Analyzing Exploration and Exploitation patterns in Multimodal Dialogue Games for preschoolers

Giorgos Evgeneiadis

Department of Electronic and Computer Engineering

Technical University of Crete

Thesis committee:
Alexandros Potamianos, Supervisor
Katerina Mania
Vassilis Digalakis

A thesis submitted in partial fulfillment for the Diploma degree of

Electronic and Computer Engineering

Chania,October 2010

Acknowledgements

First of all i would like to thank my supervisor Professor Alexandros Potamianos for his guidance and valuable directions towards accomplishing this work. Also i wish to thank Th. Kannetis for his preceding work and his guidance in the first steps of development and Vasiliki Kouloumenta for her help in the evaluation procedure.

A very special thank you to Markolefa Marilia for her Voice Acting and Recording performance. Next i would like to thank Panagiotis Giannakouras for the valuable discussions and assistance in evaluation and Markos Mellas for his new dialog concepts and designing ideas. Also, i would like to acknowledge my friends and colleagues Giorgos Vastarouhas, Vassilis Kavadias, Giannis Kyriakopoulos and Nikos Petridis for their company and psychological support during the long hours of programming time.

Next, a very big thank you to the kindergartens "Zouzounospito" and "Paidokipos" and their teachers Maria and Ioanna for their support in our evaluation and all of the teachers and kids in the kindergarten of Agrokipio for their valuable advices and memorable moments together. Finally, i could not forget to thank the kids Konstantina, Marilena, Marios, Theano, Giorgos, Despoina, Taksiarhis, Ilektra and Eftixia for participating in our evaluation and for being indescribably nice and cute.

Abstract

Multimodal dialog systems for children are constantly gaining popularity, mainly because of their naturalness, robustness and efficiency. Advancement in HCI has given the opportunity for computers to act as powerful learning and playing tools. Designing such applications for children though comes with increased responsibility and can be rather challenging considering the mannerisms of their interaction patterns which are very different than those of the adults.

The terms of Exploration and Exploitation have increasingly come to dominate many fields of scientific research. In this work we will investigate how children instinctively respond to the Exploration versus Exploitation (EvE) dilemma by interacting with the application we have developed. For this purpose we have integrated and extended the on-line multimodal platform designed by Kannetis [1] in his preceding work.

In Chapter 2 we set forth some background information on the preschoolers' cognitive development, the fields concerned with the EvE dilemma and the HCI for kids. Chapter 3 refers to Kannetis work on the Platform's Design and Implementation which we have also used. Our application supports speech and mouse usage and consists of two playing modes: those of Exploration and Exploitation.

The Exploitation mode consists of five games based on popular preschool activities which were designed by Kannetis [1] and Meliopoulos [2]. In this mode children have the opportunity to use/exploit their existing knowledge on some fundamental things (numbers/animal/shape recognition, addition task and quantity comparison). The Exploration

mode designed in this work, was implemented with the use of Embodied Conversational Agents. In this mode the children have the opportunity to engage in conversation with various characters (the main ECA, the companion ECA and secondary ECAs), participate in interactive storytelling and initiate small talk with them. Exploration mode was further separated into two different structures for the first and any subsequent playthroughs. There is no demand of knowledge and use of skills here. There is also no immediate goal or reward and the motivation lies in just exploring the many and random conversational possibilities. Our application's structure and implementation is analyzed in Chapter 4.

Nine preschool children aged 4-6 years old (three of each age) participated in our Evaluation procedure where each one had two sessions with our application. Our Evaluation Methology and Experimental Setup are analyzed in Chapter 5.

Finally in Chapter 6 we analyze the Objective (Logfiles analysis) and Subjective (Questionnaires) Metrics and Statistics and identify the preferences and patterns each age follows in their interchange among the two modes during the Evaluation sessions.

Abstract (in Greek)

Οι πολυτροπικές εφαρμογές με υποστήριξη φωνής που προορίζονται για παιδιά γίνονται ολοένα και πιο δημοφιλείς, κυρίως λογω της φυσικότιτας και αποτελεσματικότητάς τους. Η πρόοδος στις μεθόδους αλληλεπίδρασης του ανθρώπου με τους υπολογιστες (HCI) έξει δώσει στους τελευταίους την ευκαιρία να ανελιχθουν σε ισχυρά μέσα μάθησης αλλά και διασκέδασης. Ο σχεδιασμός τέτοιων εφαρμογών προορισμένων για παιδιά οστόσο προυποθέτει μεγαλύτερη υπευθυνότητα εκ μέρους του προγραμματιστή και μπορεί να αναδειχθεί ιδιαίτερα δύσκολος αν αναλογιστούμε τις ιδιομορφίες αυτης τις ηλικίας όσον αφορά τις μεθόδους αλληλεπίδρασης τους οι οποίες διαφέρουν κατά πολύ από αυτές των ενηλίκων.

Οι όροι Exploration (Εξερεύνηση νέων πιθανοτήτων) και Exploitation (εκμετάλλευση/χρήση παλιών βεβαιοτήτων) αποσχολούν ολοένα και περισσότερο μια ευρεία γκάμα επιστημονικών κλάδων. Στη δουλειά που παρουσιάζεται εδώ θα εξετάσουμε τον τρόπο με τον οποίο τα παιδιά ανταποκρίνονται στο δίλλημα κατανομής του χρόνου ανάμεσα σε Exploration και Exploitation (Exploration versus Exploitation—EvE dilemma) παίζοντας την εφαρμογή που αναπτύξαμε για το σκοπό αυτό. Για την ανάπτυξη της εφαρμογής μας, ενσωματώσαμε και επεκτείναμε την διαδυκτιακή πολυτροπική πλατφόρμα (online multimodal platform) την οποία σχεδίασε ο Θεοφάνης Καννέτης στην προηγηθείσα δουλειά του στα πλαίσια εκπλήρωσης της μεταπτυχιακής του εργασίας [1].

Στο Κεφάλαιο 2 παρουσιάζουμε κάποιες γενικές πληροφορίες ως υπόβαθρο οι οποίες αφορούν την εξέλιξη της γνωστικής/μαθησιακής διαδικασίας των παιδιών προσχολικής ηλικίας, τους επιστημονικούς κλάδους που εμπλέκονται με το EvE δίλημμα καθώς και σχεδιαστικούς κανόνες και αναφορές στην ανάπτυξη εφαρμογών για παιδια (HCI for kids). Στο κεφάλαιο 3 παρουσιάζεται η προηγηθείσα δουλειά του Καννέτη που

αφορά το σχεδιασμό και τη δομή της πλατφόρμας που χρησιμοποιήσαμε. Η εφαρμογή μας υποστηρίζει τη χρήση φωνής και ποντικιού και απαρτίζεται από δύο διακριτά μέρη (Exploration/Exploitation modes).

Το κομμάτι του Exploitation αποτελείται από πέντε παιχνίδια βασισμένα σε δημοφιλείς προσχολικές δραστηριότητες τα οποία σχεδιάστηκαν από τους Θ. Καννέτη [1] και Σ. Μελιόπουλο [2]. Σε αυτό το μέρος της εφαρμογής μας τα παιδιά έχουν την ευκαιρία να χρησιμοποιήσουν κάποιες από τις πολύ βασικές τους γνώσεις (αναγνώριση σχημάτων/ζώων/αριθμών, πρόσθεση σε μερικές περιπτώσεις και σύγκριση ποσοτήτων). Το κομμάτι του Exploration το οποίο σχεδιάστηκε στα πλαίσια αυτής της διπλωματικής, υλοποιήθηκε με τη χρήση οντοτήτων με σωματική υπόσταση (Embodied Conversational Agents). Σε αυτό το κομμάτι τα παιδιά έχουν την ευκαιρία να γνωρίσουν και να συνομιλήσουν με διάφορους χαρακτήρες, να συμμετάσχουν σε αλληλεπιδραστική αφήγηση και να εμπλακούν σε γενικές συζητήσεις με κάποια τυχαία θέματα. Εδώ δεν απαιτείται κάποιο είδος γνώσης ή επιδεξιοτήτων ούτε υπάρχει κάποιος άμεσος στόχος ή ανταμοιβή και το μόνο κίνητρο έγκειται στην εξερεύνηση των ποικίλων και τυχαίων πιθανοτήτων διαλόγου. Η δομή και η υλοποίηση της εφαρμογής μας αναλύεται στο Κεφάλαιο 4.

Εννιά παιδιά προσχολικής ηλικίας (4 έως 6 ετών —τρία από κάθε ηλικία) συμμετείχαν στην πειραματική διαδικασία αξιολόγησης, όπου κάθε παιδί άσχολήθηκε δύο φορές με την εφαρμογή μας κατά τις διαφορετικές επισκέψεις μας σε δύο νηπιαγωγεία. Η πειραματική μας διαδικασία αναλύεται στο Κεφάλαιο 5.

Τέλος, στο Κεφάλαιο 6 εξετάζουμε και αναλύουμε τα αντικειμενικά (αρχεία καταγραφής) και υποκειμενικά (ερωτηματολόγια) μας στατιστικά αποτελέσματα και γίνεται μια προσπάθεια αναγνώρισης των προτιμήσεων αλλά και των μοτίβων που ακολουθεί η κάθε ηλικία κατά την εναλλαγή ανάμεσα στα δύο κομμάτια που απαρτίζουν την εφαρμογή μας.

Contents

List of Figures

List of Tables

Chapter 1

Introduction

1.1 Introduction

Many large and small decisions we make in our daily lives —which ice cream to choose, what research projects to pursue, which partner to marry or not— require an exploration of alternatives before committing to and exploiting the benefits of a particular choice. Furthermore, many decisions require re-evaluation, and further exploration of alternatives, in the face of changing needs or circumstances. That is, often our decisions depend on a higher level choice: whether to exploit well known but possibly suboptimal alternatives or to explore risky but potentially more profitable ones. How we choose between exploitation and exploration remains an important and open question that has received relatively limited attention in the behavioural and brain sciences. The choice could depend on a number of factors, including the familiarity of the environment, how quickly the environment is likely to change and the relative value of exploiting known sources of reward versus the cost of reducing uncertainty through exploration. There is no known generally optimal solution to the exploration versus exploitation problem, and a solution to the general case may indeed not be possible. However, there have been formal analyses of the optimal policy under constrained circumstances. There have also been specific suggestions of how humans and animals may respond to this problem under particular experimental conditions as well as proposals about the brain mechanisms involved. [3]

Every researcher has personal experience with the exploration-exploitation dilemma. At some point in the conduct of a study, when the data are still inconclusive, it may become necessary to decide how to proceed. On the one hand, there is the option to continue with the experiment, in the hope that with more effort and data, the results will look more promising. Alternatively, the experiment can be scrapped in favour of a modified experimental design, a new approach to the problem, or an entirely new research topic. That is, the experimenter faces a trade-off between the value of exploitation versus exploration. This example highlights the importance of this problem in decision making, one that has typically been ignored in psychological research on cognitive control and executive function.

Exploration: —described by terms such as search, variation, risk taking, experimentation, discovery, innovation [4].

Exploitation: —described by terms such as choice, production, efficiency, selection, implementation, execution [4].

However generalized this concept might seem, it is a natural procedure in the human mind. We will focus our attention in the pre-operational stages of the human's cognitive development (ages 4-6 to be more specific) and how children choose to switch between these two "modes". In order to identify the patterns that kids use for their unbiased traverse through Exploration and Exploitation, we have designed a multimodal application/game where the child has the option to navigate through these modes. *Exploration* was implemented and materialized in the form of conversation and storytelling between the child and an Embodied Conversational Agent using Automatic Speech Recognition. The mode of *Exploitation* is a series of mini-games, targeted for toddlers and preschoolers, designed by *Theofanis Kannetis [1]* (except the 'Farm' game which was designed by *Spyros Meliopoulos)* —as part of their Master Thesis and Thesis accordingly, which preceded this work.

In the recent years lot of research has be done in the design and implementation of multimodal dialog systems for children. A number of prototype systems with advanced spoken dialog interfaces, multimodal interaction capabilities and/or embodied conversational characters have been implemented. However, almost all of these systems have focused in the age group 6-15. And that is because older children can be more easily controlled in experimental conditions than younger children [1]. To investigate the choices preschool children make in this Exploration/Exploitation environment (and how they interact with multimodal dialog systems), we have integrated an on-line (web-based) multimodal platform, in order to be able to quickly prototype, deploy and evaluate multimodal dialog systems for preschoolers. Using this platform made by *Theofanis Kannetis* for his work on Multimodal Dialog Systems For Preschoolers (along with the mini-games mentioned above) we have designed and added an environment where children can tend to their Explorative needs by engaging in interactive storytelling and conversing with the implemented ECAs.

To gather the necessary Metrics and Statistics, our evaluation process took place in two different kindergartens where nine native Greek children (four to six years old, three for each age) had the opportunity to interact with our system.

Chapter 2

Background Information

2.1 Introduction

Before we proceed to this works' analysis, let's take a quick glimpse at some necessary background information in order to gain a more spherical view of children's cognition status in pre-operational stage, researches about Human-Computer Interaction specialized for kids and finally the Exploration vs Exploitation dilemma and which fields it concerns.

2.2 Preschooler's Cognitive Development

Cognitive development is the growth of cognitive processes such as reasoning and problem solving throughout the life span. Due to the fact that psychology is not our field of research, we cannot over-analyze the following theories. However, it is somewhat important to have a superficial view at least.

2.2.1 Piaget's Theory of Cognitive Development

Jean Piaget (1896-1980) was a biologist who originally studied mollusks (publishing twenty scientific papers on them by the time he was 21) but moved into the study of the development of children's understanding, through observing, talking and listening to them while they worked on exercises he set.

His view of how children's minds work and develop has been enormously influential, particularly in educational theory. His particular insight was the role of maturation (simply growing up) in children's increasing capacity to understand their world: they cannot undertake certain tasks until they are psychologically mature enough to do so. His research has spawned a great deal more, much of which has undermined the detail of his own, but like many other original investigators, his importance comes from his overall vision.

He proposed that children's thinking does not develop entirely smoothly: instead, there are certain points at which it "takes off" and moves into completely new areas and capabilities. He saw these transitions as taking place at about 18 months, 7 years and 11 or 12 years. This has been taken to mean that before these ages children are not capable (no matter how bright) of understanding things in certain ways, and has been used as the basis for scheduling the school curriculum. Whether or not should be the case is a different matter.

Below we quote his ideas in some minimal context. It does not set out to provide a comprehensive account of Piaget's ideas. It is over-simplified [5].

2.2.1.1 Piaget's Stages of Cognitive Development

Sensorimotor stage (Infancy): In this period (which has 6 stages), intelligence is demonstrated through motor activity without the use of symbols. Knowledge of the world is limited (but developing) because its based on physical interactions / experiences. Children acquire object permanence at about 7 months of age (memory). Physical development (mobility) allows the child to begin developing new intellectual abilities. Some symbolic (language) abilities are developed at the end of this stage. Differentiates self from objects. Recognizes self as agent of action and begins to act intentionally: e.g. pulls a string to set mobile in motion or shakes a rattle to make a noise. Achieves object permanence: realizes that things continue to exist even when no longer present to the sense.

Pre-operational stage (Toddler and Early Childhood:) In this period (which has two sub stages), intelligence is demonstrated through the use of symbols, language use matures, and memory and imagination are developed, but thinking is done in a non-logical, non-reversable manner. Egocentric thinking predominates: has difficulty taking the viewpoint of others. Learns to use language and to represent objects by images and words. Classifies objects by a single feature: e.g. groups together all the red blocks regardless of shape or all the square blocks regardless of colour. This is our target stage. Two substages can be formed from preoperative thought.

The Symbolic Function Substage which occurs between about the ages of 2 and 4.

The Intuitive Thought Substage which occurs between about the ages of 4 and 7. Our target group is 4-6.

As stated in [6], children tend to become very curious and ask many questions and begin to use primitive reasoning. There is an emergence in the interest of reasoning and wanting to know why things are the way they are. Piaget called it the intuitive substage because children realize they have a vast amount of knowledge but they are unaware of how they know it. Centration and conservation are both involved in preoperative thought. Centration is the act of focusing all attention on one characteristic compared to the others. Centration is noticed in conservation; the awareness that altering a substance's appearance does not change its basic properties. Children at this stage are unaware of conservation. In Piaget's most famous task, a child is represented with two identical beakers containing the same amount of liquid. The child usually notes that the beakers have the same amount of liquid. When one of the beakers is poured into a taller and thinner container, children who are typically younger than 7 or 8 years old say that the two beakers now contain a different amount of liquid. The child simply focuses on the height and width of the container compared to the general concept. Piaget believes that if a child fails the conservation-of-liquid task, it is a sign that they are at the pre-operational stage of cognitive development. The child also fails to show conservation of number, matter, length, volume, and area as well. Another example is when a child is shown 7 dogs and 3 cats and asked if there are more dogs than cats. The child would respond positively. However when asked if there are more dogs than animals, the child would once again respond positively. Such fundamental errors in logic show the transition between intuitiveness in solving problems and true logical reasoning acquired in later years when the child grows up.

Concrete operational stage (Elementary and early adolescence): In this stage (characterized by 7 types of conservation: number, length, liquid, mass, weight, area, volume), intelligence is demonstrated through logical and systematic manipulation of symbols related to concrete objects. Can think logically about objects and events. Operational thinking develops (mental actions that are reversible). Egocentric thought diminishes. Achieves conservation of number (age 6), mass (age 7), and weight (age 9). Classifies objects according to several features and can order them in series along a single dimension such as size.

Formal operational age (Adolescence and adulthood): In this stage, intelligence is demonstrated through the logical use of symbols related to abstract concepts. Early in the period there is a return to egocentric thought. Only 35% of high school graduates in industrialized countries obtain formal operations; many people do not think formally during adulthood. Can think logically about abstract propositions and test hypotheses systematically. Becomes concerned with the hypothetical, the future, and ideological problems .

2.2.1.2 Piaget's Key Ideas

Adaptation: What it says: adapting to the world through assimilation and accommodation

Assimilation: The process by which a person takes material into their mind from the environment, which may mean changing the evidence of their senses to make it fit.

Accommodation: The difference made to one's mind or concepts by the process of assimilation. Note that assimilation and accommodation go together: you can't have one without the other.

Classification: The ability to group objects together on the basis of common features.

Class Inclusion: The understanding, more advanced than simple classification, that some classes or sets of objects are also sub-sets of a larger class. (E.g. there is a class of objects called dogs. There is also a class called animals. But all dogs are also animals, so the class of animals includes that of dogs)

Conservation: The realization that objects or sets of objects stay the same even when they are changed about or made to look different.

Decentration: The ability to move away from one system of classification to another one as appropriate.

Egocentrism: The belief that you are the centre of the universe and everything revolves around you: the corresponding inability to see the world as someone else does and adapt to it. Not moral "selfishness", just an early stage of psychological development.

Operation: The process of working something out in your head. Young children (in the sensorimotor and pre-operational stages) have to act, and try things out in the real world, to work things out (like count on fingers): older children and adults can do more in their heads.

Schema (or scheme): The representation in the mind of a set of perceptions, ideas, and/or actions, which go together.

Stage: A period in a child's development in which he or she is capable of understanding some things but not others.

2.2.1.3 Assimilation and Accommodation

Assimilation and **Accommodation** are the two complementary processes of Adaptation described by Piaget, through which awareness of the outside world is internalized. Although one may predominate at any one moment, they are inseparable and exist in a dialectical relationship. In **Assimilation,** what is perceived in the outside world is incorporated into the internal world (not Piagetian terminology), without changing the structure of that internal world, but potentially at the cost of "squeezing" the external perceptions to fit. We can think like a database example: our mind has its database already built, with its fields and categories already defined. If it comes across new information which fits into those fields, it can assimilate it without any trouble. In **Accommodation,** the internal world has to accommodate itself to the evidence with which it is confronted and thus adapt to it, which can be a more difficult and painful process. In the database analogy, it is like what happens when you try to put in information which does not fit the pre-existent fields and categories. You have to develop new ones to accommodate the new information. In reality, both are going on at the same time, so that—just as the mower blade cuts the grass, the grass gradually blunts the blade—although most of the time we are assimilating familiar material in the world around us, nevertheless, our minds are also having to adjust to accommodate it [7].

2.2.2 Neo-Piagetian Theories of Cognitive Development

Jean Piaget's theory of cognitive development was criticized on many grounds. One criticism is concerned with the very nature of development itself. It is suggested that Piaget's theory does not explain why development from stage to stage occurs. The theory is also criticized for ignoring individual differences in cognitive development. That is, the theory does not account for the fact that some individuals move from stage to stage faster than other individuals. Finally, another criticism is concerned with the nature of stages themselves. Research shows that the functioning of a person at a given age may be so variable from domain to domain, such as the understanding of social, mathematical, and spatial concepts, that it is not possible to place the person in a single stage. To remove these weaknesses, a group of researchers, who are known as neo-Piagetian theorists, advanced models that integrate concepts from Piaget's theory with concepts from cognitive and differential psychology.. The Neo-Piagetians, as they have been dubbed, attempt to preserve the best of traditional Piagetian ideas and combine them with the results of recent empirical research [8]. Their main representatives are:

Juan Pascual-Leone [9]	Robbie Case [10]
Graeme S Halford [11]	Kurt W Fischer [12]

and *Andreas Demetriou* whose theory seems rather interesting: along with his colleagues they suggest that the human mind is organized in three functional levels. The first is the level of processing potentials which involves information processing mechanisms underlying the ability to attend to, select, represent, and operate on information. The other two of levels involve knowing processes, one oriented to the environment and another oriented to the self [13].

2.3 Fields concerned with the Exploration vs Exploitation dilemma

The EvE dilemma is not limited to human behaviour. It is confronted by fungi deciding whether to concentrate growth at a local site or send out hyphae to sample more distant resources (Watkinson et al. 2005); by ant colonies exploring options for a new nest before settling on and exploiting a particular site (Pratt & Sumpter 2006); by engineers generating algorithms to deploy a fleet of automata to map the expanses of a new environment (Leonard et al. in press) and by machine learning theorists—who coined the phrase 'exploration versus exploitation'—in their efforts to improve the ability of Reinforcement Learning algorithms to function adaptively in changing environments (e.g. Kaelbling et al. 1996) [3]. James G. March set the ground rules in the EvE dillema in Organizational Learning with his article [4]. Since then, the terms "exploration" and "exploitation" have increasingly come to dominate organizational analyses of technological innovation, organization design, organizational adaptation, Organizational Learning, competitive advantage and, indeed, organizational survival [14] (e.g., Benner & Tushman, 2003; Burgelman, 2002; Holmqvist, 2004; Katila & Ahuja, 2002; Lee, Lee, & Lee, 2003; McGrath, 2001; Sigglekow & Levinthal, 2003).

2.3.1 How the human brain manages the trade-off between Exploration and Exploitation

In a landmark paper, Gittins & Jones (1974) developed a straightforward means for calculating the optimal strategy for decision making in Multi Armed Bandits problems (The Gittins Index). Bandit problems are well suited for studying the tension between exploitation and exploitation since they offer a direct trade-off between exploiting a known source of reward (continuing to play one arm of the bandit) and exploring the environment (trying other arms) to acquire information about other sources of reward (Kaelbling 1996). The Gittins Index provides a solution, but applies to restricted circumstances (e.g. only stationary environments). As yet, no general solution has been found for non-stationary environments and, depending upon the breadth and characteristics of the environment to be considered, this may not be possible. Nonetheless, empirical studies of both behaviour and neural mechanisms have begun to reveal mechanisms that animals may use to adapt to changes in the environment, by regulating the balance between exploitation and exploration. These studies appear to be converging on the view that neuromodulatory systems may play a critical role in regulating this balance within certain domains of behaviour. These systems appear to be responsive to both estimates of uncertainty and utility. However, social signals are also likely to be an important source of information. More generally, the trade-off between exploitation and exploration represents a challenge to behaviour at all levels and over multiple time-scales. It is not yet clear whether neuromodulatory mechanisms serve the same function at all of these levels and time-scales, or whether this relies on other mechanisms that remain to be discovered. Given these considerations, it seems probable that further research will require a mixed (though not yet fully informed) strategy of continuing to exploit promising lines of recent work, while considering new ones to explore.

2.3.2 Exploration and Exploitation Balanced by Norepinephrine and Dopamine

As for the strictly clinical approach of the matter, today's cutting edge networks are modeling cognitive phenomena at the level of neurotransmitters [15]. In a great example of this development, McClure, Gilzenrat & Cohen have an article [16] in Advances in Neural Information Processing Systems where they propose a role for both dopamine and norepinephrine in switching behavior between modes of "exploration" and "exploitation." The basic mechanism is that norepinephrine release has two basic modes: phasic and tonic. The phasic mode involves transient increases in norepinephrine, which facilitates processing. In tonic mode, however, the overall levels of norepinephrine are higher, which results in "unpredictable" (i.e., Explorative) behavior [15]. McClure et al. suggest that this model begins to solve the "exploration-exploitation" dilemma of intelligent agents: how do you know when to continue with your current behaviours, and when to seek out other possibilities? The fact that this solution involves norepinephrine is interesting [15]. As for the Temporal Difference Learning Algorithm, it been recognized as an efficient method of reinforcement learning, but was always associated with a cost: once a rewarding stimulus is found, it becomes the focus of attention at the expense of more exploratory behavior. Recent work in cognitive neuroscience has demonstrated how the Temporal Difference Learning Algorithm may be neurally implemented by dopamine fluctuations, and the McClure, Gilzenrat & Cohen paper [16] reviewed by [15] describes how a different neurotransmitter system is used to solve the exploration-exploitation dilemma of temporal difference learning.

2.4 HCI for kids

"In designing for children, people tend to assume that kids are creative, intelligent, and capable of great things if they are given good tools and support. If children can't use technologies we've designed, it is our failure as designers [17]."

2.4.1 Things to keep in mind when designing for preschoolers

In the pre-operational stage (ages 2-7), children's Attention Span is brief. They can only hold one thing in memory at a time. They have difficulty with abstractions. They can't understand situations from another person's point of view. While some children may begin to read at a young age, designs for this age group generally assume the children are still pre-literate. It is reasonable to expect children of this age to be able to click on specific mouse targets, but they must be relatively large. Use of the keyboard is still generally avoided by most designers [17].

HCI research includes:

1. *Dexterity,*
2. *Speech,*
3. *Reading*

4. *Backround Knowledge,*

5. *Interaction Style.*

Dexterity: Young children's fine motor control is not equal to that of adults (Thomas, 1980), and they are physically smaller. Devices designed for adults may be difficult for children to use. Compared to adults, children have difficulty holding down the mouse button for extended periods and have difficulty performing a dragging motion (Strommen, 1994). Kids have difficulty with marquee selection. They may have trouble double-clicking, and their small hands may have trouble using a three-button mouse (Bederson et al., 1996). As with adults, point-and-click interfaces are easier to use than drag-and-drop (Inkpen, 2001; Joiner, 1998) [17].

Speech: Speech recognition has intriguing potential for a wide-variety of applications for children. But, as expected, speech recognition developed for adults will not work well with very young children. Some evaluations have found a recognition rate of only 75%, resulting in a frustrating experience for their subjects. Nonetheless, there has been a remarkable progress towards creating conversational interfaces for children [18, 19]

Reading: The written word is the main vehicle for most communication between humans and computers. Consequently, designing computer technology for children with developing reading skills presents a challenge. Designing for pre-literate children presents an even greater challenge. Audio, graphics, and animation must substitute for all functions that would otherwise be communicated in writing. The higher production values required can add significantly to development time and cost [17].

Background Knowledge: Many user interfaces are based on metaphors (Erickson, 1990) from the adult world. It's helpful to choose metaphors that are familiar to kids, though kids often have success in learning interfaces based on unfamiliar metaphors if they are clear and consistent (Schneider, 1996) [17].

Interaction Style: Children's patterns of attention and interaction are quite different from those of adults. Children are easily distracted. Hanna, Risden and Alexander (1997) used a funny noise as an error message and found that the children repeatedly generated the error to hear the noise. Similarly, Halgren and colleagues (1995) found that children would click on any readily visible feature just to see what would happen, and they might click on it repeatedly if it generated sound or motion in feedback [17]. Children's curiosity and Attention Span should be taken into serious consideration.

2.4.2 What makes Things Fun to Learn? Heuristics for Designing Instructional Computer Games

According to T.W. Malone, the essential characteristics of good computer games and other intrinsically enjoyable situations can be organized into three categories: challenge, fantasy, and curiosity [20].

Challenge: In order for a computer game to be challenging, it must provide a goal whose attainment is uncertain. There are several reasons for believing that goals are important to good computer games [20].

Variable difficulty level: Good computer games should be playable at different difficulty levels. The choice of difficulty level can be either [20]:

1. Determined automatically by the program according to how well the player does.
2. Chosen by the player.
3. Determined by the opponent's skill

Uncertain outcome: A game is usually boring if the player is either certain to win or certain to lose. There are four ways to make the outcome of a game uncertain for players over a wide range of ability levels (or for the same player at different times) [20]:

1. Variable Difficulty Level,

2. Multiple Level Goals,
3. Hidden Information,
4. Randomness

Self-esteem: Goals and challenges are captivating because they engage a person's self-esteem. Success in a computer game, like success in any challenging activity, can make people feel better about themselves [20].

Fantasy: Fantasies often make computer games more interesting. In general, games that include fantasy show or evoke images of physical objects or social situations not actually present. These objects or situations may involve varying degrees of social or physical impossibility from completely possible to completely impossible [20]. According to Malone, Fantasies can be categorized as Extrinsic(fantasy depends on the use of the skill) and Intrinsic (the skill also depends on the fantasy).

Curiosity: Curiosity is the motivation to learn, independent of any goal-seeking or fantasy-fulfillment. Computer games can evoke a learner's curiosity by providing environments that have an optimal level of informational complexity (Berlyne, 1965; Piaget,1952) [20].

Sensory Curiosity: Sensory curiosity involves the attention attracting value of changes or patterns in the light, sound, or other sensory stimuli of an environment [20].

Cognitive curiosity: Cognitive curiosity can be thought of as a desire to bring better "form" to one's knowledge structures. The way to engage learners' curiosity is to present just enough information to make their existing knowledge seem incomplete or inconsistent [20].

Although the game developing industry has accomplished giant leaps towards the User-Gameplay relationship since Malone's theories, they were somewhat "prophetic" and remain contemporary.

2.4.3 Multimodal Systems for Children

Although children are early adopters of new technologies and interfaces, designing multimodal systems for children is challenging both from the core technology development and the human factors standpoint. Core technology challenges include getting speech-recognition technology to work for children users. Interface and human factor challenges have to do with the interaction patterns of children (mix of exploration and exploitation) and the variable capability in using a specific modality (e.g., language, mouse) [1].

Creating Conversational Interfaces for Children: Creating conversational interfaces for children is challenging in several respects. These include acoustic modeling for automatic speech recognition (ASR), language and dialog modeling, and multimodal-multimedia user interface design [18, 19]. In a very interesting experiment by S. Narayanan and A. Potamianos (Agent CHIMP project), about 160 children, ages 8 to 14 years, participated in the study by playing an interactive computer game using voice commands, or keyboard and mouse control. Two applications (communication agent and spelling bee) were implemented to a prototype (using a WoZ experimental setup which will be analyzed later) in order to provide design guidelines for building multimodal-input multimedia-output applications for children. During the subjective user evaluation, an exit interview took place where the vast majority of children gave very high marks to the speech interface (93% rated the interface 4 or 5) [18, 19].

Performance of ASR and SLU: Due to inherent variabilities in children's speech, acoustic modeling for automating recognition of children's speech is a challenging task [18]. Potamianos and Narayanan improvement of language and dialog modeling, greatly improved Spoken Language Understanding (SLU) performance. More specifically, an approximately overall 90% dialog action understanding accuracy was achieved for their task. These results show us that it is feasible to build successful spoken dialog systems for children using existing ASR/SLU technology [18].

Dialog Strategies: Keyboard and Mouse versus Voice: In another instance of their experiment, Narayanan and Potamianos examined the performance of 12 children players alternating on using voice and keyboard-mouse modalities. Children took fewer turns (almost 50%) using keyboard and mouse than voice to carry out the relatively high-perplexity database search and retrieval tasks. This suggests that for the database search task voice is not the most efficient modality(with the current interface). A final observation is that when using K+M, superfluous greetings at the navigation/query menu (dialog state: "Goodbye") were reduced by a factor of three compared to using voice. This reinforces the belief that *although speech might not be the most efficient modality always, it is the most natural one [19].*

2.4.4 Children's and Adults' Multimodal Interaction with 2D Conversational Agents

Another project of a multimodal game system with Embodied Conversational Agents (ECAs) by Stephanie Buisine and Jean-Claude Martin, suggests among other things that speech-only ECA systems might not be so comfortable for young children [21]. Children's preference for gestural interaction and initiatives in their experimental process might be linked to their playing habits, and also to a kind of shyness that was noticed when they spoke to the ECAs [21]. Two groups of French-speaking users participated in the experiment:

1. 7 adults (3 male and 4 female users, age range 22 —38)
2. 10 children (7 male and 3 female users, age range 9 — 15).

As far as modality usage is concerned, gesture-only was the most widely used modality (47.5% of inputs). Speech-only accounted for 31.3% of inputs and multimodality for 21.2%. The frequent use of gesture in their corpus (47.5% of inputs) is surprising regarding the literature on multimodal interfaces (much smaller percentages). Some age independent (no difference between adults and children) results regarding Verbal Behaviour were [21]:

1. User utterances (average 6 words).
2. The 91% of user's utterances with Agents would be appropriate to Human-Human interaction (natural utterances, used in Human-Human conversation).

3. Average two social cues per minute in user's behavior (politeness, feedbacks on Agent's speech or actions)

Adults and children did not differ regarding the semantic patterns they used (Redundancy between speech and gesture, "Classical" complementarity, "Dialogical" complementarity, Concurrency, Multimodal repetition) except for concurrent constructions. Indeed, 18 concurrent constructions (ex. Saying "Hello" and pointing on a cake on the floor) out of 19 in the corpus were produced by the children [21].

2.4.5 The NICE fairytale game system

The goal of the NICE project [22] is to allow users of all ages to multimodally interact with lifelike conversational characters in a fairy-tale world inspired by the Danish author H. C. Andersen. To make these characters convincing in a computer game scenario, they have to possess conversational skills as well as the ability to perform physical actions in an interactive 3D world. What primarily distinguishes the NICE fairy-tale game system from other spoken dialog systems is that the human-computer dialog takes place within the context of an interactive computer game. However, spoken and multimodal dialog is not supposed to be just an 'add-on' to the game, but the user's primary means of progression through the story [22]. Secondly, what makes NICE differ from typical spoken dialog systems is the attempt to move away from strictly task-oriented dialog. Instead, the interaction with the characters is domain-oriented. This means that the dialog concerns different subplots in the fairy-tales, but without a clear goal/orientation and without other demands than it being entertaining to the user [22]. Thirdly, a feature that differentiates NICE from other systems is that the main target user group of the system is children and young users. Previous studies have indicated that children employ partly different strategies when interacting with dialog systems than adults do, and that there are also differences between age groups [22]. The NICE project emphasizes on Explorative dialogs, in which users are encouraged to browse through information without pursuing a specific task (presented by Cassell et al. 1999 and Bell et al. 2001).

2.4.6 Ethical implications with conversational agents

Recently however, there has been a certain scepticism about whether Embodied Conversational Agents open any ethical issues. These issues according to Antonella De Angeli [23] concern in one part the ECA approach and how it found its most influential justification within the field of HCI in the Media Equation Paradigm, a research framework stating that computers are social actors to which users respond as they would do to real people (Reeves & Nass, 1996). She states [23] that this way of thinking deeply influenced the proposition of the Persona Effect (Affective Impact of Animated Pedagogical Agents) who claim that ECA's have the potential for expanding the interaction sphere between humans and computers to encompass social connections and positive emotions leading to more productive task performance (Lester et al., 1997). She expresses [23] disbelief towards the meta-analysis (Yee, Bailenson & Rickertsen, 2007) and it's empirical support to the Persona Effect. On her critical Rhetorical approach she takes into consideration the full range of human behaviour, including negative outcomes and the possibility of conflicts, disinhibited (see Disinhibition) and antisocial reactions to ECA's (Brahnam & De Angeli, 2008; De Angeli, Brahnam & Wallis, 2005). De Angeli believes towards a deontological requirement to initiate a serious reflection within professional bodies and guide the designers of ECA's in their difficult challenge, possibly to enhancing technological development within a value-based approach (Dix, 2008; Freier, 2008; Whitby, 2008) and that there is an intrinsic risk in the potential of conversational agents to elicit Disinhibition and stereotyping [23].

Chapter 3

Platform Design and Implementation

3.1 Introduction

The main advantage of building a web-based platform for multimodal game-development is that it can be used for (remote) data collection and analysis of educational software and games. The collected data can be used to train language and acoustic models for automatic speech recognition (ASR) and for analysis of user interaction patterns to improve or adapt the user interface. Educational software and games are also used extensively by linguists and psychologists, e.g., to diagnose and solve language development problems [1]. In his work, Theofanis Kannetis [1] implemented the platform to study the children-computer interaction for preschool children and to see how Malone's quality characteristics (Challenge, Fantasy, Curiosity) can be adapted in order to improve the user experience. In this project, we used —and extended by a small percentage this platform to examine the EvE dilemma and how it is perceived by preschoolers during their cognitive process of the preoperational stage. We also integrated the mini-games developed by Kannetis [1] and Meliopoulos [2] in their preceding work. These games where designed for the exactly same ages as our target group (4-6) and therefore were ideal for the Exploitation part of our application. The following sections in this chapter are mostly abstracted by Kannetis Master Thesis (chapter 4) [1] where he describes the Platform's Architecture and Modules

Implementation which we also used. All credits for the platform's design and implementation should go to him; we merely extended the WoZ GUI and changed the logfiles storage procedure used by the Application Manager (Server Part) to fit our application's needs.

3.2 Platform Architecture

The system follows a modular architecture, the full functionality of the system being the result of the collaboration between the modules. The architecture of the system is shown in Figure 3.1. Since this is a web-based platform it is (by

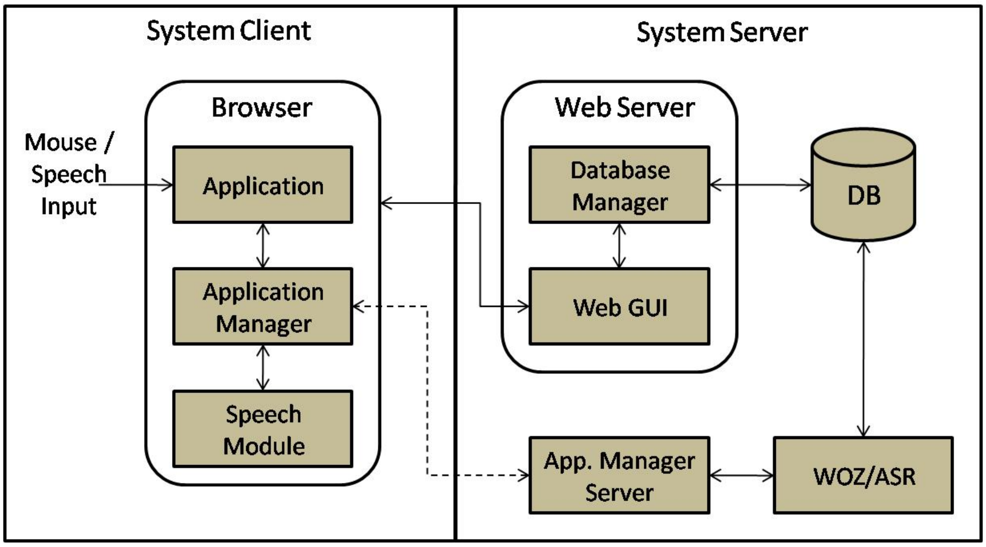


Figure 3.1: The modular architecture of the platform.

nature) distributed. The Application Manager is responsible for the synchronization and cooperation of the modules. It consists of two parts that follow the client/server architecture, i.e., the client/browser side and the server side. The two parts communicate through a two way socket connection. The Speech Module is responsible for capturing and streaming the audio, as well as, performing the voice activity detection (VAD) to determine if the user is speaking. When voice activity is detected the Speech Module starts the audio capture. At the same time, the streaming of audio data to the client part of Application Manager begins. Finally, the multimodal Application module may contain any interactive application implemented by the system designer. In our case, we have developed an environment with ECA's (Exploration part) where the children are prompted towards explorative (not goal-oriented) conversation, and integrated the aforementioned mini-games (Exploitation part). These will be thoroughly analyzed in the next chapter(4). The whole application was developed in Flash[24] in order to provide an easy and platform-independent way to manipulate sounds, animations and graphics. The only downside of implementing these pre-made games was that they were developed in ActionScript 2.0 (AS2) and therefore we were restricted from using newer versions (AS3). There was of course the option of converting the code in AS3, but as it proved, this was a time devouring process and so we decided to develop our application in AS2 too.

On the server side of the system, the Application Manager (server part) is responsible for executing the speech requests that are being received from the client side of the platform. It receives and then streams the audio data to the ASR/WoZ module for automatic recognition or manual transcription (by the wizard). The result of the recognition is send back to the client part of the Application Manager. The module also receives and stores the necessary logfiles for further processing. On the server side of the system, we have also implemented the Web Interface Module and the Database Communication Module. The web interface was designed using Java Server Pages (JSP) technology. Using the web interface users can register and login to the platform. Functionality such as profile management and preferences configuration (e.g., microphone configuration) is also provided. The Database Communication Module is responsible for all the necessary database queries. Both the Web Interface and Database Communication modules are parts of the Apache Tomcat Web Server. The ASR/WoZ module has been implemented on the server side of the system. In this study, the ASR module has been replaced by a Wizard of Oz (WoZ) module which is operated by a human transcriber. The WoZ module is actually a graphic user interface (GUI) that plays the audio stream received by the Application Manager and allows the wizard to supply the appropriate transcription via a GUI interface. Both the audio and transcription files are stored in the database.

3.3 Modules Implementation

3.3.1 Speech Module

The Speech Module is responsible for capturing and streaming the audio, as well as, performing the voice activity detection (VAD) to determine if the user is speaking. The module is implemented as a java class using the Java Sound API. Three main functions have been implemented in the module. CaptureAudio function is responsible for capturing and streaming the audio to the Application Manager(server part) using the Java Sound API. The captured audio is in PCM 16000Hz 16bit stereo format. VoiceDetection function is responsible for Voice Activity detection (VAD). Voice activity detection (also known as speech activity detection or, more simply, speech detection) is a technique used in speech processing wherein the presence or absence of human speech is detected in regions of audio (which may also contain music, noise, or other sound).

VAD is an important enabling technology for a variety of speech-based applications. Therefore various VAD algorithms have been proposed that provide different compromises between latency, sensitivity, accuracy and computational cost. Kannetis has implemented a very simple algorithm using the energy of the captured frames by computing the Root Mean Square (RMS) for each frame. When RMS is above a specific threshold (which can be changed through microphone calibration), the CaptureAudio function is called. When RMS is below that threshold for a short period of time, the StopCapture function is called, which is the function that stops the recording procedure.

3.3.2 Application Manager

The most important part of the system, is Application Manager module. As we already said, Application Manager is distributed and follows the client/server architecture. The client part of Application Manager was implemented as a java applet and it is running on the client part of the system (the browser), while the server part of the module was implemented as a java Multithreading server.

When a VAD event occurred, the client part of the module sends a "voice request" message to the server part of Application Manager so that the streaming connection could be established. Then the module waits for the appropriate response from the server part of the Application Manager. If the response is positive, the Application Manager notifies the speech module, so that the recording and the streaming of the audio data can start.

When the user stops talking and a stop speech VAD event has occurred, the client side of the Application Manager notifies the server side of the module, that the streaming is over and then waits for the audio recognition result. After it receives the answer from the server part of the module, the result is sent to the multimodal application where the appropriate action takes place, depending on the answer.

The server part of the Application Manager is responsible for receiving the audio data and send them to the Speech Recognition Module, which in our case is the Wizard of Oz module. When it has the result of the recognition, the appropriate message is send to the client part of the module. Also the server part of Application Manager is responsible for receiving and storing the logfiles of the application for further process. The server part of the module is implemented as a java Multithreading server using the network API of the Java. The communication between the client side and the server site of the module is occurred through a two way TCP/IP socket. The process of a speech request through the system is shown in figure 3.2.

In order to achieve the cooperation between the client and the server part of the Application Manager, a communication protocol has been implemented. A fairly simple protocol was built for the communication between the two part of the Application Manager based on simple text messaging. Table 3.1 shows the protocol that Kannetis used for all the communications between the client and the server part of the Application Manager.

App. Manager Client	**App. Manager Server**	**Action**
VOICEREQ#	VOICEREADY#	Audio streaming
—	AUDIODONE#	Audio streaming
ANSWEROK#	ANSWER#result	Speech Recognition result
WRITELOG#log	WRITELOGOK#	Saving Log File

Table 3.1: The communication protocol between client and server parts of the Application Manager.

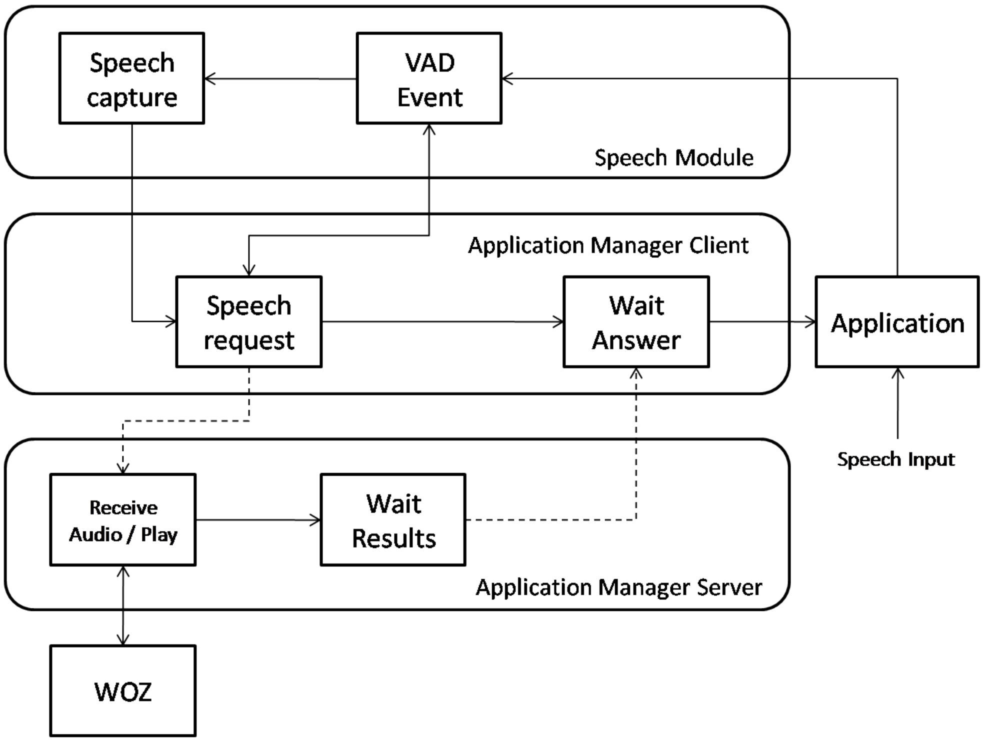


Figure 3.2: Service of a voice request process.

3.3.3 Web interface and Database

As we decide to build an on-line system, the appropriate web user interface should be implemented. Using the web interface, users could register and log-in to the platform so that they could use it. Also profile management and microphone calibration is provided through the web interface. The web interface has been implemented using the Java Server Pages (JSP) technology. In order to store information regarding to users a database communication module has been implemented. The module is responsible for all the database transactions including register and log-in. The module is actually a collection of JavaBeans that each one of them is responsible for a specific job. JavaBeans are classes written in Java conforming to a particular convention. They are used to encapsulate many objects into a single object (the bean), so that they can be passed around as a single bean object instead of as multiple individual objects. The JavaBeans classes are called through the web interface so that the communication between the database and the web interface could be established. Both JSP pages and

JavaBeans are hosted on Apache Web Server. An example of the implemented beta web interface is shown in figure 3.3.

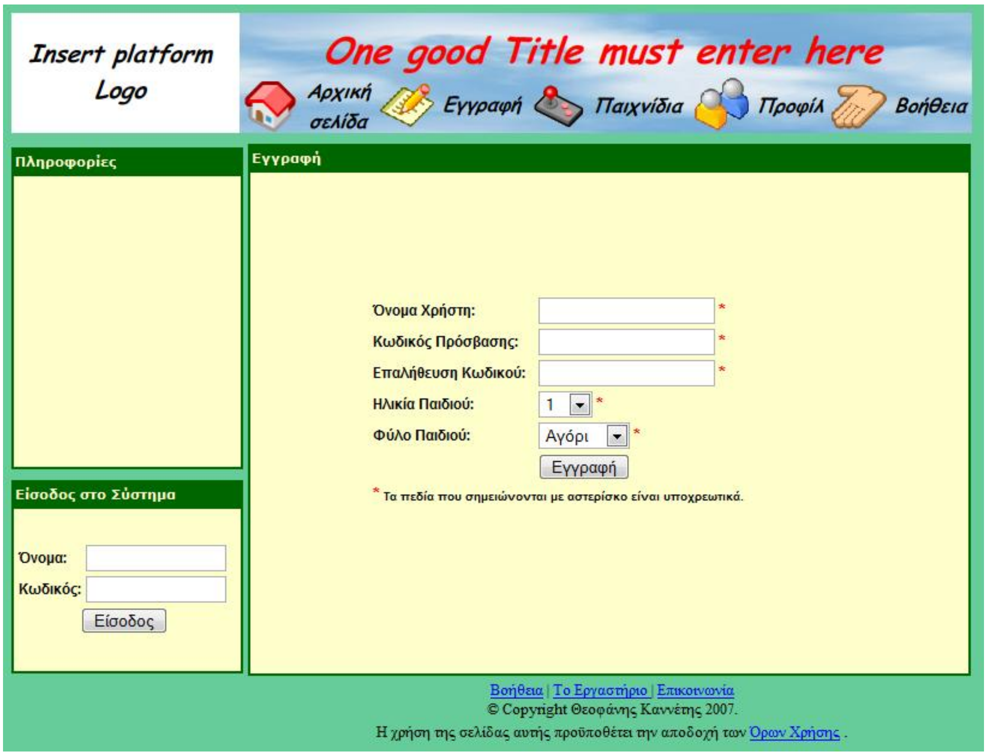


Figure 3.3: Web Interface Example

3.3.4 ASR and Wizard of Oz

The Wizard of Oz (WoZ) technique is an experimental evaluation mechanism. It allows the observation of a user operating an apparently fully functioning system whose missing services are supplemented by a hidden wizard. The user is not aware of the presence of the wizard and is led to believe that the computer system is fully operational. The wizard observes the user through a dedicated computer system connected to the observed system over a network. When the user invokes a function that is not available in the observed system, the wizard simulates the effect of the function. Through the observation of users behavior, designers can identify users needs when accomplishing a particular set of relevant tasks and evaluate the particular interface used to accomplish the tasks. In our system, the

WoZ component was implemented in JAVA and replaced the ASR component which was described above in this subsection. Our application uses the VAD component to notify the presence of speech to the Application Manager which streams the audio to the Wizard of Oz interface. It then replays the sound to the wizard who chooses the appropriate answer and sends it back to the game.For the synchronization of the various components (Application, VAD, Application Manager, WoZ) to be achieved, the Wizard of Oz application has a text flag (wait - give answer) which notifies the wizard when to give the answer. Also note that there are different tabs in the WoZ GUI, one for each of the tasks developed by Kannetis and Meliopoulos (see Figure 3.4, and one for the ECA environment we developed (see Figure 3.5. The tabs are selected automatically reflecting the user's choices.

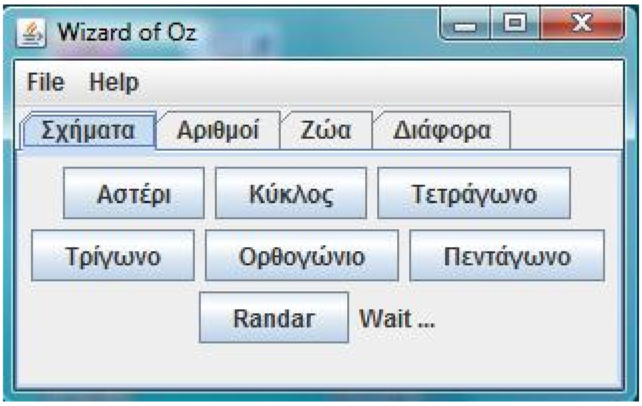


Figure 3.4: The Wizard GUI

3.3.5 Multimodal Application

The last module of the system is the multimodal application module which may contain any interactive application implemented by the system designer. These applications could be implemented in any web-based programming language. The communication between the application and the client part of the Application Manager was implemented with Javascript so that the communication between the application and the Application Manager be independent of the implemented language (almost all of the web-based languages support javascript calls). In our case, we have developed a setting where children engage in conversation with

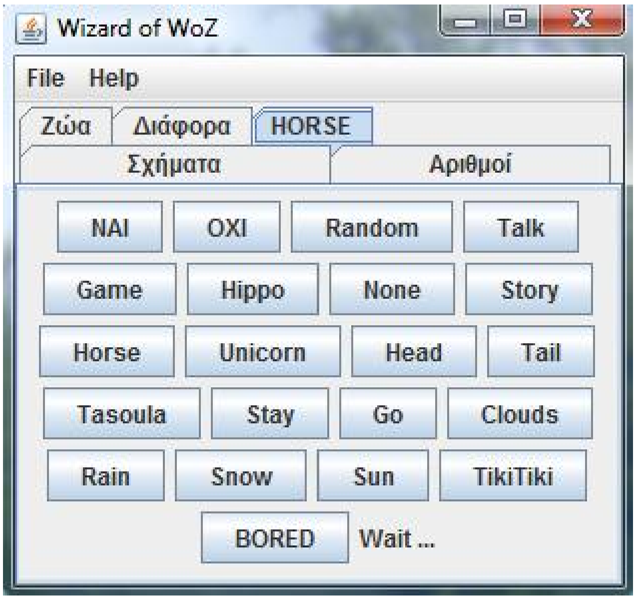


Figure 3.5: The new WoZ tab we added for the Exploration part

ECAs, and integrated the five tasks of Kannetis and Meliopoulos in a single application, using (as mentioned) Flash [24] ActionScript 2. In the next chapter we will analyze our application's structure.

Chapter 4

Application Design and Implementation

4.1 Introduction

Today's computer games are limited by the user's input options, which are often restricted to direct manipulation and simple commands. In the development of the next generation of computer games, multimodal dialog has the potential to greatly enrich the user's experience [22]. The first steps have been made towards that direction, by game developers who have integrated Voice Commands in their games, and by independent third-party applications which translate Speech Requests to compatible input methods (keyboard/mouse).

In Chapter 2 we examined some aspects of the preschooler's cognitive development, the EvE dilemma and HCI for kids. It has become clear then, that developing any kind of application for children is rather challenging. Little research has been made in this field regarding kids in the pre-operational stage and their interaction patterns with computers (and ECA's more specifically). The main goal of this work is to examine how children instinctively respond to the EvE dilemma by allowing them to freely switch between an Explorative and an Exploitative environment. The human-computer dialog takes place within the context of an interactive computer game where spoken dialog is the user's primary mean of progression through the Exploration part of the game. So far our application's multimodality supports Voice and Mouse with the perspective of implementing Touch in the future. In the following sections we will take a closer look to the game we designed and see how Exploration and Exploitation concepts were implemented in it.

4.2 The Exploration Mode

In the introduction of our game, the kid is presented with a small animated sequence with many Pandas appearing on the screen and performing some funny actions (figure 4.1). This may seem superfluous but it is a solid tactic (the

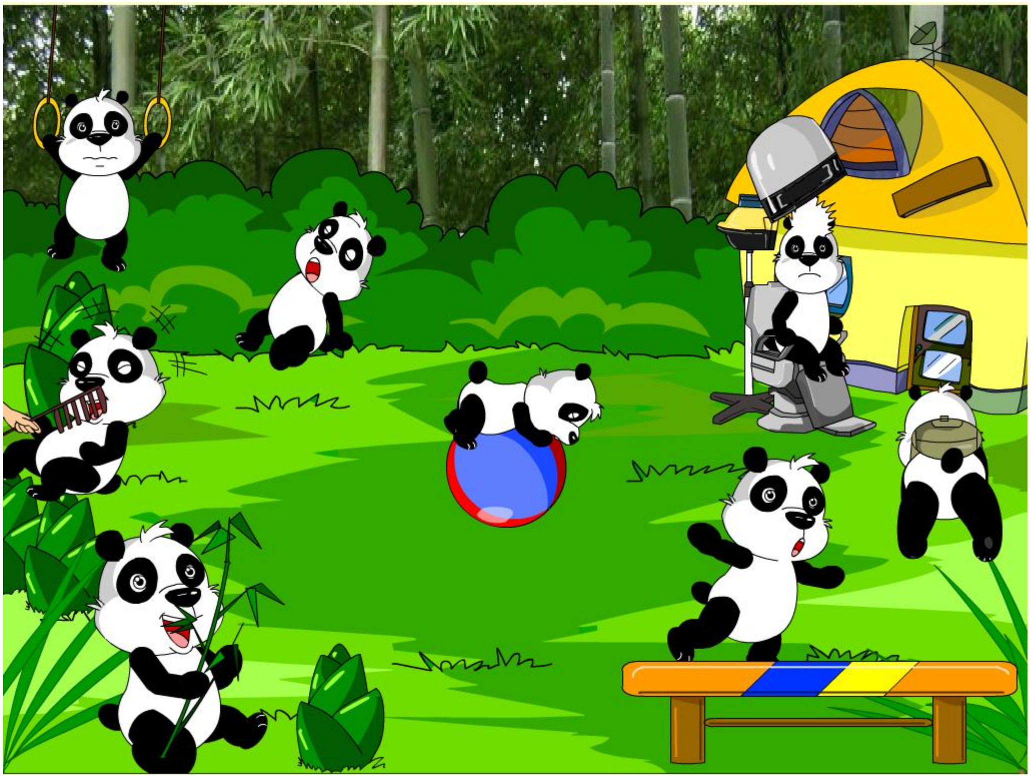


Figure 4.1: From Game: Introduction.

introduction that is) used in every modern game in order to help the user get acclimatised. Right after this sequence, all of the Pandas disappear except for one, which will be the child's *Companion* in his/her playthrough. The use of a companion is believed to be necessary, specifically for such small ages, providing hints and guidelines in cases of certain inactivity. It also serves the purpose of enhancing the user's intimacy with the environment. In the first audio listened by the child, the companion introduces herself (figure 4.2) and in a small

Figure 4.2: From Game: Companion.

tutorial explains how things work (regarding voice and talking to characters) and prompts the child to notify the supervising student about any discomforts (headphones/microphone disturbances or very loud audio) before they start their adventure.

4.2.1 Environment

The Exploration part takes place in a familiar and easily recognizable by children place....**The Stables** (figure 4.3). If the kids agree to go inside they, are presented with the use of mouse (figure 4.4 which may be used in the Exploitation part or some sequences in Exploration (where we don't demand proper use of mouse skills and forward the plot in the event of mouse difficulties; details of mouse usage in a next section). Right after, they meet the main ECA who is materialized in

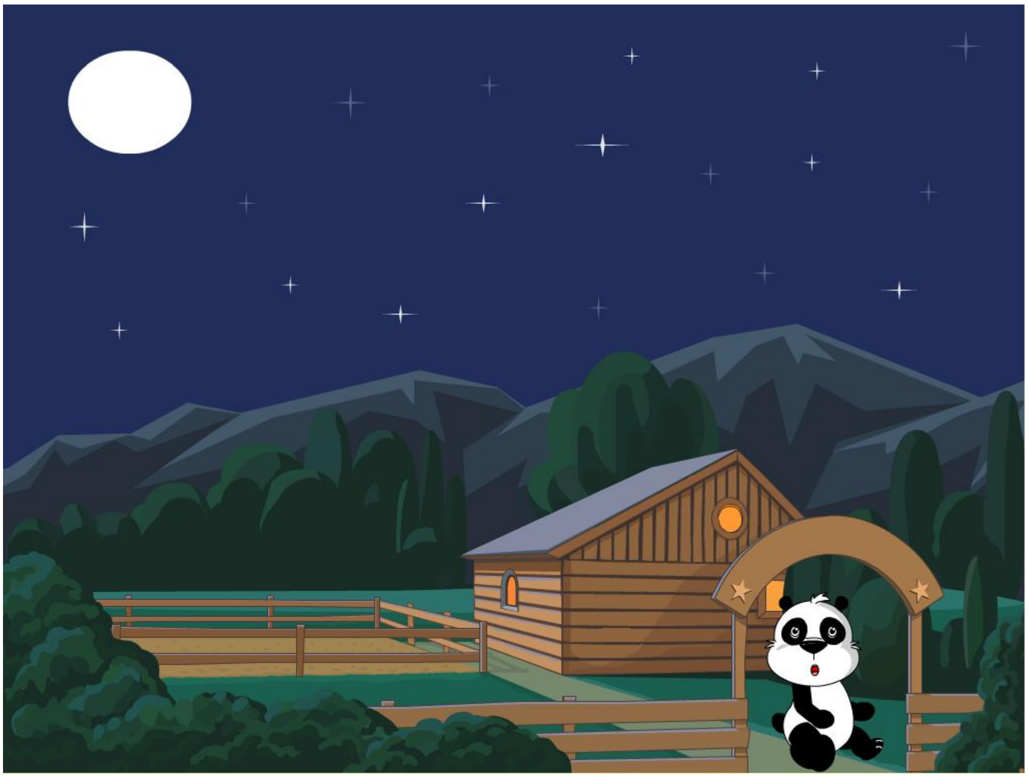


Figure 4.3: From Game: Stables Outside.

the form of a horse. There are also some other ECA's (apart from horse and panda) but they serve a limited purpose and don't have much recorded voice time (chicken and twin cats).

4.2.2 The choice of animals as our ECA's

The choice of animals as our ECA's was made due to the fact that it is widely observed that young children are fascinated by animals and care deeply about them. Recent research has revealed that animals are so important to young children that they routinely dream about them. In fact, 3 to 5 year-olds dream more frequently about animals than about people or any other topic, and animal dreams continue to be prominent at least until the age of 7 years [25]. Animals often lower a child's anxiety, and story-telling including animals is used in Animal Assisted Therapy (AAT), where the animal serves as a bridge between the therapist and the child. By lowering the child's anxiety, the animals aid in creating

Figure 4.4: From Game: On entering the Stables. Introduction to mouse usage.

a nonthreatening atmosphere whereby the child can express his or her feelings [26]. Piaget found that all children go through a stage of development when it is natural for them to ascribe human traits to animals (Piaget 1929). From a pure curiosity perspective, children may find intriguing the fact that they have the 'ability' to converse with an animal —rather than a (maybe stereotypical) human representation.

4.2.3 Exploration Mode Structure

Our Explorative part is designed by the following concept:

- In their *first playthrough*, the application has a somewhat linear form in the beginning in order for the proper introductions to be arranged (how did you get here? what's your name? etc.). There was no point in implementing randomness yet, because the user won't go through the beginning part of the first playthrough for a second time. After a few moments of dialog the child is prompted to an obligatory switch to the Exploitation part so that all possibilities of playing modes are presented to him/her. When the child decides to return from Games/Exploitation , apart from general talk, the main ECA (horse) has a random story about his world (and how he got in the stables) to tell (more on that in *Dialogs used and Story Information subsection*). If the child decides to help the horse with his problem, then our ECA is transformed into a Unicorn. After that there is the option of staying a little longer and converse with the Unicorn, going to the *Random Quotes Section* (analyzed later), or letting the Unicorn go and thus ending the whole session. Note that there are many strategically placed questions at various points of the game, where the child has to make an Explore (remain at stables) or Exploit (go to games session) decision. See figure 4.17 for the first playthrough FSM.

- In *any subsequent playthrough* (revisit) the Exploration Mode changes significantly. The ECA is now a Unicorn. The first dialogs now follow a random path of alternatives and are designed in a manner so that they are thematically correlated to each other, and keep a certain flow. They also reflect the facts that happened during and since the first playthrough. The choice of mode switching is now presented from the first dialog sequence and nothing is obligatory this time. The child has the liberty of choosing to play games, listen to the Story again or engage general talk (Random Quotes). Again, there are frequent Explore/Exploit decision breakpoints, especially during the Random Quotes section. See figure 4.18 for any subsequent playthroughs FSM.

4.2.4 How Conversation works

Children of this age may find bizarre the fact that they have to speak to a 'monitor'. It would be rather difficult to employ an ECA in the form of a 'huge children interests' database and leave the initiative of talk to the child. Instead our dialogs are inspired from a background story and the initiative of talk lies in our ECA's, which prompt the child to talk, with various questions. When the child has a basic grasp of how things work they are also prompted at certain points to ask questions themselves. These questions are undermined by the Companion. Another thing we kept in mind is the limited Attention Span that is observed at preschoolers. Especially at this age, kids become easily distracted. In one moment they may be immersed in the game experience and in another one they may lose their interest. For example it was observed during the evaluation that a long monologue (one minute or so) by the ECA sometimes ruined the immersion because of the time the kid had to remain inactive. Subsequent short conversations on the other hand keep the child vigil and immersed (more on *evaluation* in the next chapter). In the possibility of the children been distracted, we implemented Time Outs (TOs) almost in every question. So if the user remains inactive after the end of a question, the question will be repeated in a different manner after ten seconds. We say in a different manner because maybe the child was not distracted but reluctant at answering. This phenomenon was observed at some children during the first moments of their first playthrough. So, this different manner is usually a more casual one, for example adding the quote "Come on tell me! Don't be shy!.....(question)...... " before posing the question again.

4.2.5 Content, shape and progression of Dialogs and Story

In the first playthrough as mentioned before, the initial part is somewhat linear in order to make the proper acquaintances with the ECAs. The very next thing is to introduce the user (obligatory first visit) to the Exploitation part of our application so that all playing modes are revealed. After spending as much time as he/she wants, the child may return to the Exploration part again where the main ECA unfolds his *Story* . In the Story part the main ECA describes how he was a Unicorn and reveals some random information about his world-family and how he got to the Stables. The Story is divided into five parts —where each part has some alternatives which are randomly chosen , and each one of them being conceptually related to the next one (see figure 4.19 for the story's structure). During this story-telling the child is not inactive. At the end of each part the user is asked a question related to the concept of the story's part. For example, in one story part the ECA describes his home town which is located above the clouds. As a Unicorn he doesn't have trouble getting there but all non-flying visitors have to use a rather strange mean of transportation: A giant squid... Then he asks the child about his/her home location followed by the question "Is your home above clouds too?". Another thing we should mention is the visualization part of the story. During the whole story narration some 'Thought Clouds' appear, as if they represent the visualization of the ECA's thoughts. These clouds were synchronized to the ECA's speech so that they reflect his words. For example when he talks about the apple he stole from the *Tree of Life* while gazing at the *Twin Moons* of his world, the appropriate Though Clouds with the tree and the moons appear (figure 4.5). After the final part of the story, the ECA asks for the

Figure 4.5: From Game: The Thought Clouds.

user's help (they have to successfully complete a game in Exploitation mode) in order for him to transform back to a Unicorn and return to his world. Actually it doesn't matter if the kid is successful or not. After three tries we give him/her the reward so as to forward the plot and avoid possible frustration on the user's side. In any *subsequent playthroughs*, the initial part of the conversation is structured with the same principals (random parts but thematically correlated) the Story follows. The Story-telling is now optional. If the user decides he wants to hear the story again, he has also the option to specify the shape of the ECA: Horse or

Unicorn (figures 4.6 4.7).

Figure 4.6: From Game: The Thought Clouds. The Unicorn Brothers

4.2.6 The Random Quotes section

The Random Quotes (RQ's) are a state of generic small talk. According to [27], to convey personality and to build a collaborative trusting relationship with the users, the characters also have to be able to engage in small talk. The user can go to the RQ's section in the first playthrough if they decide to talk further with the Unicorn instead of letting him go home towards the end, or in any subsequent playthroughs after the initial conversations. Each RQ has a certain theme and lasts one-two minutes. At the end of each theme, the user is prompted to make an Explore/Exploit decision by the Companion. If he/she decides to stay and talk then we advance to the next RQ theme. Otherwise we switch to Exploitation mode (Games) and when the user returns, the RQs continue. This section is designed in a way so as not to hear the same segment of speech twice.

Figure 4.7: From Game: The Thought Clouds. The Shadow Forest

In the figure 4.20 we can see the structure of the Random Quotes. The RQ1(a-f) are the companion's Quotes prompting the user to make an Explore/Exploit decision. The RQs(2-6) are the themes we mentioned above. *Example: The user is asked from a random pool of utterances whether he/she would like to talk with the Unicorn some more or would prefer to play some games in the Companion's yard[RQ1(a-f)]. If he/she decides to talk, a Random Quote(2-6) is generated and if it is already listened to we re-generate a quote. Let's say RQ4 was chosen.... Here the ECA informs the child that he has a problem coordinating the movement of his head and tail in his new Unicorn form. Then, he prompts the user to ask him which of the two he/she would like to try and move. The possibility of moving the correct part (head or tail) is 50%. Then either the Companion or the Unicorn make a comment depending on the outcome (success/failure). Two more similar attempts are made and finally the Unicorn makes a final comment depending on the success/fail ratio and thus ending the RQ4 theme. Then, the*

*Companion makes a RQ1 question (Games/Talk decision) and so on....*This is a very important part of the application due to the distribution and frequency of Explore/Exploit decisions.

4.2.7 Voices and recordings

All the audio listened to in the Exploration part is pre-recorded. We made the decision of processing the audio with some effects and not leave it in it's natural form. With *pitch and speed* changes the voices became more cartoon-like and in a way that we thought kids would like (having recent memories of cartoon series). The Horse's/Unicorn's (main ECA) voice had it's pitch reduced by 10% and it's speed by 6%. This made the voice rather deep and out of the ordinary. There was an issue whether kids would find such audio scary. We shall refer to this in the Objective Metrics of Chapter 6. The Companion's voice (female) had it's pitch increased by 12.5% and resembled that of a small kid. Finally we removed any noise and clips. Even though we tempered with pitch and speed, we always had in mind to keep the audio clear and easily understandable by children.

4.2.8 ECA's personal traits

In is rather important that our ECA's have some distinct personality traits. This can have a catalytic effect in the aspect of relative realism and user's immersion. Below are the main personality traits of our application's ECAs:

Hosrse/Unicorn

- Name: Sifounas (Greek word for hurricane)
- Attitude: Friendly and kind
- Voice: Low pitched and somewhat slow.
- Likes: Apples, carrots
- Argues: with the twin cats
- Other: Eager to express unmeasured gratitude in cases where the user decides to help him with something. Apologizes when he talks a lot (during

some monologues). Homesick, likes to talk about his world. Finds our world very strange. Likes small talk.

Companion

- Name: Tiki-Tiki (Funny pronunciation in Greek and relatively easy to remember by the kids)
- Attitude: Very Friendly and helpful. Is the user's Companion and acts accordingly.
- Voice: High pitched and jolly.
- Likes: The user, her yard with the games (Exploitation Mode), throwing a nap once in a while.
- Argues: with the Horse/Unicorn occasionally.
- Other: Is the primary mean of switching between the two modes (Explore/Exploit)

Twin Cats

- Names: Litsa/Nitsa
- Attitude: Teasing.
- Voice: Very High pitched and fast.
- Likes: To make bets with the Horse/Unicorn
- Argues: with the Horse/Unicorn occasionally.
- Other: Limited recorded voice playback time

Chicken

- Names: Tasoula
- Attitude: Bored.
- Voice: Low pitched and very slow.

- Likes: To sit on her eggs and do absolutely nothing.
- Argues: with no-one.
- Other: Limited recorded voice playback time

4.2.9 The use of Mouse

Our application supports the use of mouse but in very few sequences. Furthermore, we don't demand proper use of mouse skills in these sequences. Help is given to the to the user in the form of Companion tips or by progressing the sequence nonetheless. For example: *The Unicorn is hungry and asks of the child to gather some carrots from the floor (see figure 4.8). At 10 seconds of inactivity the Companion helps the user by collecting the carrots herself and thus progressing the plot. For every carrot gathered the user gains 5 more seconds. If the user clicks on the cart full of carrots then the Companion notifies him that these are for the farmer and should gather the ones in the floor.*

Figure 4.8: From Game: Gathering some carrots for the Unicorn

This compromise was made in order to keep the 'feeling' of exploration intact. To demand and use any kind of skills is purely Exploitative. In 4.9 there's another example of minimal mouse usage: *The companion prompts the user to take the mouse and click on the (rather big) broom. The user discovers a key but that key alone can't unlock the chest.... A magic word is needed. After a small dialog sequence between the two ECAs the Unicorn remembers that magic word eventually, and the user has to reproduce it in order to retrieve the papyrus that was locked away (as a favor to the horse).*

Figure 4.9: From Game: Just a click on the broom for the key and then a magic word to unlock the chest.

4.3 The Exploitation Mode

As mentioned in previous chapters, the Exploitation part of our Application was designed by Theofannis Kannetis (except the Animal Recognition task [2]). The following section is abstracted from his Master Thesis [1] where he describes each game .

Five different popular preschooler tasks were selected for implementation. The selected tasks were: animal recognition , shape recognition, quantity comparison, number recognition and addition. For each game an embodied agent guides the child through the task. Both mouse and speech are enabled as input modalities. Animation, sounds, graphics, prerecorded prompts and synthesized text-to-speech prompts (where necessary) were used as output. The list of tasks is described next:

- The animal recognition task is taking place in a farm. First the voice of an animal is heard and then the farmer asks the child to select the appropriate animal in order to guide it into the farm. There are (up to) nine different farm animals in the game (see Fig. 4.10). The specific task has been implemented by Spyros Meliopoulos as a part of his diploma [2].

Figure 4.10: Animal recognition task.

- The number recognition task takes place at the beach where an animated character (squirrel) asks the child to identify which number (1-9) is shown on the screen (see Fig. 4.11).

Figure 4.11: Number recognition task.

- The shape recognition game takes place in a theater. Each time one of the shapes (star, circle, square, rectangle, triangle, pentagon) appears on stage and the child must identify it. The animated character (teacher) provides help and guides the child through the task (see Fig. 4.12).

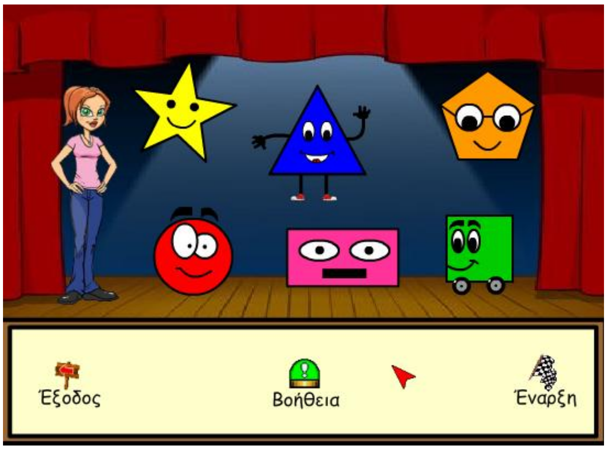


Figure 4.12: Shape recognition task

- For the comparison task, an animated character (rabbit) puts some items inside and some outside of a basket. The rabbit asks then the children to determine whether the items inside the basket are more (or less) than those outside (see Fig. 4.13).

Figure 4.13: Quantity comparison task

- For the addition task, the child must help an animated character (bear) to collect some honey from the beehives. A simple addition task appears on the screen (sum up to 9) for the child to perform. For each correct answer a bee fills a honey jar with honey (see Fig. 4.14).

Figure 4.14: Addition task

For an idea on how Fantasy, Curiosity and Difficulty were implemented to these tasks (Malone's Quality Characteristics [20]), there is a thorough analysis in [1]. Table 4.1 shows the parameters we used in our application. According to Kannetis, systems with high values of fantasy, curiosity and difficulty were the most popular among the children [1]. We used Fantasy level 1 (the next most popular) because at level 2 there was a little story implemented (thus the popularity) which we didn't need in our work (we have implemented the whole Exploration part). As for the Difficulty, during the first playthrough we used a difficulty setting of 1 and in the subsequent playthroughs a setting of 2. In figure 4.15 the popularity of various system setups is shown.

Fantasy	1
Curiosity	2
Difficulty	1-2

Table 4.1: Parameter used in Game Sessions (Exploitation Mode)

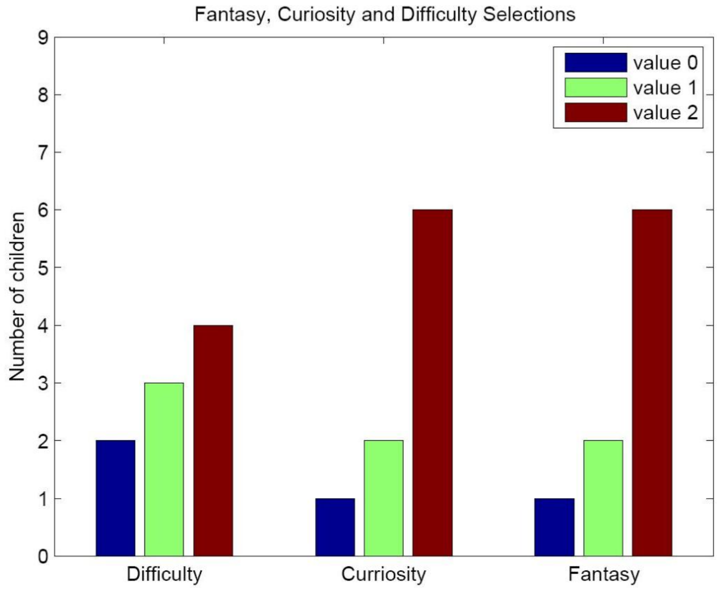


Figure 4.15: Histogram of subjective optimal levels of fantasy, curiosity and difficulty.

4.3.1 Integrating the Games

Since we had access to Kannetis and Meliopoulos source ActionScript code, this was a relatively easy task. Whenever the user decides to go to Exploitation mode, he/she is moved away from the stables environment and into the Companion's Yard. In the first visit of this mode (session's first time of either first's or subsequent's playthrough) the child is introduced (or reminded) to how things work in the Yard. The Companion informs the user that when she sleeps, she dreams of some games which the kid can play (inside a giant Thought Cloud/Frame—see figure 4.16). Whenever the user wants to return to the Stables Environment (Exploration) all he/she has to do is a mouse click on the sleeping Panda or shout her name. This can be done at any point of the games session (even while playing a game) but in most of the cases the children completed the game first before switching modes. When the Companion wakes-up after the user's prompt, she says a couple of random lines to 'smoothen' the transition between the two environments. When we return to Stables, the Horse/Unicorn says a quote too, to naturally connect the flow of the conversation with what they were saying earlier (before going to games session). The logfiles generating procedure in games is kept intact from Kannetis work.

Figure 4.16: From Game: Exploitation Mode: Game Choice

4.4 The Choice Making — Traversing through the two Modes

As mentioned before there are various points in our application where the user can switch between the modes of Exploration(converse) and Exploitation(play games). These points play a major role in the understanding of the patterns kids follow regarding Exploration and Exploitation inclinations. In the subsection above we described how the Exploit—–Explore transition is made (from games to story/talk). In the Exploration mode there are frequent questions (either by the Horse/Unicorn in the main theme, or by the Companion in the Random Quotes section) where the child has to decide whether to stay and talk further or go to the Companion's yard with the mini-games. It is quite obvious that we have a total number of four transition types:

- Exploration — Exploration
- Exploration — Exploitation
- Exploitation — Exploitation
- Exploitation — Exploration

We shall analyze these thoroughly in Chapter 6.

4.5 Flow Diagrams

Below we present the FSMs which represent the structure of the Exploration part of our application. Note that the names given to the states are the corresponding frame tags from our application. For an idea of what they contain there is a list of the dialogs witch we used at each frame (dialogs have the same tag as the frame they are used in). Also, the Random Quotes FSM (4.20) only show the outline of its structure and does not contain the various themes' exclusive FSMs.

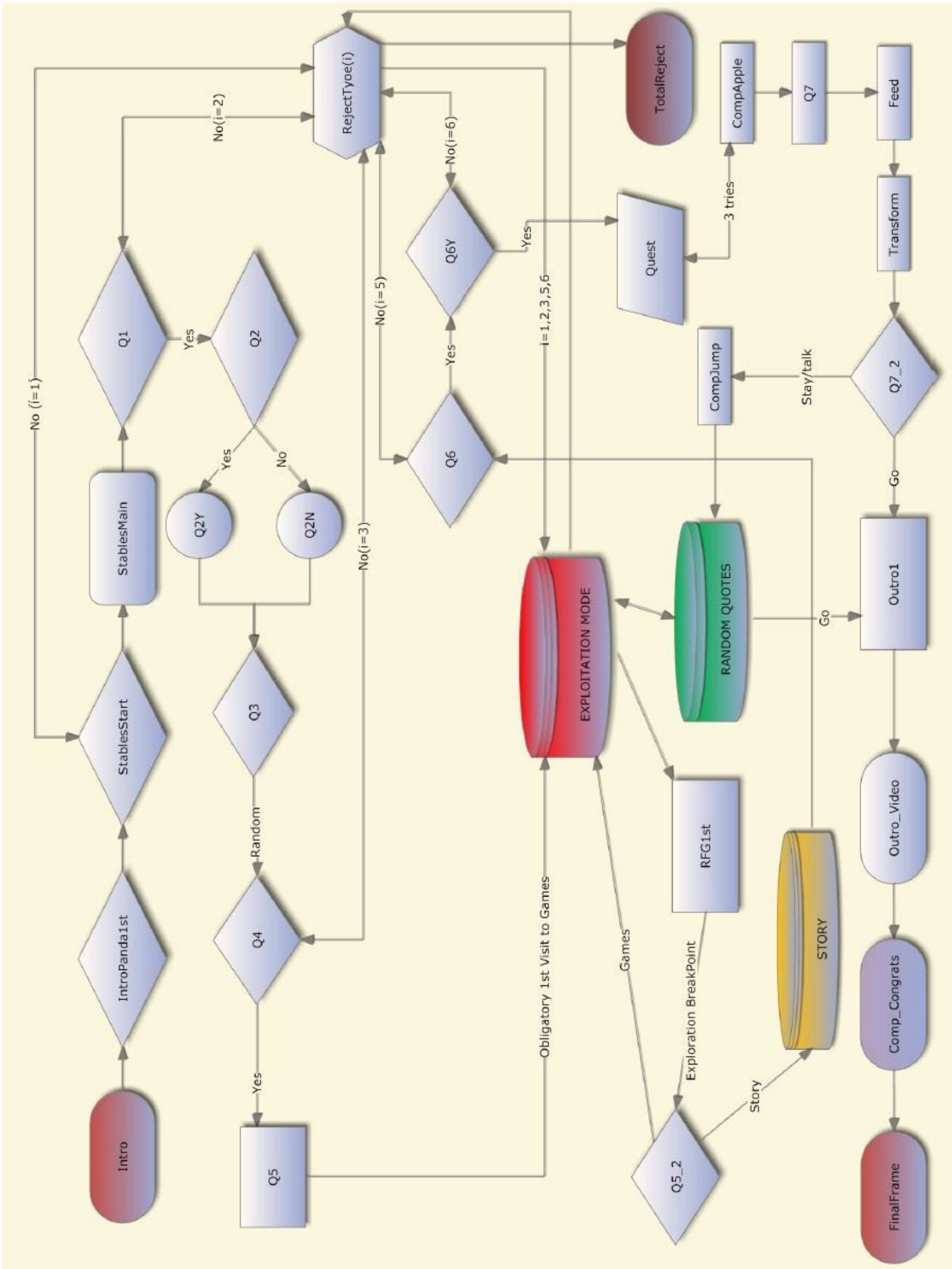


Figure 4.17: Outline FSM of the First Playthrough.

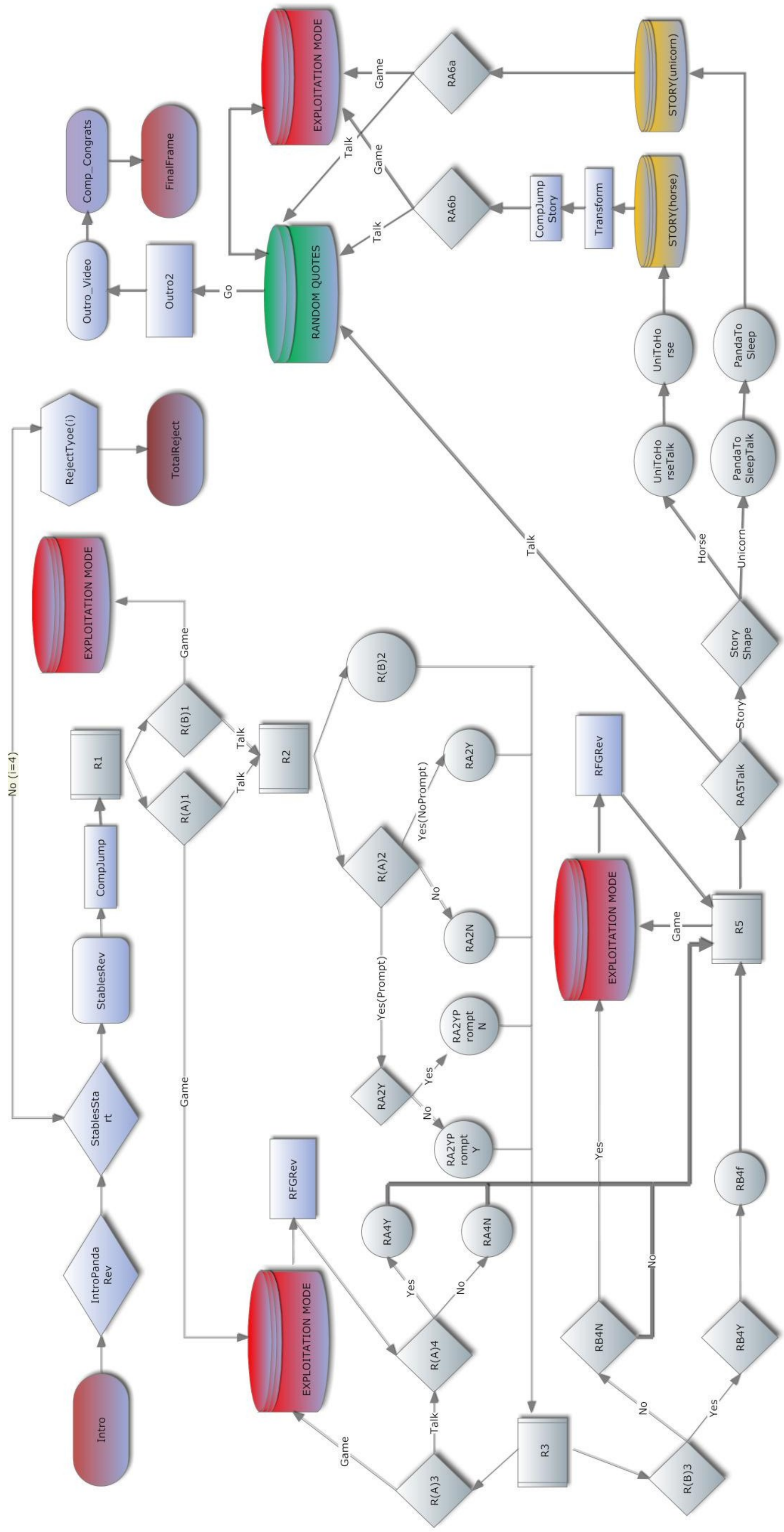


Figure 4.18: Outline FSM of any subsequent Playthroughs.

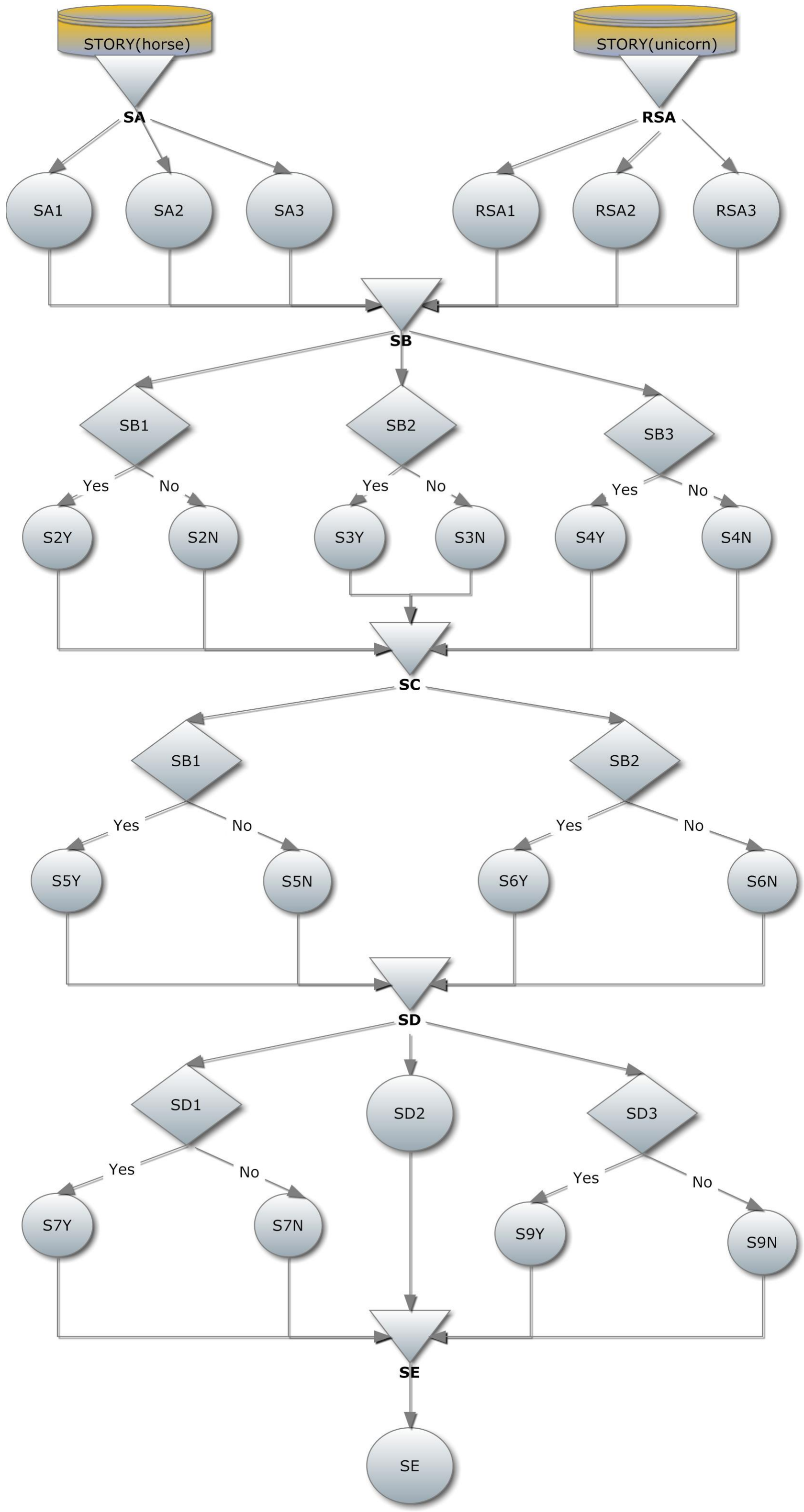


Figure 4.19: Outline FSM of the Story Section.

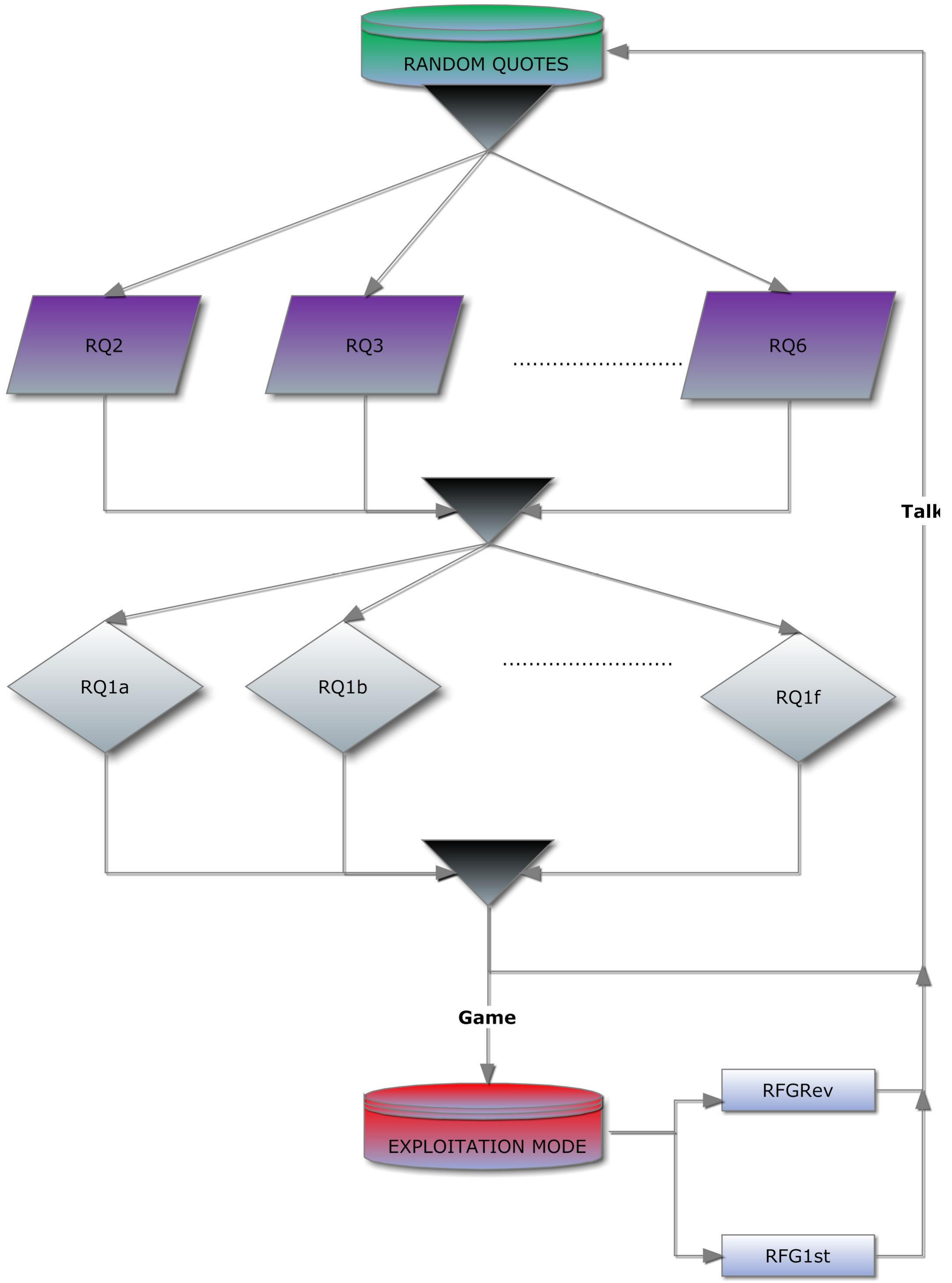


Figure 4.20: Outline FSM of the Random Quotes Section. Note that all purple boxes are separate themes with their unique FSM.

Chapter 5

Evaluation Methodology

5.1 Introduction

The evaluation of multimodal dialog systems is a complex task and different metrics (objective and/or subjective) are typically used to evaluate different aspects of such systems [28]. The goal of this work is to examine the unbiased preference preschoolers show towards Explorative (also mentioned as Exploratory) or Exploitative behaviour and recognize any patterns they follow in the transition between the two. We also want to consider how these patterns and preferences alter during the preschooler's cognitive process. For that reason (as mentioned in Chapter 2), the target group of this application are children of the preoperational stage and more specifically of the Intuitive Thought Substage. We visited two preschool environments (kindergartens) where 9 children aged 4 to 6 (three for each age) were asked to play with our game. In order to gather objective metrics, the application we developed produced a series of log files containing all the needed information and statistics. For the subjective aspect of the metrics we used a questionnaire on every child after the end of the session.

5.2 Log Files Implementation

Our application is divided into a number of frames (200 frames approximately), each one of them having one of the following contents: a dialog segment, a response or a question to the user, an animation or a mix of the above. At the end of each frame we create/attach information to our log and add the new data to any cumulative metrics we have. We store 3 general types of statistics: one for the whole application, one for the Exploration part only (Exploitation part has it's own logs from Kannetis [1] preceding work) and one for each specific frame that was played by the user in his/her session's path. Right below there is a general reference to these metrics. We will refer thoroughly to these metrics in the next Chapter and so a general aspect of what they represent is necessary.

GENERAL METRICS AND STATS

- **Total Time Elapsed:** The total time that the child played with our application. Equals to Exploration Time + Exploitation Time. *Note that there is an aberrance of +-1 second per Exploration/Exploitation mode transition. This is due to the setinterval miscounts in Flash during scene changes (insignificant if we consider that each session lasts 2000 seconds approximately)..*

- **Time spend on Exploration/Conversation:** The time that the child spent on Exploring/Conversing with the ECA's.

- **Time spend on Exploitation/Games:** This equals to Clear Games Time + Idle Games Time. Idle Games time refers to the time children spent to choose a game (Exploit—Exploit Transition), and to some small quotes the Companion says in the Exploitation Environment, for the natural connection of the two Modes and for the small tutorial in the first visit of the Exploitation Mode. This time didn't fit into any of our Modes and we calculate it separately.

- **Total number of Games Started:Completed/Abandoned:** Refers to Exploitation Mode obviously.

- **Parameters used in Game Sessions(Fantasy/Curiosity/Difficulty):** The Malone's Quality Characteristics that we used in our Exploitation Mode.

EXPLORATION MODE METRICS AND STATS

- **Total Audio-Speech Duration//Actually Listened to//Skipped:**The amount of Speech that was produced by our ECA's. Is divided into the amount that was actually listened to by the user + the amount of audio skipped (rare event which occurs when a user gives an answer and at the same time the ECA starts a Time Out prompt which is skipped).

- **Total Inactivity Time (Actual/Clear+Videos/Animations):** The time that the child remains inactive. Is the sum of Clear Inactivity time and Animation time. In the last case (animations) the user's inactivity is unavoidable and thus we measure it separately. In the first case (Clear Inactivity) we measure the time that mediates between a question from an ECA and the start of a user's response. *We remind that after 10 secs of clear inactivity a Time Out occurs where the question is repeated in a different manner (see Chapter 4). During the Time Out speech segment, inactivity time interval stops and starts again at the end of the Time Out and vice versa (for when we count the Time Outs time).*

- **Total User Voice Duration(Clear/Junk):** The duration of the child's speech. Divided further into Clear and Junk. By Clear User Voice Duration we refer to the clear answers or questions the user poses, without some ECA talking at the same time. In other words it is the user's speech duration during his/her inactivity time. Accordingly Junk time refers to the aforementioned phenomenon: the barge in of the user while an ECA still talks.

- **Total Voice Requests Detected:(Clear/Junk/SpeechCross):** The number of times the user started an utterance. Clear and Junk Requests were explained right above (for the voice duration; same logic here). SpeechCross refers to the phenomenon where the user starts to talk at Inactivity Time (Clear) and his/her utterance ends while an ECA is talking (Junk; ex. a Time Out starts) and vice versa. This will count towards a SpeechCross Request and the Voice Duration will be accordingly distributed to Clear and Junk.

- **Average Duration of Voice Activity:** This is also divided into Clear and Junk. SpeechCrosses count towards both durations.
- **Total Number of Answers Accepted:** The number of answers the user gave which were recognized and forwarded by the ASR/WoZ and accepted by our application.
- **Total Number of TimeOuts:** The number of times a Time Out occurred and help was given to the user (either a hint or a repeated question in a different way).
- **Total Duration of TimeOuts:** The duration of all the Time Outs of the session.
- **Total Delay Added:** We add some pauses/delay between some frames for natural dialog transitions. This is a part of Exploration Total Time.

FRAME BY FRAME ANALYSIS, METRICS AND STATS

- **Frame Name:** The frame's exclusive tag.
- **Total Time so far (Conversation/Games):** The time that has elapsed since the beginning of the Session and how it has been distributed so far in Exploration and Exploitation Modes.
- **Total Frame Time:** The total duration of this frame. Equals to Audio-Speech + Inactivity Time + TimeOuts Speech(Actual) + VAD Time(Clear) (+ or == Animation). Explanation: Audio-Speech is the total time of the ECA's speech in this frame. It is followed by user's the Inactivity time that mediates until he/she poses an answer(or a question less often). The duration of this answer is the VAD Time which usually ends the frame. Now as far as animations are concerned, there are three possibilities: The frame has no animations and we add nothing to the frame's total time. The frame has some animations and so their time is added to the frame's total time. The frame contains only animations and thus the frame's time equals to animation's time (all other parameters are 0).

- **Audio-Speech Duration(Actually listened to/Last Audio Time skipped):** The duration of the ECA's speech in this frame. There is a distinct possibility of the user talking with a barge in request (while an ECA talks) and it being accepted as an answer to what the ECA would ask, and thus ending the frame prematurely skipping some audio. Actually such a thing was never noticed during the experimental procedure.

- **Audio-Speech Analyzed(duration):** The turns of speech among the ECA's. For example let's assume the values in this field were: *Unicorn(3) Panda(2) Unicorn(8) Panda(1) Unicorn(16) UnicornTO(5).* This represents a conversation among the Unicorn and the Panda ECA's. Inside the parentheses is the duration of their speech. The letters TO attached next to an ECA's tag represent a Time Out audio segment.

- **Total Voice Requests Detected(Clear/Junk/SpeechCross):** The frame specific Voice Requests which were analyzed above as cumulative values for the whole Exploration Mode.

- **User Voice Duration Total (During Inactivity Time/Clear):** The frame specific Clear Voice Duration (during Inactivity Time) which were analyzed above as cumulative values for the whole Exploration Mode.

- **User Voice Duration Total (During Character Speech/Junk):** The frame specific Junk Voice Duration (barge in's) which were analyzed above as cumulative values for the whole Exploration Mode.

- **Inactivity Time:** Frame specific User Inactivity Time. We remind that this is the time between the end of a question from the part of an ECA and the beginning of an answer from the part of the User. The value of this field may be 0 if the frame was a monologue or an animation (or the frame was prematurely ended by the user in the extreme case we mentioned above).

- **Answer:** The answer that was given by the User and was forwarded by the ASR/WoZ to our application. The value of this field may be [No Question] if the frame was a monologue or an animation only.

- **TimeOuts (at):** The moment in the frame's timeline when a Time Out occurred. The value of this field may be [None] if no Time Outs occurred.

- **Mouse Clicks(Interaction/Junk):** Interaction mouse clicks are the ones which have an immediate effect (example click on the door to open it, click on the cats to talk to them etc.). Junk are all other clicks. We shall refer to the mouse spamming (junk clicks) that was observed in 4 years old kids in the next chapter.

- **Additional Info:** Just some information which usually sums up the content of the frame.

5.3 Experimental Setup

As we mentioned above, we visited two kindergartens were a total of 9 native Greek preschoolers aged 4-6 (three for each age) tested our application (all the recorded voices are in Greek). We used a set of headphones with an attached microphone as the main input/output method. Before each session, we thought necessary to dedicate some time to get to know the child and make him/her feel comfortable. It is natural for kids at this age to be hesitant towards strangers. It is crucial then to gain their trust in order to eradicate the possible feelings of apprehension, lack of comfort or awkwardness.

Each child was asked to play with our game two times but at a different day (maybe even a different week) for the reasons we explained in Chapter 4, Section 2.3: *Exploration Mode structure.* Each session lasted approximately 25-35 minutes. The kids were informed from the beginning of the session that they could play freely as long as they wanted and stop whenever they felt like it. All they had to do is say "I'm bored!" and the session would end with the appropriate break functions. Actually after 35-40 minutes of playing time we stopped the session ourselves (no need to overload the child). At the end of each session the child was asked to answer a series of questions in order for us to fill the questionnaire. We shall analyze this in the next Chapter's Objective Results. Note that a supervising student was always present with the kid for the whole session time. Many would disagree with the presence of an adult in the room as far as subjective metrics are concerned, but it would be rather irresponsible from our part to leave a child of such age unsupervised. On the other hand, close observation of the child's playing, gave us a better subjective view of things.

We should also mention that there were some crashes of the WoZ module when a child spoke in very big utterances. For that reason and having in mind the average length of a session (30 minutes), we implemented a 'continue' system in our application. See figure 5.1 for an idea of how the Main Menu of our game looks like.

Figure 5.1: From Game: The Main Menu. From here we can choose the type of the session (1st time or revisit) and shape of the main ECA (horse/unicorn), Enable/Disable Thought Clouds, Enable/Disable the Random Quotes we want, continue the session from the frame we like (we enter the frame's number corresponding to the tag from the table to the right) or enter debugging mode.

Chapter 6

Evaluation Data and Metrics Analysis

6.1 Introduction

In this chapter we set forth the evaluation results that occurred from the processing of the Logfiles that we mentioned in the previous chapter (Objective Metrics) and the questionnaire that we filled during the exit interview that took place after each session (Subjective Evaluation Results). In the following sections we shall analyze these metrics and identify any patterns that dominate each age as far as Exploration and Exploitation preferences are concerned. We shall also present some various statistics that are not directly related to the EvE dilemma and the mode switching patterns, but hold a fair amount of interest in examining Children-Computer Interaction. Finally we will comment each result based on personal observations that were made during the sessions that took place.

6.2 Objective Metrics and Statistics

As mentioned in Chapter 4, the Exploration part of our application was discriminated and designed in two parts: the first visit (first playthrough) and revisit (any subsequent playthroughs) modes. We remind that the first playthrough of our application has a mostly predefined structure so that the user gets acquainted with the modality usage and is presented with the modes of Exploration and Exploitation. In subsequent playthroughs there is a much more loose and random structure where the user has much more options. Due to this fact and also in a way to examine behavioural changes among the playthroughs, the statistics that are correlated with the EvE dilemma were separated into two categories (1st Visit and Revisit) when that mattered, while others were treated globally (both playthroughs considered). For a detailed description of what each variable measures, please refer to Chapter 5.

6.2.1 Time Distribution for the First Playthrough

First of all we present how much time was spent in each mode by each age. For this statistic the aforementioned discrimination was made. The figure 6.1 shows the first playthrough results.

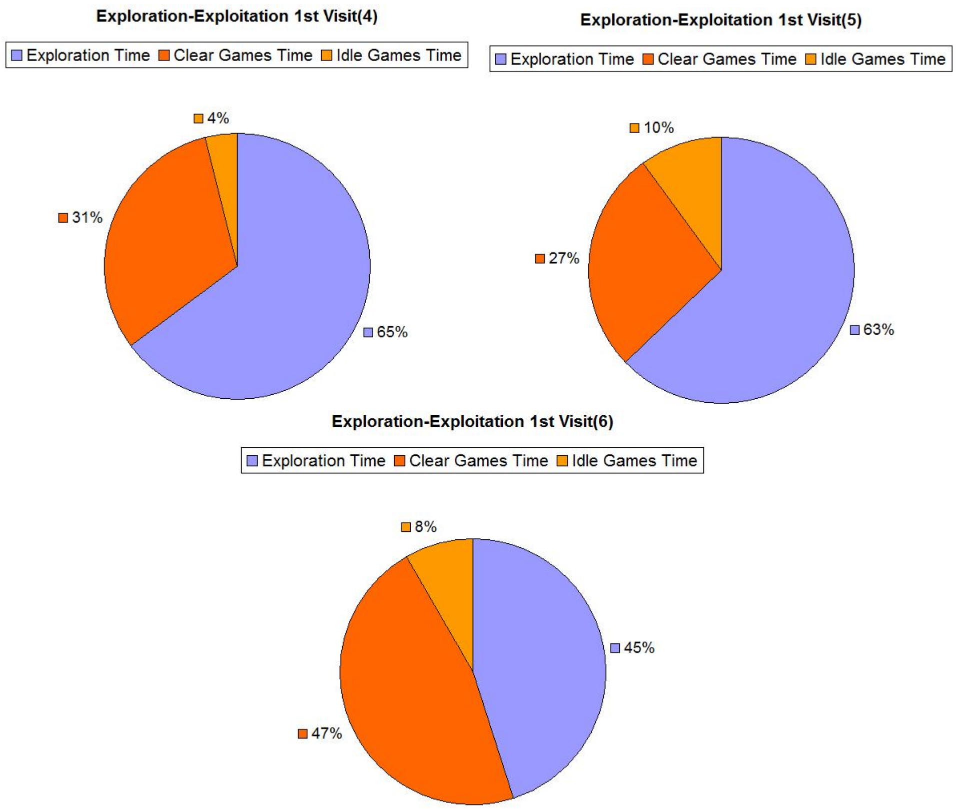
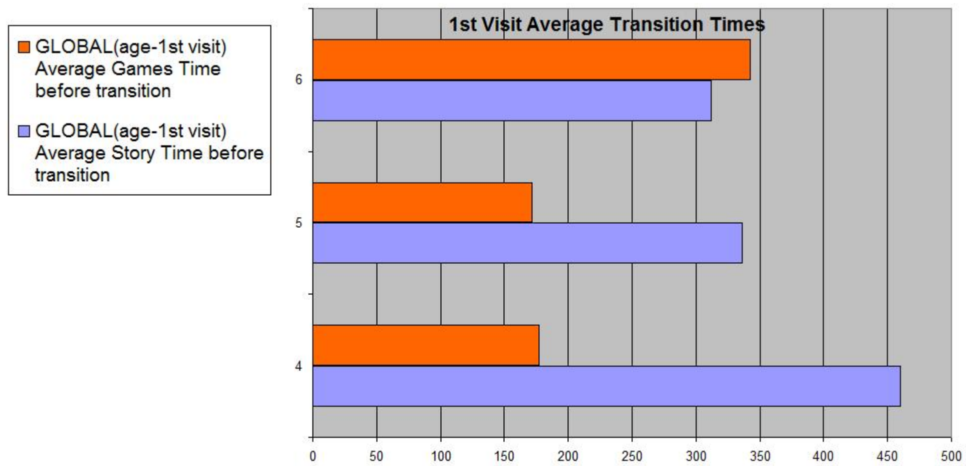


Figure 6.1: *Distribution of Exploration and Exploitation Times during the First Playthrough (age in parentheses)*

As we can notice, four and five years old children distributed their time in a similar manner. Such behaviour was expected for the first playthrough due to the application's structure we mentioned. There is a slight emphasis shown to the *Exploration* part. Six years old on the other hand showed a preference towards the Exploitation part during the first playthrough. The very small Idle Games time we notice in four years old (compared to the older ones) is due to their *'Trigger Happy'* behaviour with the use of mouse. They didn't spent much time thinking which game to choose and just clicked instinctively. This behaviour was diminished in their second playthrough. We shall refer to this again in the related subsection.

6.2.2 Average Transition Times for the First Playthrough

In figure 6.2 we can see the average amount of time each age dedicated in one mode before making a transition to the other. It is rather interesting that five



	4	5	6
GLOBAL(age-1st visit) Average Games Time before transition	176.83	171.33	342.67
GLOBAL(age-1st visit) Average Story Time before transition	460.33	336.25	312.22

Figure 6.2: *Average Time spent on each mode before switching to the other. First Playthrough Results (seconds)*

years old children appear to have the mean behavioural pattern in transition times. They have very similar values in average Exploration time with six years olds and almost exactly the same average Exploitation time with four years old ones. Four years old kids had the biggest average Exploration time while six years old the biggest average Exploitation time. Six years old dedicated a fair amount of time in Exploitation mode from their first contact with the games while the others did not. Considering now the Time Distribution per age we mentioned above (four and five years old are similar) we need to explain the deviation of the four years old average Exploration Time since the structure of the first playthroughs' Exploration mode is mostly predefined. In one small part this is due to their adaptation curve with speech as modality usage. They had the highest Inactivity Times in Exploration mode. Again, this will be analyzed in a separate subsection. In another part, the way they distributed these times through the session is responsible.

6.2.3 Timelines for the First Playthough

In order to gain a better perspective of the time distribution, we set forth the figures 6.3 (age 4), 6.4 (age 5) and 6.5 (age 6) which visualize how each child distributed it's time in the first playthrough (3 timelines for each age).

Note that the X-axis values (time in seconds) are centered for each session in a category (age) but not centered across categories, meaning that each age has different values in the X-axis and that the visualization (size of bar) is not always representative of the actual time when making cross-categories comparisons. Blue color represents Exploration/Story Time, light orange Idle Games/Exploitation Time and dark orange Clear Games/Exploitation Time. Numbers in orange areas show the number of games played. The 'Q' represents the Quest game in the first playthrough which progresses the story.

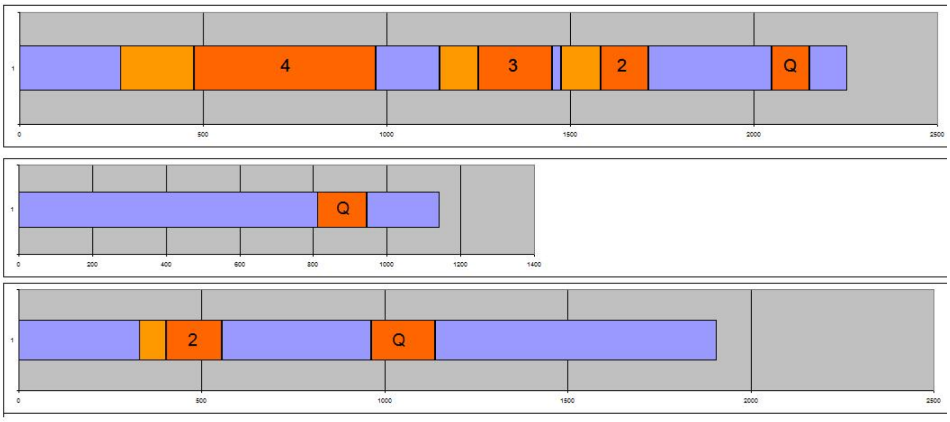


Figure 6.3: *The Timeline of First Playthrough for 4 year olds (seconds)*

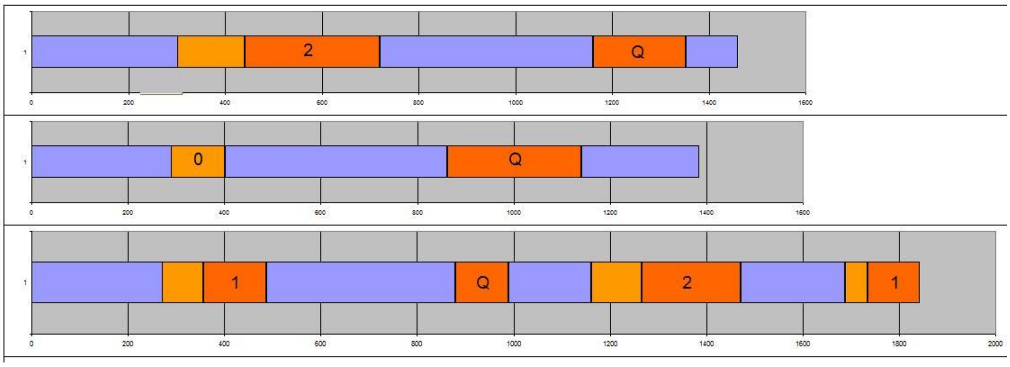


Figure 6.4: *The Timeline of First Playthrough for 5 year olds (seconds)*

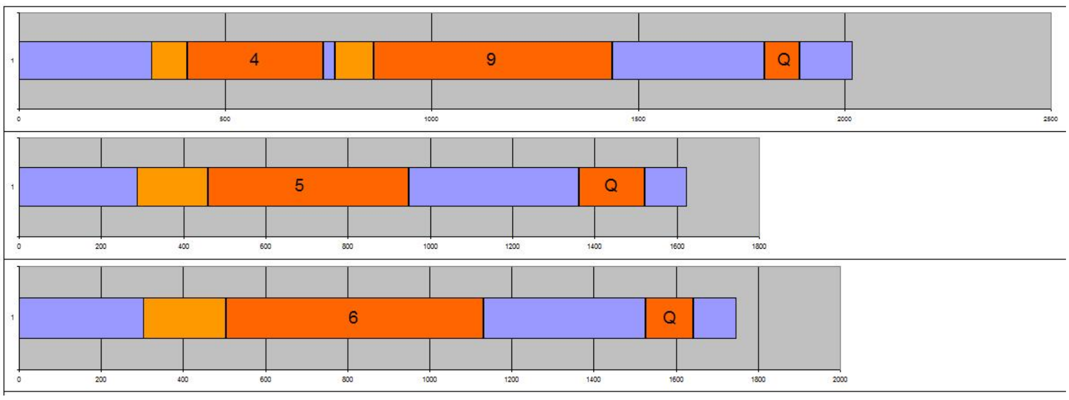


Figure 6.5: *The Timeline of First Playthrough for 6 year olds (seconds)*

There is a rather obvious and consistent pattern here that dominates the age of six. Four and five years old children either played a couple games once in a while or even completely ignored the Exploitation mode in few cases. There was only one exception out of six for this behaviour (the first timeline of figure 6.3). Six years old kids on the other hand showed a 'heavy' Exploitation performance from the very start. They dedicated much more time in games than the other ages and with minimal transitions, thus justifying their bigger average Exploitation time we mentioned. They played all the games available at least one time. If they enjoyed one game they would sometimes play it again but not consecutively (interchanging between non played games).

6.2.4 Time Distribution for subsequent Playthroughs

In the second playthrough there were some major changes in the behaviour of the children aged four and six. Five years old kids sustained their behaviour with short transition times and a preference towards the Exploration part. Compared with the first visit, there was only a very slight deviation (3%) in their second playthrough which certifies this preference, considering the structure of our application's subsequent playthrough (more options, randomness, generic structure). In figure 6.6 we can see the Time Distribution in the two modes during our revisit.

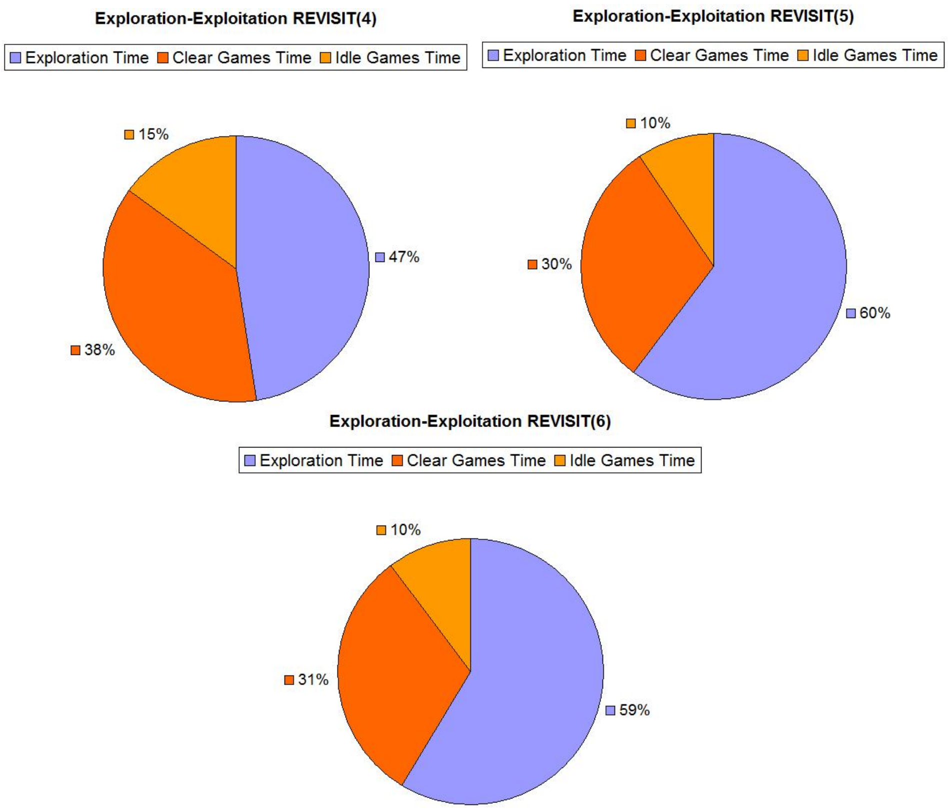
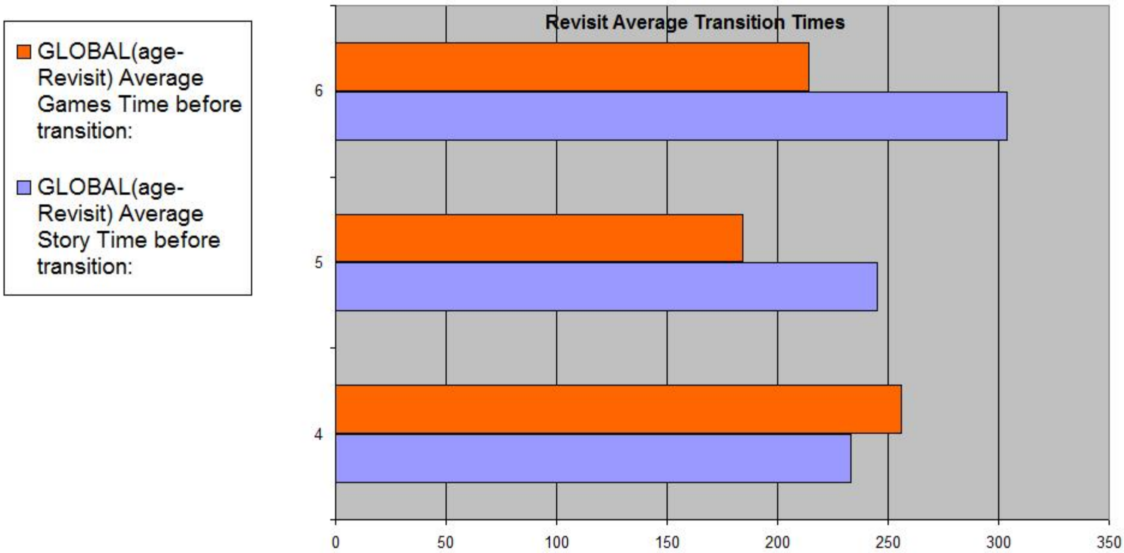


Figure 6.6: *Distribution of Exploration and Exploitation Times during the subsequent Playthrough (age in parentheses)*

Four year olds dramatically reduced their Exploration time and focused more on playing games. When they got a firm grasp of them and understood how they could use their knowledge to advance in the games, they preferred to remain in the 'safety' of what they knew well instead of exploring many of the conversational possibilities. The mouse 'trigger happy' effect we noticed in the first playthrough is now minimized because they now tend to think in advance and wonder which game to play, thus increasing the Idle Games/Exploitation Time to normal levels (5% more compared to other ages). Six year olds inverted their behaviour too. Their time distribution now highly resembles that of the five years old (only 1% difference), which means an inclination towards Exploration. The patterns they followed though are different and we shall see them in the two following subsections.

6.2.5 Average Transition Times for subsequent Playthroughs

In figure 6.7 we present the Average Transition Times for the subsequent playthrough. It is interesting to note how the average Exploration time increases here in proportion with age, while the case was exactly the opposite in the first playthrough.



	4	5	6
GLOBAL(age-Revisit) Average Games Time before transition:	256	184	214
GLOBAL(age-Revisit) Average Story Time before transition:	233	245	304

Figure 6.7: *Average Time spent on each mode before switching to the other. Subsequent Playthrough Results (seconds)*

Four year olds almost halved their stay in Exploration mode before going to Exploitation, where they increased their average playing time which was the highest of all ages. Five years old children had an average stay time in games very similar to the first playthrough results. However, they decreased their average Exploration time before a toggle to Exploitation mode, which means they made more frequent transitions, as a break from the interaction with our ECAs. Six year olds had the highest average Exploration time before a transition to the games, but they greatly reduced their average Exploitation time. In figures 6.8 and 6.9 we present each mode's Average Transition Time for both playthroughs in a stacked representation for a better visualization of the behavioural change. Note that in the second playthrough all transition times as a sum in every age where reduced, which translates as more frequent mode toggling.

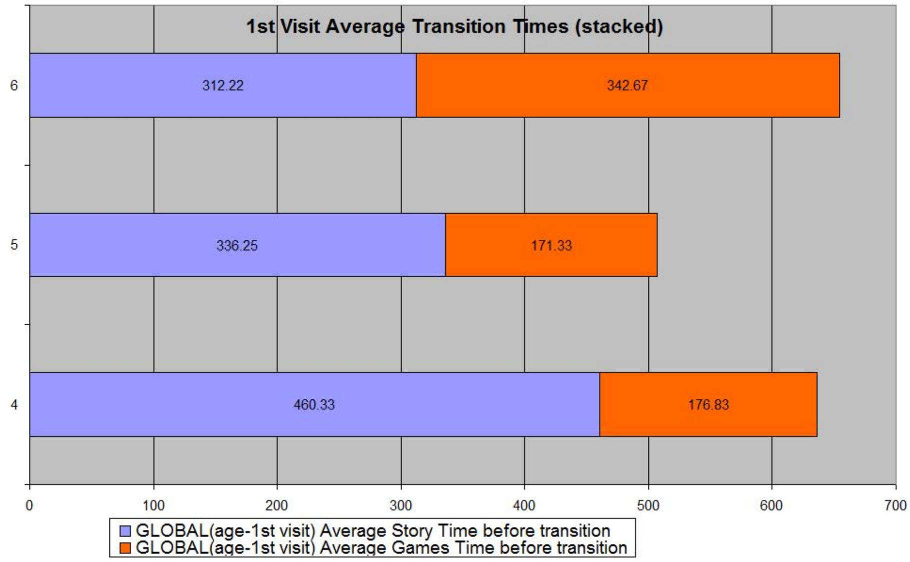


Figure 6.8: *Average Transition Times (stacked) for first playthrough*

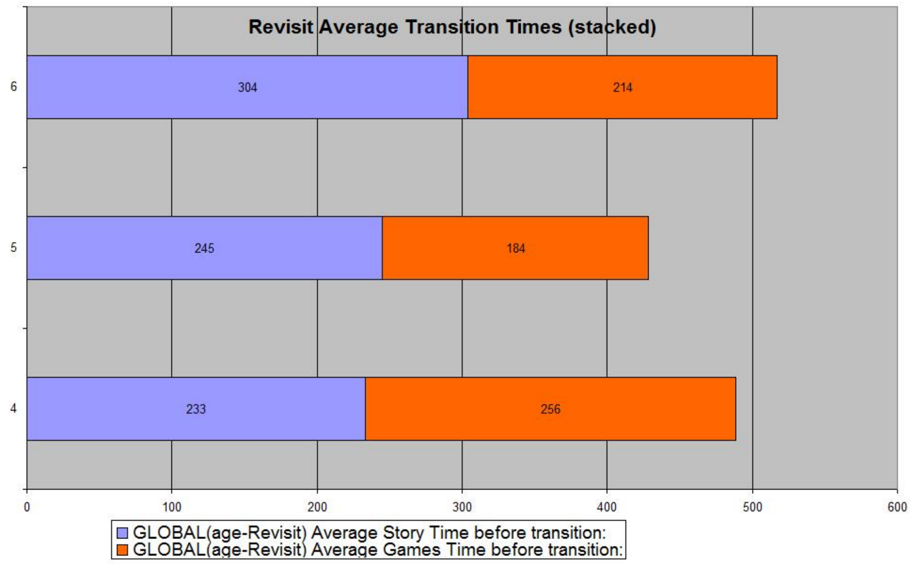


Figure 6.9: *Average Transition Times (stacked) for second playthrough*

6.2.6 Timelines of subsequent Playthoughs

In the figures 6.10, 6.11 and 6.12 we can see each session's Timeline for the second playthrough. The same rules regarding the X-axis we mentioned in the previous timelines apply here as well.

Six year olds have once again the more obvious pattern. They play almost every game available in their first chance, and when they feel there is nothing new for them in the Exploitation mode they focus on Story/Exploration with minor Exploitation breaks (2 games maximum).

As we mentioned, four year olds showed an increased preference in Exploitation in the second playthrough. The frequent mode toggling is present here too but in an inverted way compared to six year olds. They had the smallest average in Exploration time and the biggest average in Exploitation time. It was observed during the experimental procedure that when they started completing the games in a reasonable amount of time, they felt more confident and wanted to repeat their success. Conversation with the ECA's is now limited to a break from gaming sessions. Whenever they felt overwhelmed conversing with the ECAs, they would toggle to Exploitation mode in their first chance.

For the five years old children, it is worthy of note the fact that they once again almost completely ignored the Exploitation part at the beginning. For the first 1000 seconds approximately (15-17 minutes) they played either one or none games at all, while other ages had an average of 3,3 (age 4) and 6,3 (age 6) games in that time. Afterwards there is once again the frequent mode toggling phenomenon with a heavier focus on Exploration. In fact they had the smallest sum of the two Average Transition Times. Games where also used here as a break from conversation as with the six year olds, but with the difference that they never exploited the full potential of the games as six years old children did in their first Exploitation chance. There was only an increase in Exploitation time by the end of the session (except the first timeline).

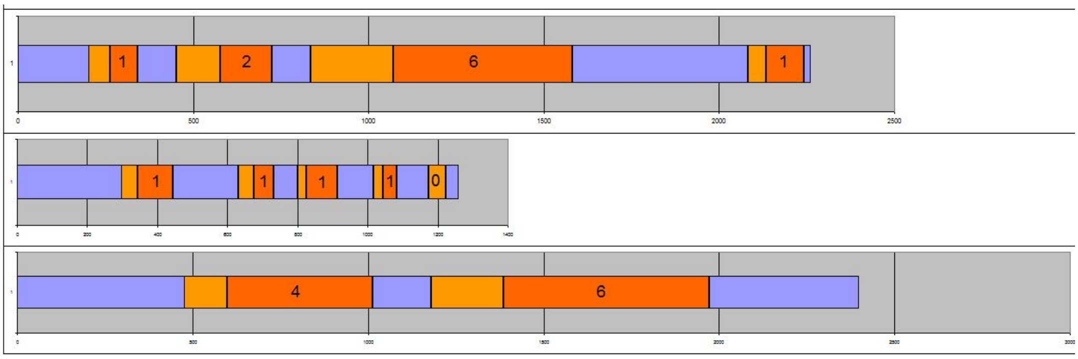


Figure 6.10: *The Timeline of subsequent Playthrough for 4 year olds (seconds)*

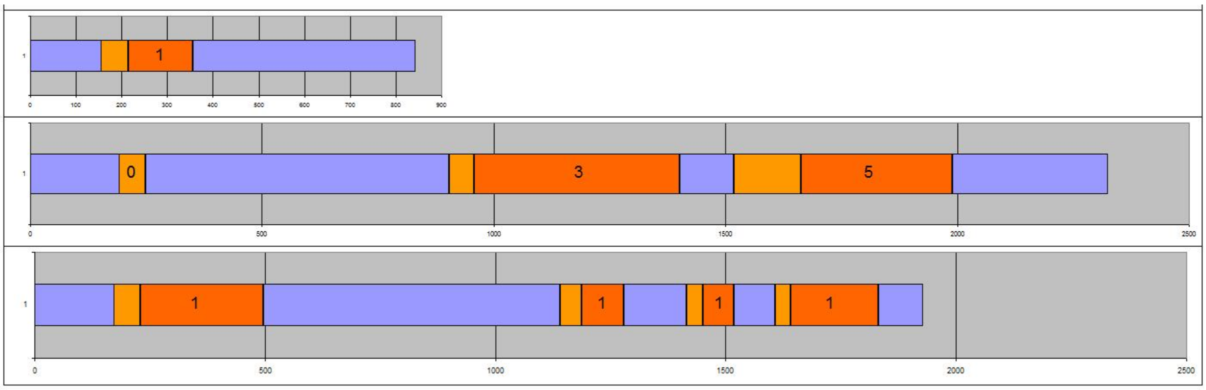


Figure 6.11: *The Timeline of subsequent Playthrough for 5 year olds (seconds)*

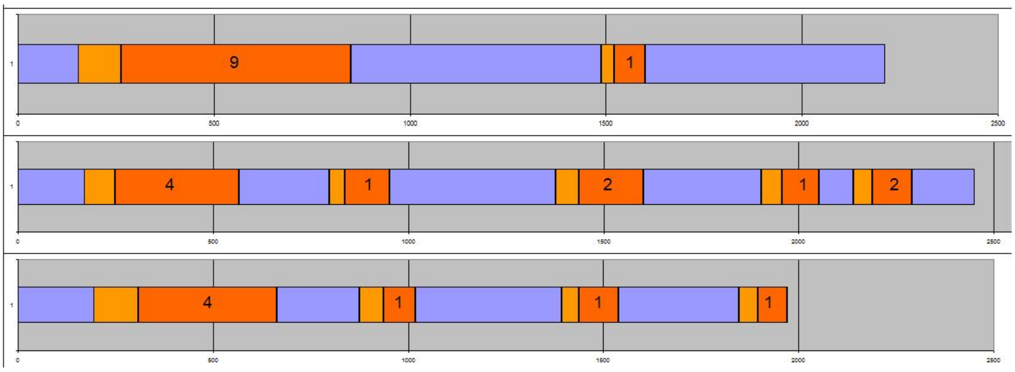
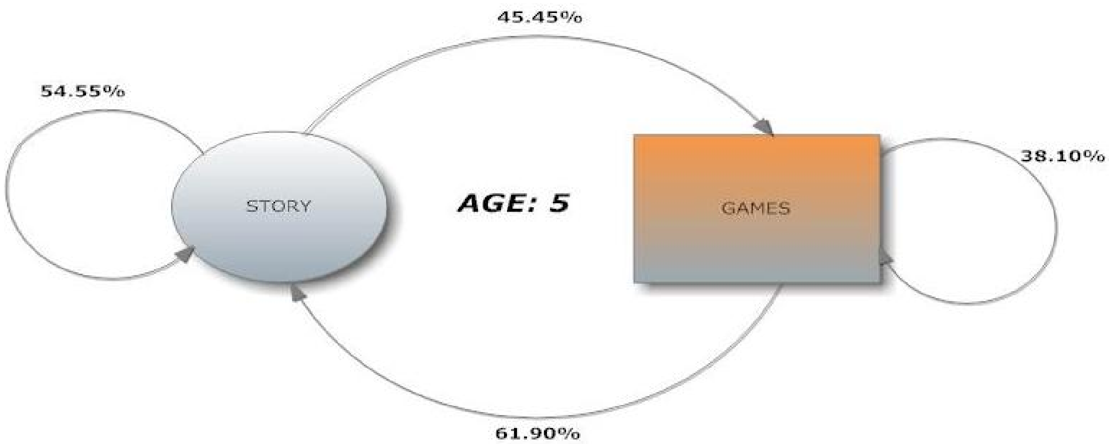


Figure 6.12: *The Timeline of subsequent Playthrough for 6 year olds (seconds)*

6.2.7 Decision making and mode choice probabilities

Taking the previous metrics into consideration, it is also interesting to see the decision making behaviour each age followed, regarding the mode switching. The bigrams presented in figures 6.15 (age 4), 6.13 (age 5) and 6.14 (age 6) represent the mode choice probability for each age and the percentage of each type of transition out of the total number of transitions. These bigrams refer to both playthroughs. We did not make a separate analysis here for each playthrough because we thought useful to interpret these global metrics by considering the individual playthrough conclusions (total times, average mode times and timelines) of the previous subsections. We must note here once more the fact that Exploration choice making frames where distributed in a way so that there was an average time of two minutes (approximately) between them. Each game lasts one to two minutes depending on age, difficulty and familiarity with the specific game. Taking into consideration the Idle Exploitation time this could be more.

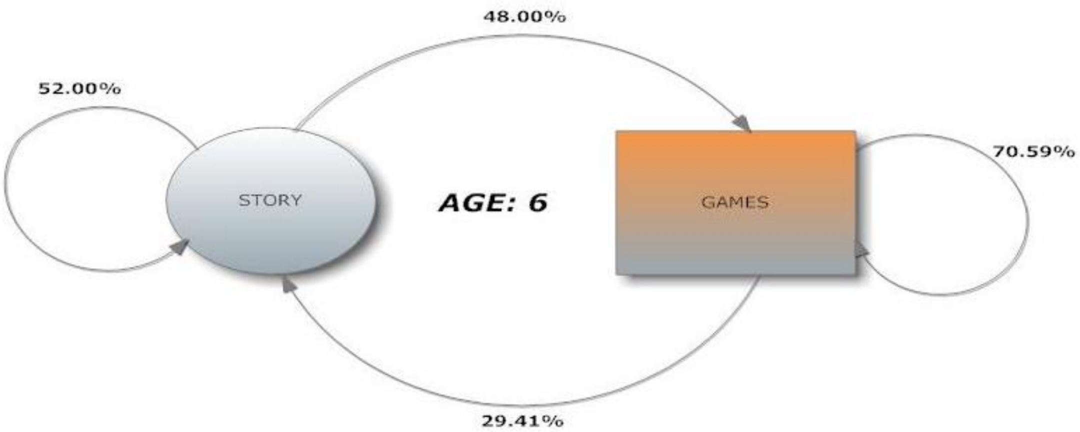


Type of Transition and Total percentage

Story-Story	27.91%
Story-Games	23.26%
Games-Games	18.60%
Games-Story	30.23%

Figure 6.13: *Bigrams of choice probability for all playthroughs (age 5)*

All ages had almost the same percentage in Story/Exploration mode decisions which were very close to a 'coinflip' 50-50 distribution. The pattern of frequent mode toggling was present in all ages but was used at different times and for different reasons. Five year olds (figure 6.13) had the lowest percentage (38.1%) in Games-Games transition witch reinforces our conclusion of their Exploitation mode usage which they tended to ignore, at least in the beginning of the session. Their higher Exploration Time percentage along with the lowest cumulative average transition time suggest that they followed the mode toggling pattern more intensely and used games as a break from conversation with the ECA's.

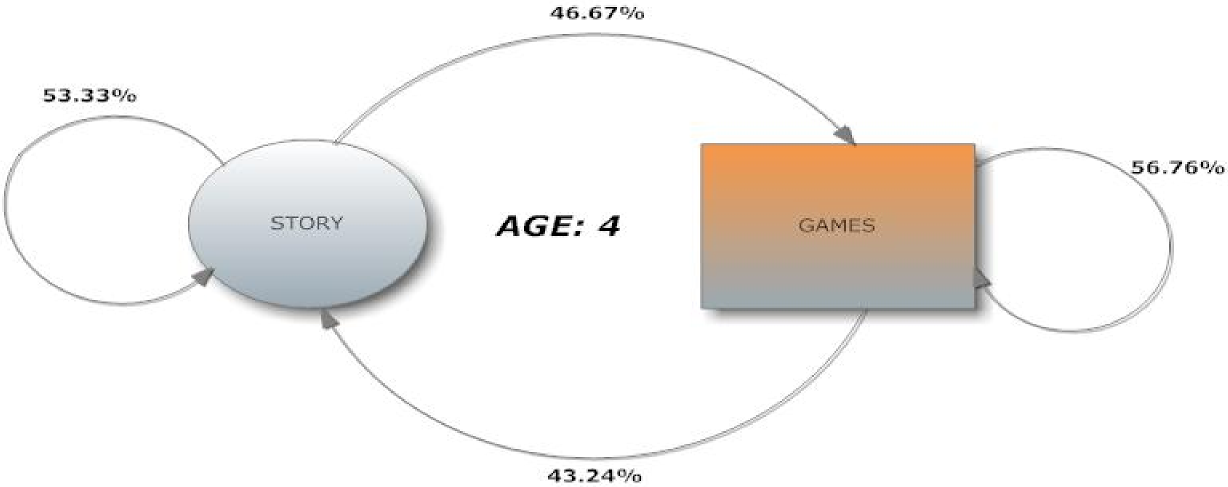


Type of Transition and Total percentage

Story-Story	17.11%
Story-Games	15.79%
Games-Games	47.37%
Games-Story	19.74%

Figure 6.14: *Bigrams of choice probability for all playthroughs (age 6)*

Six year olds (figure 6.14) had an almost doubled percentage for that type of transition (70.59%). We saw in their timelines (figures 6.5 and 6.12) that once they were in Exploitation mode they played games in a consecutive manner. In their second playthrough they repeated this successive game playing in the beginning and when they realized there was no potential or challenge for them they focused solely on Exploration with only minor Exploitation breaks.



Type of Transition and Total percentage

Story-Story	23.88%
Story-Games	20.90%
Games-Games	31.34%
Games-Story	23.88%

Figure 6.15: *Bigrams of choice probability for all playthroughs (age 4)*

Four year olds (figure 6.15) had the most balanced distribution in their choice making. All types of transitions were almost equally probable (close to 50% for each decision). There is only a very slight inclination towards Games-Games transitions. We mentioned that in the first playthrough there was a preference towards Exploration which was counterbalanced by the Exploitation preference in the second one. This inverted behaviour was noticed when the games in Exploitation mode became more comprehensible for them while story-telling and conversing became overwhelming. Apart from that, the equiprobability of their decisions could be related to their limited Attention Span compared to other ages.

6.2.8 Miscellaneous Metrics and Statistics

The following subsections contain some miscellaneous information about the children's interaction with our application and especially with the Exploration part which we developed. A thorough analysis of the Exploitation mode interaction has been made by Kannetis in his Master Thesis [1]. All of them are Exploration mode metrics except for subsections 6.2.8.5 (Total Sessions Time) and 6.2.8.6 (Game Statistics).

6.2.8.1 Inactivity Times

As we mentioned in Chapter 5, the term Actual Inactivity Time refers to the time that mediates between the end of an ECA's question and the start of the user's response. Animations Inactivity Time is not standard but determined by the random path that was followed in Exploration. We are only concerned in Actual Inactivity Time. In figure 6.16 we present the average of these times for each age, considering both playthroughs.

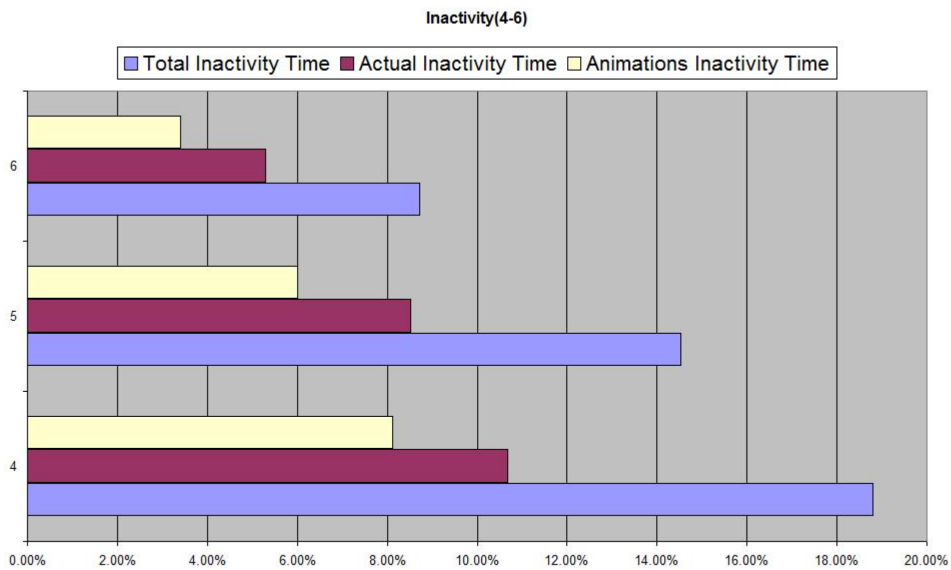


Figure 6.16: *Inactivity Times (Exploration Mode) for both playthroughs*

The results are predictable enough: Actual Inactivity Time decreases with age. Four year olds had the longest times before a response. It was observed that their adaptation to the concept of conversing with an ECA took longer compared to other ages. In the beginning this concept was somewhat incomprehensible for them but that did not discourage them from Exploration Mode in the first playthrough (quite the contrary, they were immersed and motivated towards conversation at start, even though not so effective). Ages five and especially six adapted to the conversation concept much faster and started giving immediate responses almost from the beginning of the session.

6.2.8.2 Voice Activity Detections and Durations

In figure 6.17 we set forth the number of Voice Activity Detections for each age. We should clarify once more that by the term 'Clear' (for either VADetection or VA Duration) we refer to the Voice Activity which occurred during the Clear Inactivity Time, when a verbal action is expected from the user. 'Junk' refers to Voice Activity during ECA audio playback where the user is not expected to talk. As 'Barge In' (also mentioned as 'SpeechCross') we named the rare phenomenon where voice activity starts at 'Junk' status (when an ECA talks) and ends during 'Clear' Inactivity Time (ECA has stopped talking) or vice versa. These times are distributed accordingly then to each category (Clear-Junk).

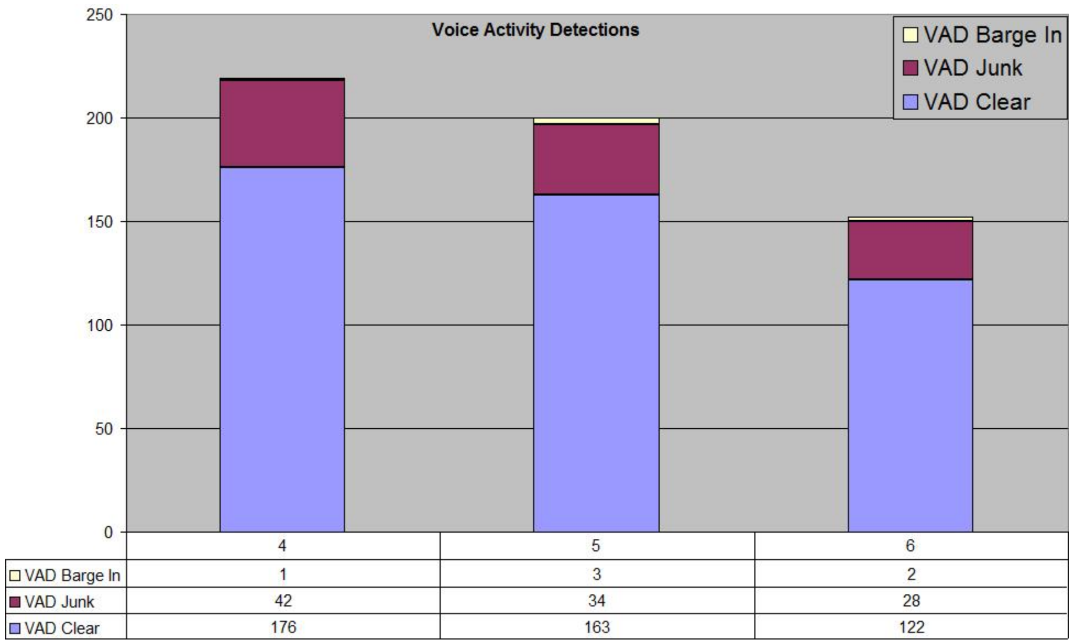
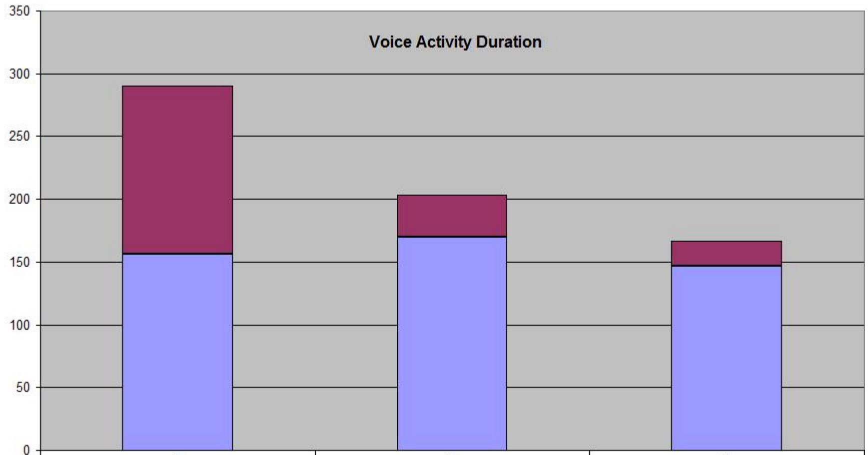


Figure 6.17: *Total number of Voice Activity Detections (Exploration Mode)*

As we can see VA Detections for both types decrease with age. Statistics such as this though (that are not mean values), can be valid only if all ages had the same Exploration Time. But that was the case exactly in our sessions: Oddly enough, no matter the different patterns each age followed and the preferences they showed towards one mode or the other, they all ended up having almost the same Exploration Times. Age four has a sum of 104 Exploration minutes, age five a sum of 102 and age six a sum of 105. We should mention though that one of the five year old kids had a very brief session (said the application terminating word 'Bored' in 13 minutes of playing time—first timeline in figure 6.11) in the second playthrough and if she was to follow the pattern of her peers, their Exploration Time would be higher. In the next figure (6.18) we can see the Duration of Voice Activity as a total for each age.



	4	5	6
VA duration Junk	134	33	20
VA duration Clear	156	170	147

Figure 6.18: *Total Duration of Voice Activity (Exploration Mode)*

Five year olds had the longest Clear VA Duration. When prompted to speak in Clear (Inactivity) Time, in many cases they made several small and different utterances to respond, which results in increased VA Detections and Duration. Six year olds had the shortest Junk VA Duration. They tended to pay more attention to the ECA's speech content and would rarely talk during that time. They responded with fewer but somewhat larger utterances which resulted in decreased VA Detections. Four year olds presented the longest by far Junk VA Duration. Their interaction with the ECA was not always smooth. They were often over-energetic in talk and used many and big utterances during the ECA's audio segment. The aforementioned statistics can be combined to a mean value which is presented in figure 6.19.

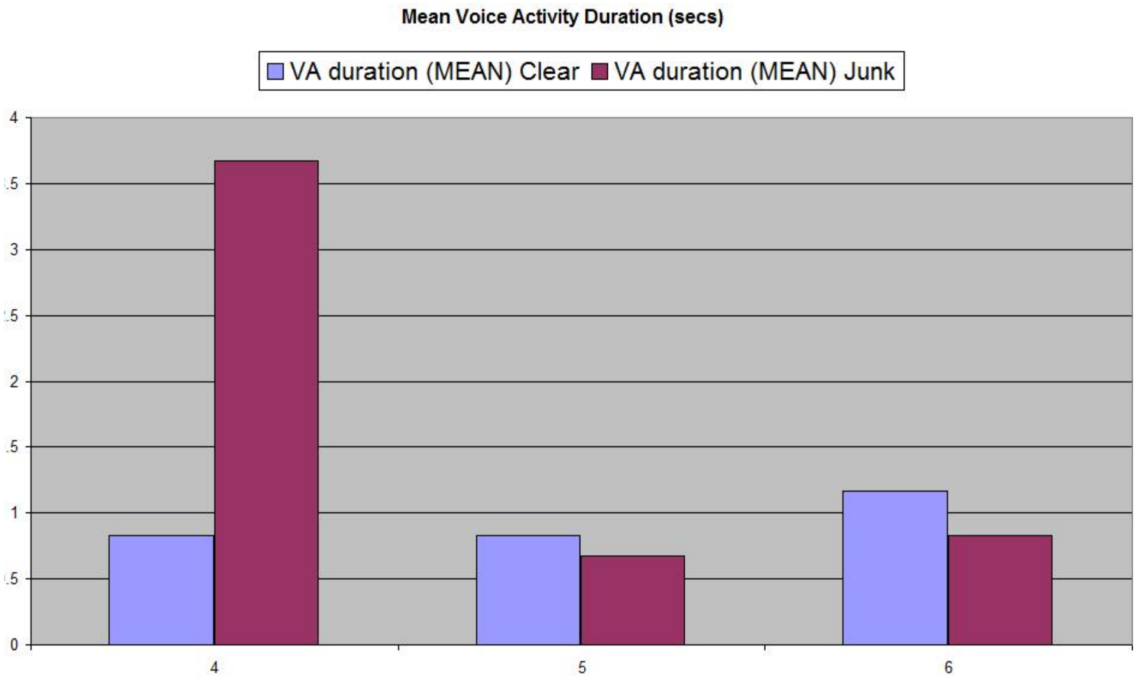


Figure 6.19: *Mean Duration of Voice Activity (Exploration Mode)*

It is obvious that four year olds could not show conservation in speech usage as other ages did. Even though highly immersed, it is contradictory how easily they got distracted by the ECA's words, and how they were very eager to engage in conversation with many random and unpredictable questions (a well known characteristic for children of such small age) even while the ECA was talking to them. That resulted in abnormally high mean 'Junk' VA Duration and in missing many of the ECA's speech content thus not knowing how to respond in many cases and triggering a Timeout (next subsection, see figure 6.20). Five and six years old children were much more focused when an ECA was talking, grasped the theme of turned based talking (conversation that is) more quickly and gave more graphic responses —especially the six year olds which where very selective in their utterances and had the highest mean Clear VA Duration.

6.2.8.3 Timeouts

In this subsection we present the number of Timeouts that occurred during Exploration time for each age (figure 6.20). Since all ages happened to have similar Exploration Times (the oddity we referred to in the previous subsection), an average Timeout occurrence for each age was not necessary and a Timeout number as a total is representative enough.

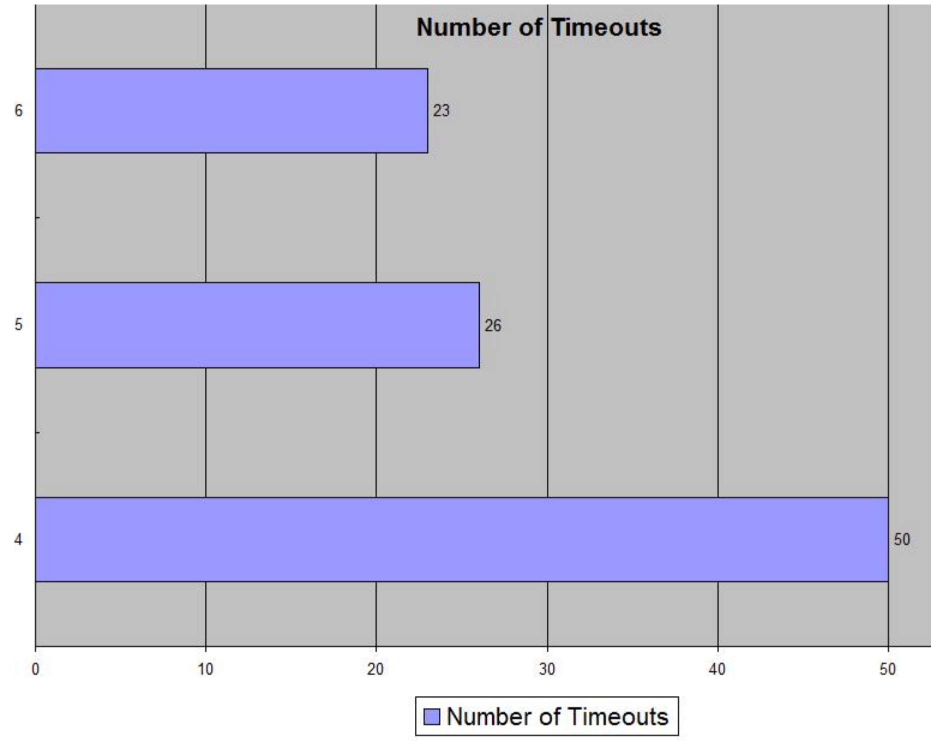


Figure 6.20: *Total number of Timeouts (Exploration Mode)*

Five and six years old children had almost the same amount of Timeouts which occurred rather rarely. It is the four year olds that have an almost double number of Timeouts for the reasons we explained in the previous subsection when analyzing their Voice Activity. Their limited (compared to other ages) Attention Span combined with increased impulsiveness and rash behaviour leaded in distraction from the ECA's speech content and eventually in a Timeout because they did not comprehend what they were supposed to do/say. This behavior was even more intense during some rare long ECA monologues (1 minute of speech).

6.2.8.4 Audio Duration Listened to

The following statistic (figure 6.21) may seem rather useless but if we consider the same total Exploration duration for all ages, it can be interpreted in a different manner.

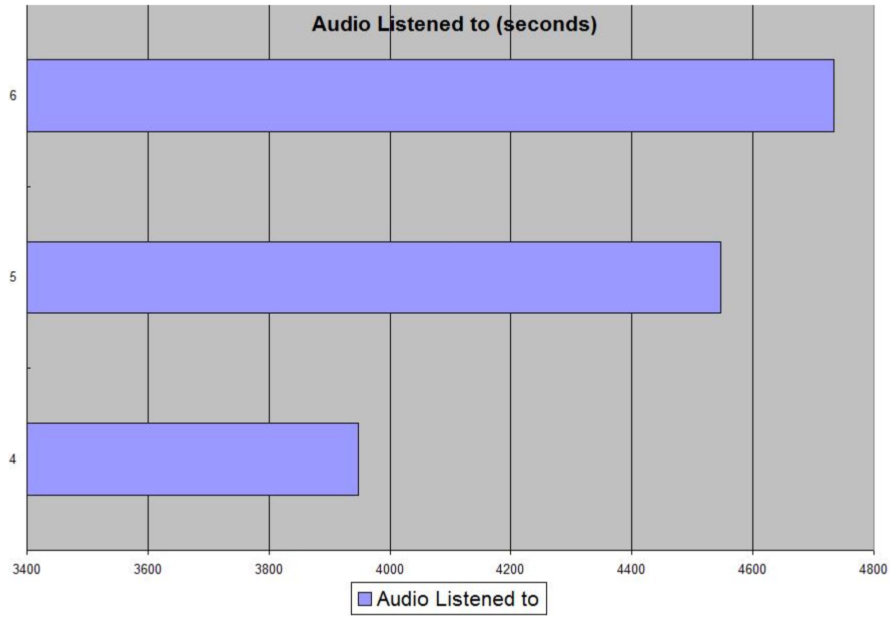


Figure 6.21: *Total pre-recorded audio time listened to (Exploration Mode—seconds)*

It is one fact that our application has a random and generic structure (*especially for subsequent playthroughs— figures 4.17, 4.18, 4.19 and 4.20 from Chapter 4 represent that structure*) and so is the audio content and duration that could have been heard. Considering though the total time all sessions lasted for each age (presented in figure 6.22 in the next subsection) and our application's length of speech segments, such deviations between ages could not be attributed to the randomness of audio playback. The case here is each age's adaptation curve and progression pace in Exploration mode. Larger audio duration listened to can be perceived as faster pace and progression in Story/Exploration where more dialog sequences will be triggered. Actual Timeout Durations and Inactivity Times count negatively towards this progression. Four year olds had the smallest Audio Duration Listened to, the highest Inactivity Times and the largest (double) number of Timeouts (figure 6.20). All this sums up to the lowest Story/Exploration percentage covered and played through and thus the slowest pace. Six year olds were the most effective in Story progression followed by five year olds with an only slight difference in effectiveness which is accounted to their higher by 5.81% Inactivity Times (both animation and clear—figure 6.16). Timeout differences between 5 and 6 year olds are minimal (only 3 Timeouts more for children in the age of 5) and so is VA Duration (23 more seconds of Voice Activity for 5 year olds). To see how Exploration Time is splitted into other times refer to chapter 5.

6.2.8.5 Total Sessions Time

In the next figure (6.22) we can see the Total Time (Exploration + Exploitation) each age dedicated to our application in all of our visits to the kindergartens.

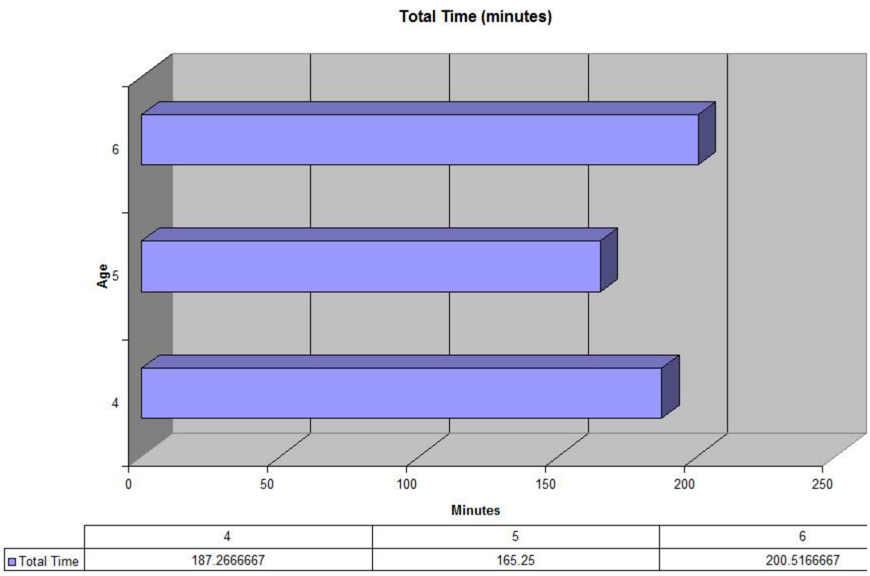


Figure 6.22: *Total time of sessions (minutes)*

Six year olds had the highest playing time and were the most effective in both modes as well (Game effectiveness is due to the next subsection —figure 6.23). Four year olds come second in Total Times but as we mentioned were the least effective in Story/Exploration progression. Five year olds come last in Total Times because one session was prematurely (compared to average playing times) terminated by user (first timeline of 6.11) but their effectiveness in Exploration was much higher than that of the four year olds. These times could also be interpreted as a success of our application to keep children immersed. The prematurely ended session we mentioned was one case out of 18 sessions. In the other sessions, average playing times exceeded 30 minutes and children were prompted by the supervisor to stop if they reached 40 minutes of playing time.

6.2.8.6 Games Statistics

Figure 6.23 shows the average completion time for each game in Exploitation mode by each age. The addition task was not available in four years old children and partially available in five years old ones (they were asked if they had knowledge of numbers addition before the session started).

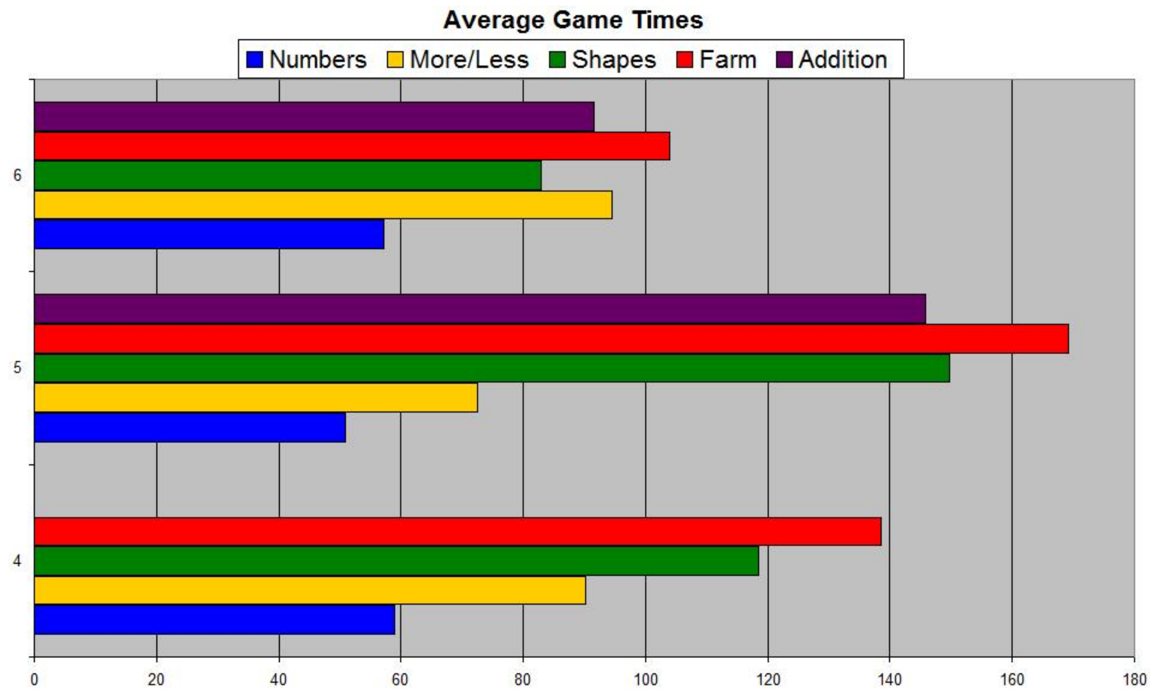


Figure 6.23: *Average Time per Game for each age (addition task not available in 4)*

Six year olds had the best performance almost in all games. As we can see they only have the highest completion time in the More/Less task. This is attributed to the fact that they had the habit of counting the carrots inside and outside of the basket just to make sure (which took significantly more time) whereas other ages gave answers relying on their visual awareness. Five years old children had bigger completion times in Farm (animal recognition) and Shapes tasks than those of the four year olds. This is due to their lower Exploitation times. There were cases where they could not recognize an animal sound or the pentagon shape whereas four year olds would found the answer faster based on their memory of previous gaming sessions. In figure 6.24 we can see an average game completion time considering all games.

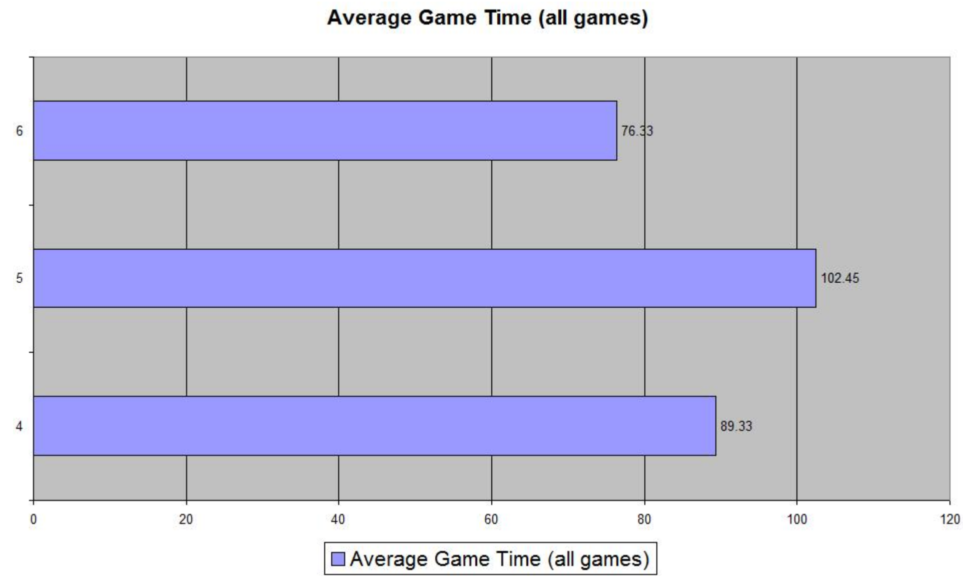


Figure 6.24: *Average Game Completion Time (all games) in seconds*

It is once more clear that six year olds were the most effective in games. The highest average times belong to five year olds. This mannerism regarding higher times than four year olds can be attributed to the following facts: First of all they had access to the Addition task which was rather time consuming (145.8 seconds average) . Secondly they had worse performance in two tasks (Farm, Shapes) for the reasons we mentioned above. Finally they dedicated the least time in Exploitation Mode which they tended to ignore. In figures 6.25 and 6.26 we can see the clear amount (not considering Idle Exploitation Times) of gaming time and the number of games each age played and completed.

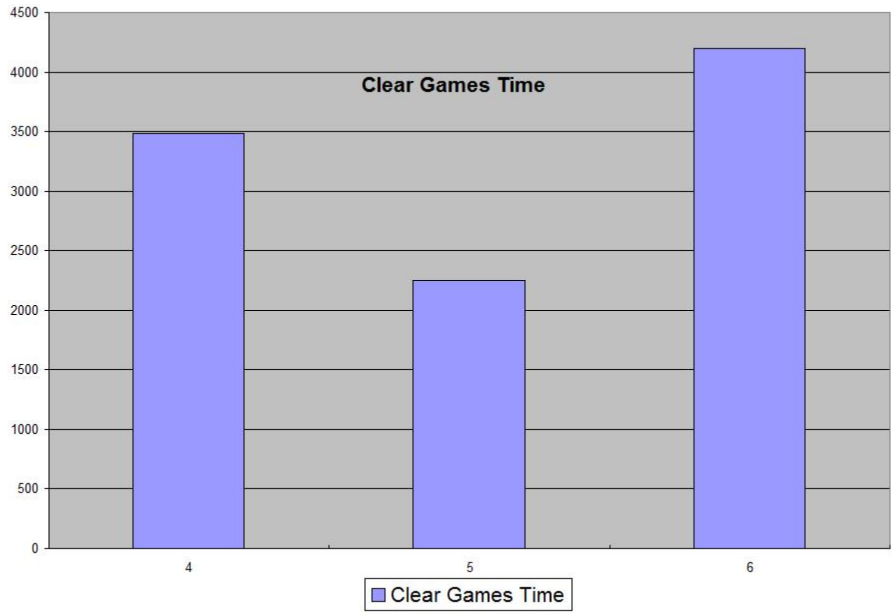


Figure 6.25: *Clear Games Time (excluding Idle Games Time) in seconds*

The differences regarding the Exploitation activity of the five year olds are obvious.

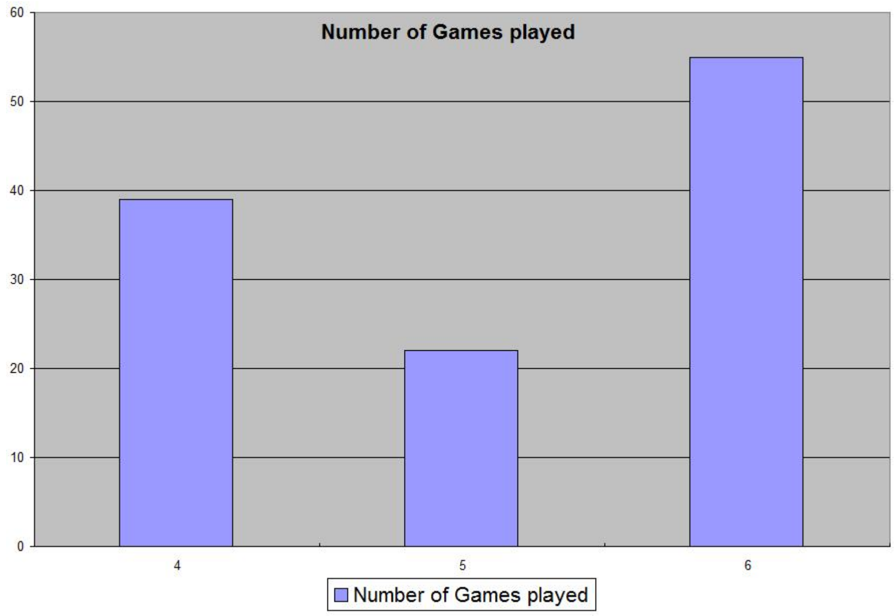


Figure 6.26: *Number of Games each age played*

6.3 Subjective Metrics

Each session was followed by an exit interview with the children where they were asked a series of questions about their experience with our application. In the following two subsections we will present the results of the questionnaires we filled and comment the objective metrics based on personal observations made during the evaluation process.

6.3.1 Questionnaire

A total of 22 questions were made after each session (18 questionnaires total). Some of them regarded mode preferences and what the children liked best in our application, while other questions were meant as a way to check the children's attention and memory of some basic Exploration events. Finally we made some questions to learn what they did not enjoy and how comfortable or not they felt with our experimental setup. As we did with Objective Metrics, we shall also separate some of the questions here into age (4,5,6) and playthrough (first, second) categories where that matters, while treat others globally. We should note here that the questions were made in Greek and in a manner such as not to prejudice children towards an answer. If the children could not perceive the content of what we have asked we would repeat the question in a different way. So, the English translation of the questions that follow does not always represent our exact words in Greek but outlines the content of them.

One of the most significant things we asked the children was: *"Which of the two did you like the best? The conversation with the horse/unicorn (the main ECA) or the games in the panda's (companion ECA) yard?"* This question was meant to check their opinion on the two modes. If the kid was over-thinking it and was reluctant to answer, it was told that he/she could also answer both. In table 6.1 we can see how many children of each age preferred each mode after their First Playthrough.

Age	Exploration	Exploitation	Both
4	3	0	0
5	3	0	0
6	1	1	1

Table 6.1: Subjective Mode Preference of each age in their First Playthrough

There is a major preference here towards Exploration Mode. As we saw in our Subjective Metrics, only six year olds dedicated some extra time in Exploitation (where they had the double Average Transition Time too) and were the only ones (two of them) who answered 'both' or 'Exploitation/Games'. All others with no exception chose Exploration as their preferred mode. Table 6.2 shows the answers that were given after the Second Playthrough.

Age	Exploration	Exploitation	Both
4	1	2	0
5	2	1	0
6	1	0	2

Table 6.2: Subjective Mode Preference of each age in their Second Playthrough

Four years old children (two of them) switched their opinions and expressed a preference for the Exploitation mode. Of the five year olds only one changed her opinion and it happened to be the one that prematurely ended the session (half playing time compared to average) as we have already mentioned. Six year olds now seem to have a somehow increased preference in Exploration by answering 'both' (two of them) and 'Story/Exploration' (one) whereas First Playthrough results (6.1) where equally distributed.

Below we set forth the rest of the questions. Due to the mannerism of some questions and the answers that were given, we shall present and comment only the most popular ones.

- **Did you like/enjoy the game?** All children answered yes in both playthroughs with no exception (even not the one that had the smallest/half session).

- **What differences did you notice since the last time you played?**This question was made only after the revisit's session. Some answers regarded certain themed dialogs that caught the attention of the children (answer ex. 'that i helped him open the chest'). The most popular answer though was a general reference to the difference in the main ECA's speech content.

- **What made you an impression out of the things you have seen or done?**The most popular answer here was *'that i was talking to the Horse/Unicorn' (9 answers with relevant content).* Some other answers regarded the games (2 of them), the outro video and some various specific themed frames (carrot gathering-feeding theme, chicken eggs theme).

- **Which character did you like the most?**The following table(6.3) shows the popularity of our application's characters among the three age groups considering the questionnaires of both visits. In one case a child (age 5) chose both ECA's.

Age	Horse	Panda	Other
4	2	4	0
5	6	1	0
6	4	2	0

Table 6.3: Favorite Character/ECA

The children showed an exclusive preference towards the two main ECAs. There were other characters as well but with minor pre-recorded audio duration and none of them seemed to have been etched in their memory. Five years old kids were rather fond of the Horse/Unicorn (main ECA) and as we mentioned were the ones who maintained their playing patterns and had increased Exploration percentages in both playthroughs.

- **What do you think of Sifounas (Hurricane-the main ECA)?**All children with no exception commented the main ECA as 'very nice/good'.

- **Did you like his story?**Same case here. Of those that traversed through the Story part (Story is optional in subsequent playthroughs) all of them gave positive answers. We shall see in a next question though that long monologues (there were two of them which could have randomly been played in the Story part) seem to overwhelm and tire the children.

- **Did you like the pictures you saw in his story?**The visualizations of the main ECA's thoughts in the story part were highly liked by the children and all of whose who traversed through the Story part answered 'yes'.

- **Tell me one thing that you remember from his story:** Almost all answers were different in this question. That was expected, considering the randomness of the Story part. Four year olds were reluctant to answer to that question. Only 2 out of 6 gave an answer. As we mentioned they had the most limited Attention Span so the somewhat longer Story dialogs/monologues seemed to overload them.

- **What do you think of the Panda (companion ECA)?**Once again all of the children made nice comments on the companion ECA.

- **Was she helpful?**All children answered 'yes'. The help we mention regards some Timeout prompts that were triggered in cases of user inactivity and were meant to guide the children or forward the plot (in cases of low mouse skills).

- **Do you remember her name?**This question was meant as a minor memory/attention test. In tables 6.4, 6.5 we present the answers given after each playthrough.

Age	**YES**	**NO**
4	1	2
5	0	3
6	2	1

Table 6.4: Memory of Panda's name after the First Playthrough

We notice a very low percentage (33.3%) when children tried to remember the Panda's name after the First Playthrough. The name of the Panda was mentioned in the introductory frames and in some random dialogs with the main ECA.

Age	**YES**	**NO**
4	3	0
5	3	0
6	3	0

Table 6.5: Memory of Panda's name after the Second Playthrough

After the Second Playthrough we were just expecting some increased attention/memory performance but, surprisingly enough, all of the children were able to remember the Panda's name.

- **Were the various characters listening to your words?**This question was meant to evaluate the implementation of theWoZ /ASR in our application. All children unanimously answered 'yes'. The children were never aware of the existence and functionality of the WoZ. Such a thing would surely ruin the immersion for them. On the contrary they perceived the ECAs as a separate existence with personality and understanding capabilities (especially the four years old ones). There were even cases where children whispered something to the supervisor so as not to be heard by the ECAs.

- **How did you like the voices of the characters?**All answers were 'nice/good' in this question.

- **Was some character voice funny or scary?**In the First Playthrough Questionnaire one child of each age group found the main ECA's voice somewhat odd and scary but entertaining nonetheless. This was intended from our part considering the voices of many animated cartoon series. As we said, that did not discourage the children (on the contrary it somehow motivated their curiosity) from voting the main ECA as the most popular one. In the revisit's playthrough there was only one comment on the main ECA's voice as 'a bit odd'. All of the other comments regarded the Panda's voice which they found rather funny.

- **Where you tired or bored at any point of the session (and if yes at what point)?**Table 6.6 shows the answers.

Age	YES	NO
4	5	1
5	5	1
6	0	6

Table 6.6: Answers to the 'where you bored/tired at some point?' question

Six year olds where the most enthusiastic users of our application and the most tolerant ones also. Four and five year olds became overwhelmed faster. The most popular points the kids mentioned where 'towards the end' (during the Random Quotes part of Exploration which usually is traversed in the second half of the session –5 year olds), 'when i played games' (during the Exploitation part which became rather tiresome after a certain amount of time for the four year olds) and 'during his Story when he talked much' (the long monologues as we have mentioned are somewhat tiresome for 4 and 5 year olds and render them unable to focus, leading in increased inactivity times and a possible Timeout in the question that follows a monologue).

- **Is there something that seemed odd to you or you didn't like?**Six year olds stated that they liked everything in our application every time they were asked. Some of the four(4 out of 6 questions) and five (1 out of 6 questions) year olds mentioned some random things that they seemed odd to them (ex. the barking sound in the farm game, the fact that the horse was behind wooden bars, the scene were the Panda takes a nap, the colours of the Panda).

- **What would you like to see in this game if you play again?** Although we expected a unique answer from each child in this question, there was one witch was rather popular: 'The Birds'. The birds were the only NPCs in the main Exploration theme with no pre-recorded audio and the only interaction that was available was a mouse click which triggered an animation. Children highly anticipated a conversation with them. Other answers regarded the main ECA, some games from Exploitation mode, the house of the companion ECA and the twin cats.

- **Were you comfortable during the session?**A unanimous answer of 'yes' in this question also.

- **Were the headphones annoying?**Table 6.7 shows the answers.

Age	YES	NO
4	3	3
5	3	3
6	0	6

Table 6.7: Headphones annoyance or discomfort

There was a minor discomfort regarding the headphones at some points. Once again six years old children where the most tolerant and did not express any kind of annoyance while half of the four and five year olds where a bit more eager to express their feelings and answered 'yes' or 'a little bit'.

- **Were you able to hear the animal voices clearly?**A question meant for audio-playback evaluation. All children were able to hear the voices clearly enough and gave 'yes' as an answer (no exception here too).

- **Would you like to play again when i come back?**Only one children (out of 18 total questionnaires) gave 'no' as an answer (the outlined behaviour— first timeline of figure 6.11).

6.3.2 Observations and personal notes

6.3.2.1 The Exploration value of Exploitation

In this Chapter we analyzed the Subjective and Objective metrics of our multimodal application which primarily concern the EvE dilemma in preschoolers. There was a discrimination of Exploration and Exploitation modes and we examined the patterns and preferences of children aged 4-6 towards them. There is however another point of view when examining the content of each mode. We could consider that the Exploitation mode contains Explorative elements from the perspective that it is something new and unexplored for the children in the first moments of game-playing. From the point that the kid has explored all of the possible games, this mode gains its literal meaning of Exploitation. Taking the aforementioned into consideration along with our subjective and objective metrics, we can reach to a different conclusion, especially for six year olds. Although their time distribution and mode choice probability did not favour Exploration mode as that of the five year olds did, the patterns they followed indicate a much more Explorative behaviour:

- In the First Playthrough they had the highest Average in Exploitation Time (before a transition) and the lowest one in Exploration (figure 6.8).
- In the subsequent Playthrough this behaviour is changed. They have the highest average in Exploration time while the Exploitation Average is reduced by 1/3 approximately and lies in the middle among the other two age groups (figure 6.9).
- In both playthroughs they played the games of Exploitation mode in a consecutive manner (figures 6.5 and 6.12). When they realized there was no escalation of knowledge usage in Exploitation and once they had cycled through all of the games, they showed a major preference in Exploration mode, especially in the second playthrough.

What the six year olds did was that they dedicated some extra time (compared to ages 4 and 5) in Games/Exploitation in their first playthrough which has a mostly predefined Exploration mode structure (meaning that all ages will traverse through the same Exploration frames more or less, for the reasons we explained in Chapter 4). In their second playthrough (generic and random Exploration mode structure) they dedicated some time in Exploitation mode during the beginning of the session so as to remember the games (second session was usually after one week) which they once more played consecutively (thus explaining the very high Games-Games Transition Type percentage — figure6.14). After that point, the Exploitation mode acts only as a break from Exploration (1-2 games maximum in each transition). So actually, from the perspective that we mentioned, this behaviour of the six year olds (exploring all gaming possibilities and then focusing on conversing with the ECAs) seems highly predisposed towards Exploration. The subjective and objective metrics of the five year olds indicate the highest Exploration mode preference. But there was almost no point where they examined the full potential of Exploitation mode and thus their behaviour could be characterized as less Explorative than that of the six year olds. Four years old children presented the least Explorative behaviour because from the point that they felt comfortable and capable in the Exploitation mode of the second playthrough, they increased their gaming times (totals and averages) which were the highest and decreased their Exploration ones (which were the lowest). As a conclusion, if we were to consider and accept that there is an Exploration value in the first moments of Exploitation, we could therefore deduce that *Explorative behaviour is amplified with age progression (4-6 years old).*

6.3.2.2 Stereotyping and Verbal Disinhibition issues with ECA's

In Chapter 2 we mentioned some ethical implications which arise with the use of ECAs. However important (or even exaggerating by some opinions) these may seem, it is not part of this work to study and analyze these potential implications. Nonetheless we should mention that there was always in mind that our application is intended for very young children and that there are major responsibilities for that fact. There was also a major effort to avoid any kind of stereotyping in our application. There seems to be a belief though, that stereotyping cannot always be avoided due to the fact that a large number of applications and products designed for children (games, animated series, comic books, songs) draw their themes from social phenomena (excluding fantasy worlds/themes/songs). It is the Society's structure that enforces stereotyping and such applications just reproduce and imitate these facts in order to become more familiar and intimate for the children, thus enhancing realism and immersion. Realism though, in aspects of gameplay, is not that important for children games (it may even be perceived as a negative attribute) as it is for adult game development. Stereotyping must be avoided in children applications, considering how 'malleable' their minds are. Finally we would like to point that there were no indications of verbal Disinhibition in any of our sessions and it would be rather surprising and upsetting if there was any for such young children.

Chapter 7

Conclusions and Future Work

7.1 Conclusions

As mentioned in Chapter 2, the terms Exploration and Exploitation have increasingly come to dominate many fields of scientific research. In this work we have investigated how preschool children (ages 4-6) instinctively respond to the EvE dilemma. For that purpose we have integrated and extended the on-line multimodal platform designed by Kannetis in his Master Thesis [1]. Our application is divided into two discrete modes: The Exploitation mode which consists of five games based on popular preschool activities which were designed by Kannetis [1] and Meliopoulos [2]. In this mode children have the opportunity to use/exploit their existing knowledge on some fundamental things (numbers/animal/shape recognition, addition task and quantity comparison). The Exploration mode designed in this work, was implemented with the use of ECAs. In this mode the children have the opportunity to engage in conversation with various characters (the main ECA, the companion ECA and secondary ECAs), participate in interactive storytelling and initiate small talk with them. Exploration mode was further separated into two different structures for the first and any subsequent playthroughs. There is no demand of knowledge and use of skills here. There is also no immediate goal or reward and the motivation lies in just exploring the many and random conversational possibilities. The modality usage of our application supports speech and mouse.

Preliminary experiments in preschool environments (kindergardens) provided us with various subjective and objective metrics and statistics. Nine children aged 4-6 (three of each age) participated in our evaluation procedure where each one had a total of two sessions (the second session was after a week approximately). In Chapter 6 there is a detailed analysis of all metrics and statistics concerning the EvE dilemma and some miscellaneous aspects. However we shall summarize the outline of the most important results:

First of all, all ages followed the pattern of frequent mode toggling. The Attention Span of each child combined with personal preferences can affect the frequency of mode switching. Decisions regarding transitions from Exploration to Exploitation mode had almost the same 'coinflip' percentage in all ages. In the first playthrough four and five year olds showed an increased preference towards the Exploration mode (regarding time distribution, average transition times and questionnaire results) whereas six years olds dedicated some extra time in Exploitation thus shaping a more balanced distribution of time among the two modes. In their second playthroughs four and six years old children seem to invert their behaviour: Four year olds now have become familiar with the games in Exploitation mode and seem to prefer remaining in the 'safety' of what they comprehend thus reducing all Exploration times and increasing the Exploitation ones. Storytelling and the concept of conversing with ECAs can become overwhelming for them at some point. They had the slowest Exploration progression pace (highest inactivity times and number of timeouts) and used this mode as a break from games. Six year olds now increase their Exploration interest. They had the most consistent patterns in both playthroughs. What they did was that they dedicated some time in the first moments of the session to cycle through all (or the most at least) of the games which they played in a consecutive manner. When they realized that there was no challenge or new potential for them they focused on Exploration mode with minor (1-2 games maximum) Exploitation breaks. Five year olds maintained their patterns and preferences and once more almost completely ignored Exploitation mode for the first half of the session.

There is however the concept that Exploitation has an Exploration value in the first moments of engagement. Considering this fact and the patterns six year olds followed, we could interpret their behaviour as the most Explorative one. They explored all possibilities of Exploitation mode (successive game playing) in the beginning and afterwards massively increased their interest in Exploration mode. Five year olds, no matter their preference towards Exploration in both playthroughs, almost never dedicated enough time in Exploitation to test their potentials. Four year olds and their patterns are not affected by this concept and maintain their Exploitative (compared to other ages) behaviour.

7.2 Future Work

More experiments are needed in order to achieve better results and safer conclusions. The Evaluation Methology was rather time consuming and there were a total of 18 sessions (two for each child). Due to the random and generic structure of our application after the first play-through, the same user can have multiple sessions (which is highly advised for more robust behavioural patterns) without repeated occurrence of same dialog segments. To further enforce this structure more themes could be integrated into the Random Quotes (small talk with ECAs) section of the Exploration mode.

Audio feature extraction from the collected data could lead in classifying the child's emotions into categories which can be corresponded to certain modes and themes of our application where we could investigate engagement, attention and immersion. Furthermore the Wizard of Oz should be replaced with an Automatic Speech Recognition System. With the necessary acoustic and language models we could replace the Wizard of Oz module with an Automatic Speech Recognizer. The ASR module is already exist and functional for adult users. Further improvement of the platform could be done by supporting more modalities in the future. The touch modality and gesture recognition is believed to be a very natural choice for children and might prove very interesting for them.

Finally, a more humanist approach to the matter would be rather important. It is crucial to examine any kind of ethical implications that may arise with the extended use of ECAs, especially when children are involved.

Appendix A

System setup manual

In this Appendix the installation manual of the platform is provided.

Instructions:

- Install the Java Development Kit (JDK) 1.5 or later. In this work we used JDK 1.6.0.11
- Install the Apache Tomcat web server. Tomcat 5.5 is recommended.
- Install the My SQL Database and run the children.db script from the thesis cd in order to create the database.
- Copy the files from the JSP directory of the thesis cd to the /Webapps/ROOT/children folder in the tomcat directory.
- Compile the java beans using the following commands:

  javac web/persistence/*.java

  javac web/children/*.java

  You have to be sure that the username and password variables in the javabeans are the same with the ones you put during the database installation.
- In the /Webapps/ROOT/WEB-INF/ folder of the Tomcat, create a directory with the name “classes” and copy the web folder with the compiled javabeans in it.

- Copy the mysql connector from the thesis cd to the /Webapps/ROOT/WEB-INF/lib and the web.xml to the Tomcats conf folder.
- Compile the Controller Applet using the following commands: javac Controller.java

  jar cvf controller.jar *.class

  jarsigner -keystore myKeyStore controller.jar me

  Copy the controller.jar file to the /Webapps/ROOT/children folder.
- Startup the web server. In order to be sure that it is working, open the following address in your browser: http://localhost:8080/. In order to have access to the platform open the following address: http://localhost:8080/children

Glossary

ActionScript is a scripting language developed by Adobe. It is a dialect of ECMAScript (meaning it has the same syntax and semantics of the more widely known JavaScript), and is used primarily for the development of websites and software targeting the Adobe Flash Player platform, used on Web pages in the form of embedded SWF files. Originally developed by Macromedia, the trademark in the language is now owned by Adobe (which acquired Macromedia in 2005), though the language itself is "open source" in the sense that its specification is offered free of charge and both an open source compiler (as part of Adobe Flex) and open source virtual machine (Mozilla Tamarin) are available. 23, 29, 48

Animal Assisted Therapy (AAT) is a type of therapy that involves an animal with specific characteristics becoming a fundamental part of a person's treatment. Animal-assisted therapy is designed to improve the physical, social, emotional, and/or cognitive functioning of the patient, as well as provide educational and motivational effectiveness for participants. AAT can be provided on an individual or group basis. During AAT, therapists document records and evaluate the participant's progress. 34

ASR Abbreviation for Automatic Speech Recognition. 17, 23, 28, 58, 59, 91, 99

Attention Span is the amount of time that a person can concentrate on a task without becoming distracted. Most educators and psychologists agree that the ability to focus one's attention on a task is crucial for the achievement of one's goals. 13, 15, 36, 76, 81, 90, 98

Bootstrapping : The computer term bootstrap began as a metaphor in the 1950s. In computers, pressing a bootstrap button caused a hardwired program to read a bootstrap program from an input unit and then execute the bootstrap program which read more program instructions and became a self-sustaining process that proceeded without external help from manually entered instructions. As a computing term, bootstrap has been used since at least 1958. 106

Disinhibition is a term in psychology used to describe a lack of restraint manifested in several ways, including disregard for social conventions, impulsivity, and poor risk assessment. Disinhibition affects motor, instinctual, emotional, cognitive and perceptual aspects with signs and symptoms similar to the diagnostic criteria for mania. Hypersexuality, hyperphagia, and aggressive outbursts are indicative of disinhibited instinctual drives. 20, 95

EvE Abbreviation for Exploration vs Exploitation. 11, 21, 31, 63, 64, 93, 97, 98

Media Equation : The term 'media equation' means that Media equal real life. It implies that people (Audience) process technology-mediated experiences in the same way as they would do nonmediated experiences, because an 'individual's interactions with computers, television, and new media are fundamentally social and natural, just like interaction in real life' (Reeves & Nass 1996, 5). In 1996, Reeves and Nass first introduced this theory, and carefully summarized their previous media equation studies in the seminal book The media equation. Since the publication of this book, Nass, Reeves, and their colleagues have expanded the domain of the theory from traditional communication settings to radically different areas, from e-commerce to human–robot interaction.In general, media equation studies can be categorized as having one of two themes: audience responses to physical features of traditional media (Audience Research), and user responses to social characteristics of computers and software agents (Human—Computer Interaction; Computer—User Interaction; Avatars and Agents). 20

Multi Armed Bandits : In statistics, particularly in the design of sequential experiments, a multi-armed bandit takes its name from a traditional slot machine (one-armed bandit). Multiple levers are considered in the motivating applications in statistics. When pulled, each lever provides a reward drawn from a distribution associated with that specific lever. The objective of the gambler is to maximize the sum of rewards earned through a sequence of lever pulls. In practice, multi-armed bandits have been used to model the problem of managing research projects in a large organization, like a science foundation or a pharmaceutical company. Given its fixed budget, the problem is to allocate resources among the competing projects, whose properties are only partially known now but may be better understood as time passes. In the early versions of the multi-armed bandit problem, the gambler has no initial knowledge about the levers. The crucial tradeoff the gambler faces at each trial is between "exploitation" of the lever that has the highest expected payoff and "exploration" to get more information about the expected payoffs of the other levers. 11

Multithreading computers have hardware support to efficiently execute multiple threads. These are distinguished from multiprocessing systems (such as multi-core systems) in that the threads have to share the resources of a single core: the computing units, the CPU caches and the translation lookaside buffer (TLB). Where multiprocessing systems include multiple complete processing units, multithreading aims to increase utilization of a single core by leveraging thread-level as well as instruction-level parallelism. As the two techniques are complementary, they are sometimes combined in systems with multiple multithreading CPUs and in CPUs with multiple multithreading cores. 24, 25

NPC Abbreviation for Non Playable Character. 92

Organizational Learning is an area of knowledge within organizational theory that studies models and theories about the way an organization learns and adapts. 11

Persona Effect : As a result of rapid advances in animated agent technology, the prospect of deploying animated pedagogical agents on a broad scale is quickly becoming a reality. Because these agents can provide students with customized advice in response to their problem-solving activities, their potential to increase learning effectiveness is significant. In addition, however, these agents can also play a critical motivational role as they interact with students. As a result, students may choose to use interactive learning environments frequently and for longer periods of time.To investigate the affective impact of animated pedagogical agents on students' perception of their learning experiences, there was an empirical study with 100 middle school students. The study revealed that well crafted lifelike agents have an exceptionally positive impact on students. Students perceived the agents as being very helpful, credible, and entertaining. This persona effect held strong even for an agent whose communicative behaviors were muted. The study also found that combinations of types of advice can increase students' positive perception of the agent and increase learning performance. This work represents a promising first step toward developing an understanding of the impact that animated pedagogical agents can have on children's learning experiences. Perhaps the greatest challenge lies in determining precisely which characteristics of these agents are most effective for particular age groups, domains, and learning contexts. 20

Temporal Difference Learning Algorithm : Temporal Difference (TD) learning is a prediction method. It has been mostly used for solving the reinforcement learning problem. TD learning is a combination of Monte Carlo ideas and dynamic programming (DP) ideas. TD resembles a Monte Carlo method because it learns by sampling the environment according to some policy. TD is related to dynamic programming techniques because it approximates its current estimate based on previously learned estimates (a process known as Bootstrapping). The TD learning algorithm is related to the temporal difference model of animal learning. As a prediction method, TD learning takes into account the fact that subsequent predictions are often correlated in some sense. In standard supervised predictive learning, one learns only from actually observed values: A prediction is made, and when the observation is available, the prediction is adjusted to better match the observation. The core idea of TD learning is that we adjust predictions to match other, more accurate, predictions about the feature. This procedure is a form of Bootstrapping. 13

The Gittins Index : For an n-armed bandit problem, an agent is required to choose between n options, each of which delivers reward with a probability pi. The probability of obtaining reward from a bandit, pi, may change through time but only when a choice is made for that bandit. The goal for the agent is to maximize expected rewards, Vi, where rewards earned in the future are discounted by an exponential discount factor Gittins & Jones proved that optimal performance can be obtained by tracking a single index Vi of the form of:

$$V_i = sup_{T>0} \frac{\langle \sum_{t=0}^{T} \delta^t * R_i(t) \rangle}{\langle \sum_{t=0}^{T} \delta^t \rangle}$$

. 11, 12

VAD Abbreviation for Voice Activity Detection. 24, 25, 28, 58, 78

WoZ Abbreviation forWizard of Oz. 17, 22, 23, 28, 58, 59, 61, 91

References

[1] T. Kannetis, “Multimodal dialogue systems for preschoolers,” Master’s thesis, Technical University of Crete, 2009. ii, iv, v, 2, 3, 17, 21, 44, 47, 56, 77, 97

[2] S. Meliopoulos, “Multimodal system for preschool children,” 2008. ii, v, 21, 44, 45, 97

[3] S. M. M. Jonathan D Cohen and A. J. Yu, “Should i stay or should i go? how the human brain manages the trade-off between exploitation and exploration,” *Philosophical Transactions of the Royal Society Biological Sciences*, vol. 362(1481), May 2007. 1, 11

[4] J. G. March, “Exploration and exploitation in organizational learning,” *JSTOR*, vol. 2, pp. 71–87, February 1991. 2, 11

[5] Atherton J.S. (2010) Learning and Teaching; Piaget’s developmental theory [On-line] UK: Available `http://www.learningandteaching.info/learning/piaget.htm`. 6

[6] Santrock, J.W. (2007). Cognitive Development Approaches. In Emily Barrosse (Ed.), A Topical Approach to Life-Span Development (pp. 225-230). New York, NY: Beth Mejia. 7

[7] Atherton J S (2010) Learning and Teaching; Assimilation and Accommodation [On-line] UK: Available: `http://www.learningandteaching.info/learning/assimacc.htm`. 10

[8] A. Demetriou, A. Efklides, and D. M. Shayer, eds., *Neo-Piagetian Theories of Cognitive Development Implications and Applications for Education*. Routledge, UK, 1992. 10

[9] J. Pascual-Leone, *A mathematical model for the transition rule in Piaget's developmental stages*. Acta Psychologica, 1970. 10

[10] R. Case, "Intellectual development. birth to adulthood.," *New York: Academic Press*, 1985. 10

[11] G. S. Halford, "Children's understanding: The development of mental models," *Hillsdale, NJ: Erlbaum*, 1993. 10

[12] K. W. Fischer, "A theory of cognitive development: The control and construction of hierarchies of skills," *Psychological Review*, 1980. 10

[13] Demetriou, A., Christou, C., Spanoudis, G., & Platsidou, M. (2002). The development of mental processing: Efficiency, working memory, and thinking. Monographs of the Society of Research in Child Development, 67, Serial Number 268. 11

[14] A. K. Gupta and K. G. Smith, "The interplay between exploration and exploitation," *Academy of Management Journal*, vol. 49, no. 4, pp. 693–706, 2006. 11

[15] C. Chatham, "Exploration & exploitation balanced by norepinephrine & dopamine," *Developing Intelligence*, 2007. 12, 13

[16] M. S. G. Samuel M. McClure and J. D. Cohen, "An exploration-exploitation model based on norepinephrine and dopamine activity," *Center for the Study of Brain, Mind, and Behavior*, 2005. 12, 13

[17] A. Bruckman and A. Bandlow, "Hci for kids," in *The Human-Computer Interaction Handbook: Fundamentals, Evolving Technologies, and Emerging Applications* (J. Jacko and A. Sears, eds.), Lawrence Erlbaum and Associates, 2002. 13, 14, 15

[18] S. Narayanan and A. Potamianos, "Multimodal systems for children," in *Proc. European Conf. on Speech Communication and Technology*, (Budapest, Hungary), September 1999. 14, 17

[19] S. Narayanan and A. Potamianos, "Creating conversational interfaces for children," vol. 2, IEEE TRANSACTIONS ON SPEECH AND AUDIO PROCESSING, February 2002. 14, 17, 18

[20] Malone, T.W. What makes things fun to learn? Proceedings of Conference on Video Games and Human Development: A Research Agenda for the 80's, Harvard University Graduate School of Education, Cambridge, Massachusetts, May 24, 1983. 15, 16, 47

[21] S. Buisine and J.-C. Martin, "Children's and adults' multimodal interaction with 2d conversational agents," April 2005. 18, 19

[22] J. B. A. L. Joakim Gustafson, Linda Bell and M. Wiren, "The nice fairy-tale game system," in *Proceedings of SIGdial, Boston*, April 2004. 19, 31

[23] A. D. Angeli, "Ethical implications of verbal disinhibition with conversational agents," *PsychNology Journal*, vol. 7, no. 1, pp. 49–57, 2009. 20

[24] Macromedia Flash support page. Available at: http://www.adobe.com/products/. 23, 29

[25] W. Crain, "Animal feelings: Learning not to care and not to know," *Green Money*, 2009. 34

[26] Reichert E. (2003) Animal Assisted Therapy for Sexually Abused Children in: Catherine Ford Sori and Lorna Hecker: The Therapist's Notebook for Children and Adolescents. Homework, Handouts, and Activities for Use in Psychotherapy. Haworth Clinical Practice Press. New York. pp. 163-168. 35

[27] Cassell, J., Bickmore, T. "Negotiated Collusion: Modeling Social Language and its Relationship Effects in Intelligent Agents", User Modeling and Adaptive Interfaces(12) 1-44, 2002. 39

[28] N. Beringerb,U. Kartal, K. Louka, F. Schiel, and U. Turk, "Promice: A procedure for multimodal interactive system evaluation," *In Proceedings of the*

*LREC Workshop on Multimodal Resources and Multimodal Systems Evaluation,*Las Palmas, Spain, pp. 271-280, 2002. 55