table 16):

Table 16. Error for all interface patterns for group with no ASD.

User Interface Patterns	Average	Std. Deviation	N
Pattern 1: Accordion	1.58	1.88	29
Pattern 2: Carrousel	0.66	0.87	33
Pattern 3: Scrolling menu	1.433	0.98	30
Pattern 4: Navigational drawer	1.82	1.01	29
Pattern 5: Tabs	0.81	1.31	33
Pattern 6: One - window drilldown	1.41	0.89	29
Pattern 7: Two - panel select	0.375	0.819	32
Pattern 8: Wizard	1.54	377	31

Results for the error rate metric for the group with ASD level 1 (Table 17):

Table 17. Error for all interface patterns for group with ASD level 1.

User Interface Patterns	Average	Std. Deviation	N
Pattern 1: Accordion	3.085	2.54	35
Pattern 2: Carrousel	0.885	1.73	35
Pattern 3: Scrolling menu	2.34	1.97	35
Pattern 4: Navigational drawer	3.09	2.16	33
Pattern 5: Tabs	0.8	1.008	35
Pattern 6: One - window drilldown	2.08	1.18	35
Pattern 7: Two - panel select	0.4	0.59	35
Pattern 8: Wizard	3.03	1.78	33

Results for the error rate metric for the group with ASD level 2 (Table 18):

Table 18. Error for all interface patterns for group with ASD level 2.

	Average	Std. Deviation	N
Pattern 1: Accordion	7.5	6.51	12
Pattern 2: Carrousel	2	0.81	12
Pattern 3: Scrolling menu	5.63	3.33	11
Pattern 4: Navigational drawer	8.22	6.1	9
Pattern 5: Tabs	2	2.27	12
Pattern 6: One - window drilldown	7.8	3.95	11
Pattern 7: Two - panel select	0.75	0.72	12
Pattern 8: Wizard	4.71	2.05	7

Cognitive load

Cognitive load measurements were taken for each prototype, but not all participants were able to reliably take the test for cognitive load. The ones that were unable to properly take the test were the participants of the group with ASD level 2 since they could not properly follow the instructions. The results shown here are from the participants that were able to complete the cognitive load test, represented by the letter N.

The results for the group of participants with no ASD were (Table 19):

Table 19. Cognitive load test results for all interface patterns for group with ASD level 0.

Interface Patterns	Average	Median	Geometric mean	Upper bound	Std. Deviation	Lower bound	Confidence Interval	N
Pattern 1: Accordion	26.1	22	22.31	32.06	16.38	20.13	5.96	29
Pattern 2: Carrousel	17.21	15	14.64	20.81	10.56	13.6	3.6	33
Pattern 3: Scrolling menu	24.93	22.5	23.18	28.48	9.91	21.38	3.54	30
Pattern 4: Navigational drawer	31.93	29	29.42	36.501	12.773	27.36	4.57	30
Pattern 5: Tabs	13.33	11	11.02	16.09	8.1	10.56	2.76	33
Pattern 6: One - window drilldown	24.44	22	21.79	28.17	10.91	20.72	3.72	33
Pattern 7: Two - panel select	9.71	8	8.26	11.86	6.19	7.57	2.14	32
Pattern 8: Wizard	24.45	25	20.67	29.06	13.11	19.83	4.61	31

For the group of ASD level 1 the results were (Table 20):

Table 20. Cognitive load test results for all interface patterns for group with ASD level 1.

	Average	Median	Geometric mean	Upper bound	Std. Deviation	Lower bound	Confidence Interval	N
Pattern 1: Accordion	33.59	30	31.52	37.66	12.1	29.52	4.06	34
Pattern 2: Carrousel	20.06	20	16.35	24.07	11.92	16.05	4.008	31

	Average	Median	Geometric mean	Upper bound	Std. Deviation	Lower bound	Confidence Interval	N
Pattern 3: Scrolling menu	25.34	22.5	23.25	28.85	10.13	21.83	3.51	32
Pattern 4: Navigational drawer	38.71	32	36.33	43.44	13.63	33.99	4.72	32
Pattern 5: Tabs	17.46	15	14.82	21.03	10.29	13.9		32
Pattern 6: One - window drilldown	33.71	33	31.59	37.78	11.72	29.65	4.06	32
Pattern 7: Two - panel select	11.93	10	10.33	14.1	6.24	9.77	2.16	32
Pattern 8: Wizard	34.65	30	32.7	39.35	13.57	29.95	4.7	32

Overall, the patterns with the least amount of perceived cognitive load on average were two-panel select and tabs, matching up with the usability test metrics.

Interaction Rules

We used classification trees on the datasets; the following rules were simplified from all the rules obtained from the trees. To validate the rules, we tested the accuracy by fitting them to another dataset consisting of 30 random participants in the same age range that did not participate previously.

Only the trees where a decent accuracy level of prediction are presented since some of the obtained trees show too much variance and don't give any useful information.

The tests for Auditory Attention (AA) and Response Set (SS) have a maximum score of 30. Animal sorting has no maximum value, but a value of around eight is considered very high. Digits Span (DS) has a maximum score of 16, so do Numbers and Letters (NL).

From the best performing interface patterns (Figure 8):

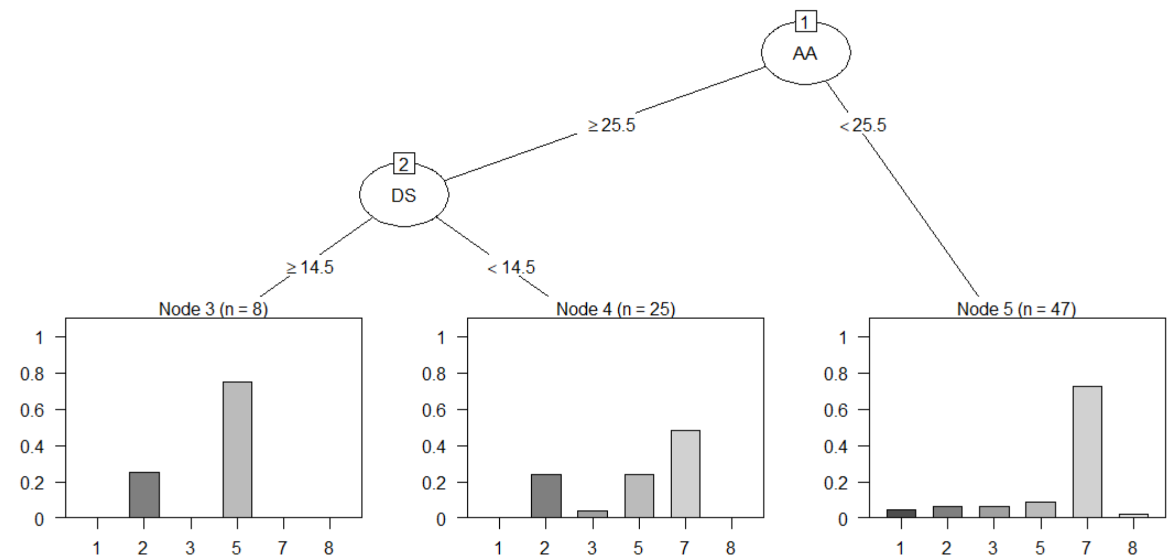


Figure 9. Decision tree for the best performing patterns based on usability testing scores.

IF AA < 25.5 THEN

(pattern1 4.3%, pattern2 6.4%, pattern3 6.4%, pattern5 8.5%, pattern7 72%, pattern8 2.1%)

IF AA >= 25.5 && DS >= 14.5

(pattern1 0%, pattern2 25%, pattern3 0%, pattern5 75%, pattern7 0%, pattern8 0%)

IF AA >= 25.5 && DS < 14.5 && SS<26.5

(pattern1 0%, pattern2 57%, pattern3 0%, pattern5 0%, pattern7 43%, pattern8 0%)

IF AA >= 25.5 && DS < 14.5 && SS>=26.5

(pattern1 0%, pattern2 11%, pattern3 5.6%, pattern5 33%, pattern7 5%, pattern8 0%)

In this case, we can see an overwhelming better result for the pattern 7, this being two-panel elect. We can see that users with a score of 25.5 or lower of Auditory Attention (AA), which is one of the tests we used related to cognitive flexibility, performed a lot better with two- panel select, closely followed by the tabs pattern when the users had a better AA score. AA turned out to be the most useful predictor and served as the main factor of the executive functions for the classification tree to determine segmentation.

Other influential user attributes were digit span (DS) and response set (SS). With a decent digit span score over 14.5, the users improve usability scores on several other patterns such as tabs and carousel. When considering set shifting users with a score of less of 26.5 performed better with tabs by a small margin (difference of 13%) instead of two-panel select. This ruleset showed a 53.33% on average accuracy when tested for accuracy with another dataset.

We can generalize with this tree that as a rule of thumb that users with a medium level to low auditory attention will perform better using the two-panel select, with tabs and carrousel closely behind when AA is a little higher than 25 and DS is around 14.

For the second best performing interface patterns (Figure 9):

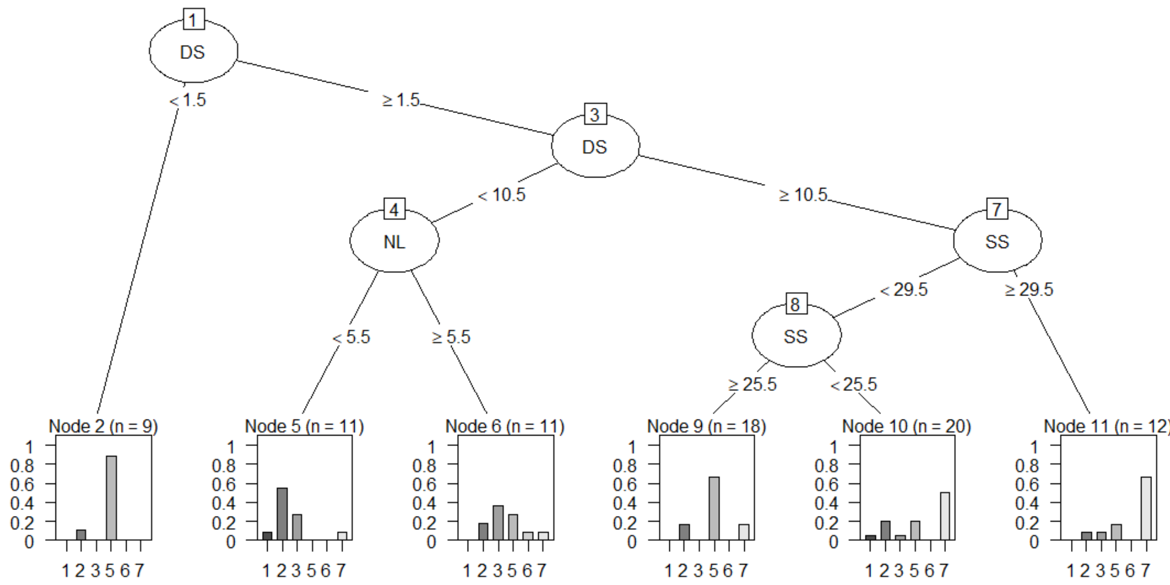


Figure 10. Decision tree for the second best-performing patterns based on usability testing scores.

IF DS < 1.5

(pattern1 0%, pattern2 1.1%, pattern3 0%, pattern5 89%, pattern6 0%, pattern7 0%)

IF DS >=1.5 && DS <10.5 && NL < 5.5

(pattern1 0.9%, pattern2 55%, pattern3 27%, pattern5 0%, pattern6 0%, pattern7 9.1%)

IF DS >=1.5 && DS <10.5 && NL >= 5.5

(pattern1 0%, pattern2 18%, pattern3 36%, pattern5 27%, pattern6 9.1%, pattern7 9.1%)

IF DS >=1.5 && DS >=10.5 && SS < 29.5 && SS >=25.5

(pattern1 0%, pattern2 17%, pattern3 0%, pattern5 67%, pattern6 0%, pattern7 17%)

IF DS >=1.5 && DS >=10.5 && SS < 29.5 && SS <25.5

(pattern1 5%, pattern2 20%, pattern3 5%, pattern5 20%, pattern6 0%, pattern7 50%)

IF DS >=1.5 && DS >=10.5 && SS >= 29.5

(pattern1 0%, pattern2 8.3%, pattern3 8.3%, pattern5 17%, pattern6 0%, pattern7 67%)

In the case of the second ones, we start to see more variance in the results. A user with a low score of 1.5 in the digit span test performed better with the Tabs pattern in 89% of cases. When digit span is greater than 1.5 and numbers and letters less than 5.5 we see that carrousel and scrolling menu have significant gains.

After this, we see the influence of Response set when it has a somewhat high value between 25.5 and 29.5 again the most usable pattern being Tabs. The accuracy obtained on average was 41.37%.

From here we jump to eighth place performing patterns trees since the other ones showed very low accuracy (values less than 20%) (Figure 10).

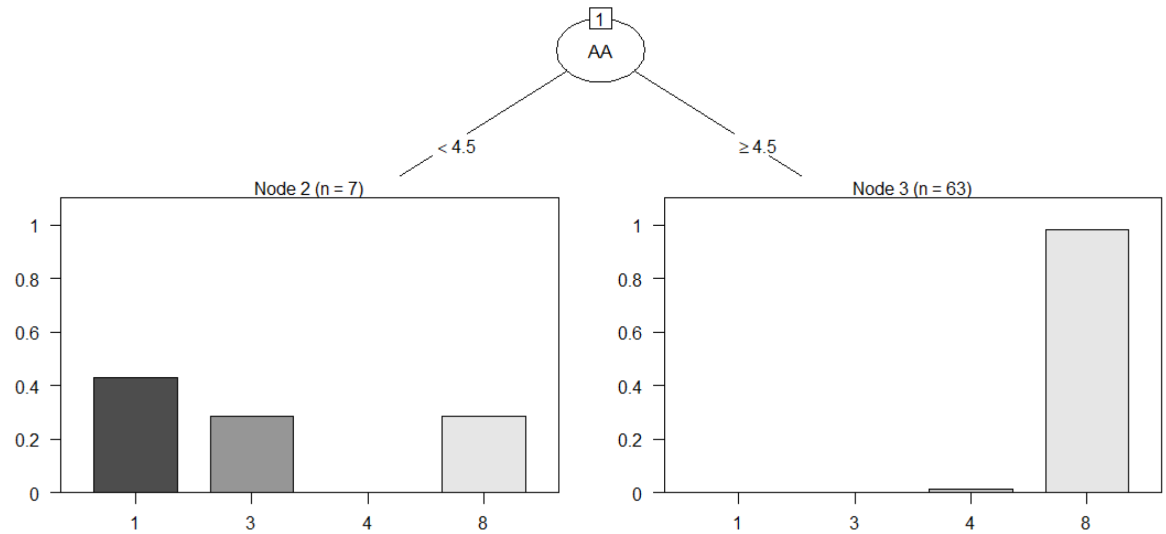


Figure 11. Decision tree for the eight best performing patterns based on usability testing scores.

IF AS < 4.5

(pattern1 42.8%, pattern3 28.5%, pattern4 0%, pattern8 28.5%)

IF AS >= 4.5

(pattern1 0%, pattern3 0%, pattern4 1.5%, pattern8 98%)

As can be seen in the case of the task that was considered the wizard pattern seemed to be the worst performing in most cases. The accuracy of this classification tree was a high 98.4%.

In summary, we can determine the best performing classification tree, that the great majority of users with a medium to low AA (AA<25.5) will perform better with two-panel select, followed closely by tabs when the user has an AA > 25.5, signaling that the tabs pattern might be slightly harder to use.

From the second best patterns we can see similar results while considering other metrics, being digit span the main one since those with a near score of 0 in digit span performed better using tabs, which was one of the best evaluated in the tests. Once the value of digit span incremented above 1.5, we start to see some variance between carrousel, scrolling menu, and tabs, with influence in part of by response set.

The surprise was the performance of the wizard pattern, which was overall the worst one, at least for this task. One theory is that since the participants had deficiencies in working memory they had a hard time remembering the previous step and committed mistakes, unlike two-panel select, they had at all times the information on the screen.

Analysis of Variance of Cognitive Load and Executive Functions

Applying a repeated measures ANOVA, the data violates the assumption of sphericity, to correct this issue we use the Huynh-Feldt correction (Epsilon = 0.868).

With the correction the mean scores for cognitive load for each interface design pattern were statistically significantly different with an $F(6.073, 364.36) = 64.355$, $p<0.05$.

When considering the executive function component as a covariate we can see a significant interaction $F(6.073, 364.36) = 7.369$, $p<0.05$. Which is expected since our mental resources are affected by cognitive load, in this case working memory and cognitive flexibility.

Regression Analysis

To see the impact of executive functions on the cognitive load produced by each of the patterns, after doing all the necessary data transformations we proceeded to do a linear regression analysis, where the executive function component as the independent variable, and the cognitive load from the modified NASA-TLX.

For the accordion pattern the results of the regression analysis were (Table 21):

Table 21. Regression analysis of the accordion pattern.

R	R Squared	Adjusted R Squared	Std. Error of the Estimate
.369	.136	.122	.37885

The regression equation with the coefficients being:

Cognitive Load = 2.742 – .149EFComponent

For this analysis, the cognitive load metric was transformed by a λ of 0.303 via Box-Cox transformation to achieve normality.

For the regression analysis of the carousel pattern the results were (Table 22):

Table 22. Regression analysis of the accordion pattern.

R	R Squared	Adjusted R Squared	Std. Error of the Estimate
.333	.111	.097	.12354

With the regression equation being:

Cognitive load = 1.477 – 0.43EFComponent

The cognitive load value was previously transformed with a λ of 0.1414 to achieve normality.

For the scrolling menu pattern, the regression analysis results were (Table 23):

Table 23. Regression analysis of the accordion pattern.

R	R Squared	Adjusted R Squared	Std. Error of the Estimate
.629	.396	.386	11.330

The equation being:

Cognitive load = 26.799 - 9.027EFComponent

The analysis for the navigational drawer yielded the following result (Table 24):

Table 24. Regression analysis of the accordion pattern.

R	R Squared	Adjusted R Squared	Std. Error of the Estimate
.595	.354	.343	13.915

With the equation being:

Cognitive load = 36.875 – 10.132EFComponent

For the tabs design pattern, we have the following regression result (Table 25):

Table 25. Regression analysis of the accordion pattern.

R	R Squared	Adjusted R Squared	Std. Error of the Estimate
.446	.199	.186	.02991

The regression equation being:

Cognitive load = .859 + .015EFComponent

For the one-windows drilldown pattern, the regression results were (Table 26):

Table 26. Regression analysis of the accordion pattern.

R	R Squared	Adjusted R Squared	Std. Error of the Estimate
.509	.259	.247	14.702

With the regression equation being:

Cognitive load = 31.515 – 8.562EFComponent

For the two-panel select, the regression results were (Table 27):

Table 27. Regression analysis of the accordion pattern.

R	R Squared	Adjusted R Squared	Std. Error of the Estimate
.382	.146	.132	.08846

With the equation being:

Cognitive load = .399 + .036EFComponent

And finally, for the wizard design pattern, the regression results were (Table 28):

Table 28. Regression analysis of the accordion pattern.

R	R Squared	Adjusted R Squared	Std. Error of the Estimate
.559	.312	.301	.80048

The regression equation being:

Cognitive load = 4.120 - .535EFComponent

The regression analysis showed in that in some cases, the executive function moderately influenced the cognitive load produced by the patterns (between .4 and .5 r-squared), in some cases, it was almost negligible. This could be explained by the difficulty presented by each pattern, some could be so easy that they used very little mental resources to complete the task, while some of the harder ones to use showed more influence of the executive function component.

Discussion

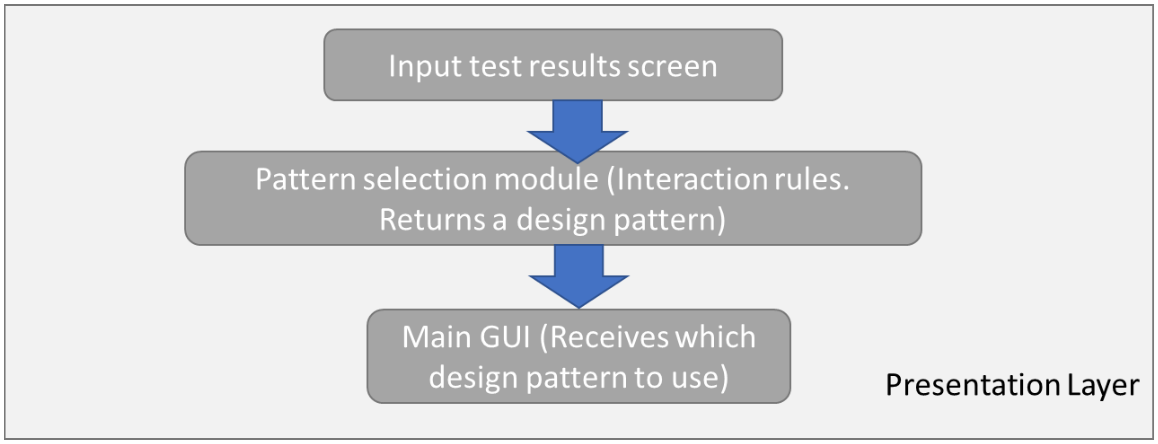
Some of the tests measurements were significant in determining how the rules behaved. Although not all of the classification trees were significant. It was mostly in the cases where near the middle of the pack regarding usability testing results of the patterns that the trees weren't a good fit, but as seen in the results, the trees of the top patterns and the worst performing ones were the more accurate.

Auditory attention (AA) turned out to be the main predictors used on the rules of the classification trees, followed by set shifting, digit span, number and letters. Auditory attention and response set (SS) represent test results from the NEPSY II test of cognitive flexibility, while the digit span (DS) and number and letters (NL) test scores come from the WISC-IV tests for working memory.

The obtained rules should be useful as a guideline if you have a target audience with known cognitive impairments. Even more so if you have access to neuropsychological test results of WISC-IV and NEPSY II. If that is the case, the rules can be integrated to develop an application with an adaptable interface based on the values of the tests.

A user would need to insert the test values in a separate interface or screen, for the next screens to be adapted with the appropriate interface pattern (Figure 11).

Figure 12. Basic layered architecture of the presentation layer of an app with the integrated interaction rules.



Since we tested patterns that are used mostly as a form of navigation, the rules could be useful for application that deals with some task decomposition.

In the repeated measure ANOVA it showed a statistically significant difference between means of the cognitive load produced by each pattern, even with the covariate of executive functions, although not as significant. Even then, it shows that there is some relationship between the executive functions of working memory, cognitive flexibility, cognitive load, and the interface design pattern used.

The regression analysis is another story, some patterns showed very little influence from the executive function component, while a few showed a moderate influence nearing a 50% accuracy rate. It could probably be factored to other user attributes that for this study could not be controlled, for example, general IQ and a smaller difference in age.

Chapter VI:

Conclusion

"I think and think for months and years. Ninety-nine times, the conclusion is false. The hundredth time I am right."

- Albert Einstein

Conclusion

In this dissertation we presented a series of rules obtained from extensive usability testing of several interface design patterns with participants that presented impairments in the executive functions of working memory and cognitive flexibility. We presented how we obtained the rules and tested the accuracy of the rules with a separate dataset composed of random participants of the same age range that did not participate in the original set of tests.

Executive functions are an integral part of our cognitive capabilities by managing several functions for any goal-directed behavior. From how we store and manipulate information in our memory, to organizing our thoughts and actions, even managing our feelings.

It comes to no surprise that they also influence how we use software applications. The executive function that is most considered in interface design is working memory, since it is related to how much information we can handle at any given time. But there are other executive functions that could also have an impact since they use working memory as a component, such as planning and cognitive flexibility, which could have an impact on how a user manages with the workflow of the software.

One of the main problems of interface design is tied to working memory: the problem of cognitive load produced by the interface design itself. We all have finite resources in our minds, if the difficulty of the task, plus the complexity of the interface design could overwhelm the user causing usability problems.

But cognitive load could affect other executive functions, since those same mental resources also serve for organizing thoughts, planning, shifting focus, and other functions. With this in mind we set out to develop rules to help us better select design patterns for different levels of working memory and cognitive flexibility and see the impact of each on cognitive load produced.

The rules could be used as a design guideline when designing an interface, specially for users with some form of cognitive impairments, such as Autism Spectrum Disorder, since that in most impairments, there is some form of executive function impairment. This way we can choose design patterns better suited for particular group of users or a particular user, since the rules are more specific than the general design guidelines.

They also could be integrated into an Adaptive User Interface if results of the same neuropsychological tests are available of the users, but the rules should be complemented with values of parameters of the different interface components at a given interface pattern uses.

Based on the results of the experiment, we can conclude that there is a non-trivial relationship between the usability results of different design patterns and the measurements of working memory and cognitive flexibility, and the perceived cognitive load produced by the pattern upon the users.

We saw a statistically significant difference in the cognitive load measurements of the design patterns when applied a repeated measure ANOVA. Also, when using a component of the executive functions as a covariate it was still found a significant impact, although less so. Even then, it shows that there is some relationship between the executive functions of working memory and cognitive flexibility with cognitive load, and the interface design pattern used.

There were some minor conflicting results. The regression analysis showed very little influence from the executive function component in most cases, while a few showed a moderate influence nearing a .50 R-Squared. It could probably be factored to other user attributes that for this study could not be controlled, for example, general IQ and a smaller difference in age.

There is much future work to be done, this study can serve as a base in order to continue with a more throughout experiment. Due to limitations of the available population, there were aspects that we could not properly control, such as IQ of the participants, and a minimal standard deviation of the age, since there is a very limited number of, in this case, children with autism in our community where we made the study. All of those factors could have greatly impacted the results, so the results obtained of this study can’t be said to be a good standard, but can serve as a base for future work.

References

Anderson, J. R. (1996). ACT: A simple theory of complex cognition. *American Psychologist*, *51*(4), 355.

Anderson, V., Jacobs, R., & Anderson, P. J. (2010). *Executive functions and the frontal lobes: A lifespan perspective*. Psychology Press.

Antonenko, P., Paas, F., Grabner, R., & Van Gog, T. (2010). Using electroencephalography to measure cognitive load. *Educational Psychology Review*, *22*(4), 425–438.

Association, A. P. (2013). *Diagnostic and statistical manual of mental disorders (DSM-5®)*. American Psychiatric Pub.

Ayres, P. L. (2001). Systematic mathematical errors and cognitive load. *Contemporary Educational Psychology*, *26*(2), 227–248.

Ayres, P., & Sweller, J. (2014). The split-attention principle in multimedia learning. *The Cambridge Handbook of Multimedia Learning*, *2*.

Baddeley, A. (2000). The episodic buffer: a new component of working memory? *Trends in Cognitive Sciences*, *4*(11), 417–423.

Baddeley, A. D., Baddeley, A. D., & Braddlely, A. (1986). *Working memory* (Vol. 11). Elsevier.

Baddeley, A. D., & Hitch, G. (1974). Working memory. *Psychology of Learning and Motivation*, *8*, 47–89.

Barkley, R. A. (1997). Behavioral Inhibition, Sustained Attention, and Executive Functions" Constructing a Unifying Theory of ADHD, *121*(1), 65–94. Retrieved from http://sciences.ucf.edu/psychology/childrenslearningclinic/wp-content/uploads/sites/91/2013/08/Barkley-1997-Psych-Bulletin.pdf

Bauer, F., Bolliet, L., & Helms, H. J. (1968). Report of a conference sponsored by the NATO

Science Committee. In *NATO Software Engineering Conference* (Vol. 1968, p. 8).

Biswas, P., Bhattacharya, S., & Samanta, D. (2005). User model to design adaptable interfaces for motor-impaired users. In *TENCON 2005 2005 IEEE Region 10* (pp. 1–6). IEEE.

Biswas, P., & Robinson, P. (2008). Automatic evaluation of assistive interfaces. In *Proceedings of the 13th international conference on Intelligent user interfaces* (pp. 247–256). ACM.

Biswas, P., & Robinson, P. (2009a). Modelling perception using image processing algorithms. In *Proceedings of the 23rd British HCI Group Annual Conference on People and Computers: Celebrating People and Technology* (pp. 494–503). British Computer Society.

Biswas, P., & Robinson, P. (2009b). Predicting pointing time from hand strength. In *Symposium of the Austrian HCI and Usability Engineering Group* (pp. 428–447). Springer.

Biswas, P., Robinson, P., & Langdon, P. (2012). Designing inclusive interfaces through user modeling and simulation. *International Journal of Human-Computer Interaction*, *28*(1), 1–33.

Bratfisch, O. (1972). Perceived Item-Difficulty in Three Tests of Intellectual Performance Capacity.

Britton, B. K., & Tesser, A. (1982). Effects of prior knowledge on use of cognitive capacity in three complex cognitive tasks. *Journal of Verbal Learning and Verbal Behavior*, *21*(4), 421–436.

Byrne, M. D. (2001). ACT-R/PM and menu selection: Applying a cognitive architecture to HCI. *International Journal of Human-Computer Studies*, *55*(1), 41–84.

Caldwell, B., Cooper, M., Reid, L. G., & Vanderheiden, G. (2008). Web content accessibility guidelines (WCAG) 2.0. *WWW Consortium (W3C)*.

Calvary, G., Coutaz, J., Thevenin, D., Limbourg, Q., Souchon, N., Bouillon, L., …

Vanderdonckt, J. (2002). Plasticity of user interfaces: A revisited reference framework. In *In Task Models and Diagrams for User Interface Design*. Citeseer.

Card, S. K., Moran, T. P., & Newell, A. (1983). *The psychology of human-computer interaction*. L. Erlbaum Associates. Retrieved from https://dl.acm.org/citation.cfm?id=578027

Chan, A. S., Cheung, M., Han, Y. M. Y., Sze, S. L., Leung, W. W., Man, H. S., & To, C. Y. (2009). Executive function deficits and neural discordance in children with Autism Spectrum Disorders. *Clinical Neurophysiology : Official Journal of the International Federation of Clinical Neurophysiology*, *120*(6), 1107–15. https://doi.org/10.1016/j.clinph.2009.04.002

Chandler, P., & Sweller, J. (1992). The split-attention effect as a factor in the design of instruction. *British Journal of Educational Psychology*, *62*(2), 233–246.

Delis, D. C. (2012). Delis-Rating of Executive Function (D-REF). *Bloomington, MN: Pearson*.

Dieterich, H., Malinowski, U., Kühme, T., & Schneider-Hufschmidt, M. (1993). State of the art in adaptive user interfaces. *Human Factors in Information Technology*, *10*, 13.

Dix, A. (2009). Human-computer interaction. In *Encyclopedia of database systems* (pp. 1327–1331). Springer.

Figueroa, A. M., Juárez-Ramírez, R., Inzunza, S., & Valenzuela, R. (2014). Implementing adaptive interfaces: a user model for the development of usability in interactive systems. *International Journal of Computer Systems Science and Engineering*, *29*(1), 95–104.

Freyhoff, G., Hess, G., Kerr, L., Tronbacke, B., & Van Der Veken, K. (1998). Make it Simple, European Guidelines for the Production of Easy-toRead Information for People with Learning Disability. *ILSMH European Association, Brussels*.

Gajos, K., & Weld, D. S. (2004). SUPPLE: automatically generating user interfaces. In

Proceedings of the 9th international conference on Intelligent user interfaces (pp. 93–100). ACM.

Godefroy, O., Cabaret, M., Petit-Chenal, V., Pruvo, J.-P., & Rousseaux, M. (1999). Control functions of the frontal lobes. Modularity of the central-supervisory system? *Cortex*, *35*(1), 1–20.

Hill, E. L. (2004a). Evaluating the theory of executive dysfunction in autism. *Developmental Review*, *24*(2), 189–233. https://doi.org/10.1016/j.dr.2004.01.001

Hill, E. L. (2004b). Executive dysfunction in autism. *Trends in Cognitive Sciences*, *8*(1), 26–32.

Hill, E. L. (2004c). Executive dysfunction in autism☆. *Trends in Cognitive Sciences*, *8*(1), 26–32. https://doi.org/10.1016/j.tics.2003.11.003

Hongwanishkul, D., Happaney, K. R., Lee, W. S. C., & Zelazo, P. D. (2005). Assessment of hot and cool executive function in young children: Age-related changes and individual differences. *Developmental Neuropsychology*, *28*(2), 617–644.

Jacko, J. A. (2012). *Human computer interaction handbook: Fundamentals, evolving technologies, and emerging applications*. CRC press.

Jorge, J. A. (2001). Adaptive tools for the elderly: new devices to cope with age-induced cognitive disabilities. In *Proceedings of the 2001 EC/NSF workshop on Universal accessibility of ubiquitous computing: providing for the elderly* (pp. 66–70). ACM.

Kaklanis, N., Biswas, P., Mohamad, Y., Gonzalez, M. F., Peissner, M., Langdon, P., … Jung, C. (2014). Towards standardisation of user models for simulation and adaptation purposes. *Universal Access in the Information Society*. https://doi.org/10.1007/s10209-014-0371-2

Kieras, D. (1994). GOMS modeling of user interfaces using NGOMSL. In *Conference companion on Human factors in computing systems* (pp. 371–372). ACM.

Kieras, D. E., & Meyer, D. E. (1997). An overview of the EPIC architecture for cognition and performance with application to human-computer interaction. *Human-Computer Interaction*, *12*(4), 391–438.

Kim, G. J. (2015). *Human-Computer Interaction: fundamentals and practice*. CRC press.

Korkman, M., Kirk, U., & Kemp, S. (2007). NEPSY-II: Clinical and interpretive manual. *San Antonio, TX: The Psychological Corporation*.

Laurie-Rose, C., Frey, M., Ennis, A., & Zamary, A. (2014). Measuring perceived mental workload in children. *The American Journal of Psychology*, *127*(1), 107–125.

Lavie, T., & Meyer, J. (2010). Benefits and costs of adaptive user interfaces. *International Journal of Human-Computer Studies*, *68*(8), 508–524.

Lezak, M. D. (2004). *Neuropsychological assessment*. Oxford University Press, USA.

Luria, A. R. (1973). An introduction to neuro-psychology. In *The working brain*. Penguin Books Harmondsworth.

Mayer, R. E., & Moreno, R. (2003). Nine ways to reduce cognitive load in multimedia learning. *Educational Psychologist*, *38*(1), 43–52.

Mejía, A., Juárez-Ramírez, R., Inzunza, S., & Valenzuela, R. (2012). Implementing adaptive interfaces: a user model for the development of usability in interactive systems. In *Proceedings of the CUBE International Information Technology Conference* (pp. 598–604). ACM.

Miller, E. K., & Cohen, J. D. (2001). An integrative theory of prefrontal cortex function. *Annual Review of Neuroscience*, *24*(1), 167–202.

Miyake, A., Friedman, N. P., Emerson, M. J., Witzki, A. H., Howerter, A., & Wager, T. D. (2000). The unity and diversity of executive functions and their contributions to complex “frontal lobe” tasks: A latent variable analysis. *Cognitive Psychology*, *41*(1), 49–100.

Nielsen, J. (1994). Enhancing the explanatory power of usability heuristics. In *Proceedings of the SIGCHI conference on Human Factors in Computing Systems* (pp. 152–158). ACM.

Nielsen, J. (1995). 10 usability heuristics for user interface design. *Fremont: Nielsen Norman Group.[Consult. 20 Maio 2014]. Disponível Na Internet*.

Paas, F., Ayres, P., & Pachman, M. (2008). Assessment of cognitive load in multimedia learning. *Recent Innovations in Educational Technology That Facilitate Student Learning, Information Age Publishing Inc., Charlotte, NC*, 11–35.

Paas, F. G. (1992). Training strategies for attaining transfer of problem-solving skill in statistics: A cognitive-load approach. *Journal of Educational Psychology*, *84*(4), 429.

Palinko, O., Kun, A. L., Shyrokov, A., & Heeman, P. (2010). Estimating cognitive load using remote eye tracking in a driving simulator. In *Proceedings of the 2010 symposium on eye-tracking research & applications* (pp. 141–144). ACM.

Pribram, K. H. (1973). The primate frontal cortex—executive of the brain. *Psychophysiology of the Frontal Lobes*, 293–314.

Rice, A. D., & Lartigue, J. W. (2014). Touch-level model (TLM): evolving KLM-GOMS for touchscreen and mobile devices. In *Proceedings of the 2014 ACM Southeast regional conference* (p. 53). ACM.

Shneiderman, B. (2010). *Designing the user interface: strategies for effective human-computer interaction*. Pearson Education India.

Stephanidis, C., Paramythis, A., Sfyrakis, M., Stergiou, A., Maou, N., Leventis, A., … Karagiannidis, C. (1998). Adaptable and adaptive user interfaces for disabled users in the AVANTI project. *Intelligence in Services and Networks: Technology for Ubiquitous*

Telecom Services, 153–166.

Sweller, J. (1988). Cognitive load during problem solving: Effects on learning. *Cognitive Science*, *12*(2), 257–285.

Sweller, J., Ayres, P., & Kalyuga, S. (2011). *Cognitive load theory* (Vol. 1). Springer.

Tarpin-Bernard, F., & Habieb-Mammar, H. (2005). Modeling elementary cognitive abilities for adaptive hypermedia presentation. *User Modeling and User-Adapted Interaction*, *15*(5), 459–495.

Thevenin, D., & Coutaz, J. (1999). Plasticity of User Interfaces: Framework and Research Agenda. In *Interact* (Vol. 99, pp. 110–117).

Van Gerven, P. W. M., Paas, F., Van Merriënboer, J. J. G., & Schmidt, H. G. (2004). Memory load and the cognitive pupillary response in aging. *Psychophysiology*, *41*(2), 167–174.

Van Gog, T., & Scheiter, K. (2010). Eye tracking as a tool to study and enhance multimedia learning. Elsevier.

Varma, S., & Just, M. A. (2006). 4CAPS: An Adaptive Architecture for Human Information Processing. In *AAAI Spring Symposium: Between a Rock and a Hard Place: Cognitive Science Principles Meet AI-Hard Problems* (pp. 91–96).

Wagner, E. (1992). A system ergonomics design methodology for HCI development. In *Proc. of East-West International Conference on Human-Computer Interaction EWHCI'92* (pp. 388–407).

Wechlser, D. (2003). Wechsler Intelligence Scale for Children (WISC-IV). *San Antonio, TX: Psychological Corporation*.

Zhang, P., Carey, J., Te'eni, D., & Tremaine, M. (2005). Integrating human-computer interaction development into the systems development life cycle: A methodology. *Communications of* *the Association for Information Systems*, *15*(1), 29.

Zudilova-Seinstra, E. (2007). On the role of individual human abilities in the design of adaptive user interfaces for scientific problem solving environments. *Knowledge and Information Systems*, *13*(2), 243–270.