ABSTRACT

Research into gameplay can contribute to more self-conscious approaches to design, allowing designers to create effective gameplay with less testing, or to target specific cognitive and emotional affects of gameplay for serious games applications. Self-conscious design includes theoretically motivated design of game systems to facilitate gameplay motivated by cognitive, scientific and/or rhetorical theories of game affect and functionality. Deepening the understanding of gameplay requires a consideration of basic epistemological questions about the nature of understanding. Understanding gameplay is a matter of generating mappings to explanatory frameworks in alternative interpretation paradigms. All games are cognitive skill learning environments, and an especially useful approach that may aid in the creation of more self-conscious game design practices is to conduct research into gameplay using theories and methods of cognitive science and cognitive psychology. On this basis, a framework is proposed based upon the integration of schema theory with attention theory. Cognitive task analysis provides a foundation for developing schema descriptions, which can then be elaborated according to more detailed models of cognitive and attentional processes. This approach provides a rich explanatory framework for the cognitive processes underlying gameplay. Playing a commercial PC or console game is a highly visual activity regardless of whether the purpose is entertainment or situated learning. Information about the visual attention of the player is an important foundation for detailed schema modeling. A range of different eyetracking equipment has been used in many studies of visual cognition. However, very few studies describe dynamic stimuli involving the visual interaction of a user/player with a moving 3D scene displayed on a computer screen. In order to address this, a software interface has been developed linking a Tobii™ eyetracking system with the HiFi game engine for use in automated logging of dynamic 3D objects of gaze attention. The system has been verified in a detailed study, confirming correct operation of the system as well as providing a characterisation of its spatial and temporal accuracy. The integrated Tobii/HiFi system has been validated in a study to test three hypotheses concerning visual attention in a first-person shooter (FPS) computer game. Firstly, the cuing effect of the passive gun graphic on visual attention was tested, with no evidence being found to support this hypothesis. A second hypothesis, that a player directs their gaze at a target opponent while shooting at them, was found to be supported in most cases, while in a small percentage of cases targeting is achieved in peripheral vision. Finally, in most cases, a player targets the nearest opponent. These results provide a baseline for further investigations in which the stimulus game design may be modified to provide more detailed models of the visual cognitive processes involved in gameplay and how they are involved in player decision-making.
GAMEPLAY

(3D Game Engine + Ray Tracing =

Visual Attention through Eye Tracking)

Charlotte Sennersten
GAMEPLAY

(3D Game Engine + Ray Tracing = Visual Attention through Eye Tracking)

Charlotte Sennersten

School of Techoculture, Humanities and Planning
Blekinge Institute of Technology
SWEDEN
# Table of Contents

Abstract 9  
Acknowledgments 10  
List of Licentiate Publications 11  
Summary of Scientific Contribution 12  

## PART I  
Background 15  
Introduction 15  
  General Approach 15  
  The First Person Shooter game 18  
Theoretical Framework 19  
  Object(ive) Knowledge 19  
  Cognition, Visual Attention and Schema Theory 24  
Methodology Development 28  
Empirical Experiments & Results 29  
Practical Notes on the Licentiate Format 31  

## PART II – THEORY, CONTEXT & METHOD  
What does it mean to understand Gameplay? 35  
An Innovation-Oriented Game Design Meta-Model Integrating Industry, Research and Artistic Design Practices 49  
A Cognitive Framework for the Analysis of Game Play: Tasks, schemas and Attention Theory 71  
Game Play Schemas: From Player Analysis to Adaptive Game Mechanics 85  

## PART III – METHODOLOGY DEVELOPMENT  
Eye Fixations in an Action Game Tutorial 101  
Verification of an Experimental Platform Integrating a Tobii Eyetracking System with the HiFi Game Engine 105  
Evaluation of Real-Time Eye Gaze Logging by a 3D Game Engine 161  

## PART IV – EMPIRICAL EXPERIMENTS & RESULTS  
An Investigation of Visual Attention in FPS Computer Gameplay 179  

Ongoing Research Directions 197  
Epilog 199  
References 200
Abstract

Research into gameplay can contribute to more self-conscious approaches to design, allowing designers to create effective gameplay with less testing, or to target specific cognitive and emotional affects of gameplay for serious games applications. Self-conscious design includes theoretically motivated design of game systems to facilitate gameplay motivated by cognitive, scientific and/or rhetorical theories of game affect and functionality.

Deepening the understanding of gameplay requires a consideration of basic epistemological questions about the nature of understanding. Understanding gameplay is a matter of generating mappings to explanatory frameworks in alternative interpretation paradigms. All games are cognitive skill learning environments, and an especially useful approach that may aid in the creation of more self-conscious game design practices is to conduct research into gameplay using theories and methods of cognitive science and cognitive psychology.

On this basis, a framework is proposed based upon the integration of schema theory with attention theory. Cognitive task analysis provides a foundation for developing schema descriptions, which can then be elaborated according to more detailed models of cognitive and attentional processes. This approach provides a rich explanatory framework for the cognitive processes underlying gameplay.

Playing a commercial PC or console game is a highly visual activity regardless of whether the purpose is entertainment or situated learning. Information about the visual attention of the player is an important foundation for detailed schema modelling. A range of different eyetracking equipment has been used in many studies of visual cognition. However, very few studies describe dynamic stimuli involving the visual interaction of a user/player with a moving 3D scene displayed on a computer screen. In order to address this, a software interface has been developed linking a Tobii® eyetracking system with the HiFi game engine for use in automated logging of dynamic 3D objects of gaze attention. The system has been verified in a detailed study, confirming correct operation of the system as well as providing a characterisation of its spatial and temporal accuracy.

The integrated Tobii/HiFi system has been validated in a study to test three hypotheses concerning visual attention in a first-person shooter (FPS) computer game. Firstly, the cuing effect of the passive gun graphic on visual attention was tested, with no evidence being found to support this hypothesis. A second hypothesis, that a player directs their gaze at a target opponent while shooting at them, was found to be supported in most cases, while in a small percentage of cases targeting is achieved in peripheral vision. Finally, in most cases, a player targets the nearest opponent. These results provide a baseline for further investigations in which the stimulus game design may be modified to provide more detailed models of the visual cognitive processes involved in gameplay and how they are involved in player decision-making.
Acknowledgements

From 2005 to now 2008 three years have passed and many people have been included in this process. I would like to give a big thank you to Högskolan på Gotland (HGO) for supporting this project that was not easy to start up because it had no established ground to lean on, and no established discipline in academia. All credit to HGO to dare to support a high risk project like this. Foremost a big thank you to Craig Lindley with his great knowledge in Games Studies, Artificial Intelligence, Computer Science, Philosophy and how to bind things together. It was space engineering, theater, dance, film, music and our former experiences in these kind of environments that made us start to understand what direction to take. My forever colleague and partner. Special thanks also to my current and past colleagues in the Game and Media Arts laboratory at BTH: Lennart Nacke, Dennis Sasse and Sophie Stelmach. Thanks and love to my beloved daughter Filippa that my heart is pumping for, and all the situations we have had to go through and how tremendously strong she has been and is. She is now in her 9th year. My parents, I truly do love you, and without you this work would never have been completed. Thanks for all the hours and all the days you have supported me, Filippa and Roxy. A big warm hug to you and we remember that alone we are weak and together we are strong. Swedish Research Defence Agency (FOI) in Linköping with Erland Svensson, Martin Castor, Björn Lindahl, Johan Hedström, Jens Alfredson and many more. Lena Trojer has been generous and warm and given me insight into the TechnoScience discipline at Blekinge Institute of Technology (BTH). From insight into another new subject area Lena has shown me how to open up for subjects that otherwise would have little or no legitimacy. FUGA, Fun of Gaming, an EU-funded project that has run in parallel with this licentiate project. I would like to thank all people involved and the insight they have given me. The aim of the project has been to “measure the impossible” which means that gameplay can be measured psychophysiologicaly. Thanks EU & Finland and foremost Niklas Ravaja! From scratch with a young discipline “Game Development” with no academic history, we have tried step by step to build the subject and improve it.
List of Licentiate Publications

This thesis will be based upon the following articles:


Summary of Scientific Contribution

The contributions of this research project are:

- Definition of a theoretical framework for the empirical analysis and study of computer gameplay based upon cognitive science
- Implementation of the real-time interconnection between a game engine and an eyetracker for automated gaze object logging
- Definition and implementation of a detailed verification methodology for the integrated game engine/eyetracker
- Spatial and temporal accuracy characterization for the integrated game engine/eyetracker
- Definition of an empirical study addressing three hypotheses regarding visual attention in computer gameplay
- Data that supports the counter hypothesis to hypothesis H1: the visually static diagonal shape and position of the gun graphic, which does not have a function in aiming the weapon, cues and directs visual attention within the screen, preceding and/or independently of the attentional demands of the player’s in-game task(s). That is, no evidence was found in support of this hypothesis.
- Data that supports the counter hypothesis to hypothesis H2: gaze is firstly directed upon an opponent prior to shooting the opponent. The gathered data confirmed that this is the case in 88% of cases, giving an indication of the magnitude of the role of peripheral vision in the remaining 12% of cases.
- Data that supports the counter hypothesis to hypothesis H3: for otherwise equivalent opponents, the closest opponent will be targeted first. The gathered data confirmed that this is the case in 77% of cases.
- Established a baseline for comparison in ongoing experimental studies.
- Established a database of experimental data for 20 experiment participants supporting rich ongoing investigation and analysis.
PART I
Background

This study has evolved from an eyetracking study of gameplay constituting the masters project carried out by the author at the Cognitive Science Department at Lund University in 2004 (Sennertsten, 2004). At that time I had to rely upon manual analysis from captured video screens. Today, in collaboration with FOI, I have developed a method where a Game Engine and Eye Tracker have been interconnected to automate the time-consuming object logging process. The masters study did not start out with a theoretical hypothesis since the game subject had no unified theory to start out with and there were no existing studies involving explanatory theory supported by empirical results. Hypotheses motivating ongoing experimental work were devised afterwards in response to the results and behavioral patterns found in the initial study.

Human-Computer Interaction (HCI) is a subfield of Computer Science. HCI is the study of the interaction between humans and computers. The definition of HCI is “The study of how humans interact with computers, and how to design computer systems that are easy, quick and productive for humans to use.”¹ Working with the Swedish Defence Research Agency, I saw that Human Factors and Human Computer Interaction methodologies were important when evaluating pilot performance and safety, and that this knowledge could feed into building knowledge about computer gameplay. Human Factors, which differs slightly from HCI, is “The science of understanding the properties of human capability (Human Factors Science), the application of this understanding to the design and development of systems and services (Human Factors Engineering) and the art of ensuring successful application of Human Factors Engineering to a program”². Hence the work here derives from related work in HCI and Human Factors. In this study our eyes are of major interest because these organs are one of the primary links between the player and the game running on the computer. Social interaction is not taken into consideration in order to simplify the task of developing a research baseline and validate the general relevance of a cognitive perspective, but will be of great importance as this approach to game research evolves.

Introduction

General Approach

This licentiate thesis presents a theoretical framework derived from theories in cognitive science that can be used to model the generative foundations of observed patterns within game interaction, a new method for studying visual behaviour in gameplay, initial empirical results gathered using the method, and publications written in collaboration with others. The aim is to present the foundations upon which ongoing studies within the doctoral project will be based. This project started out with the working name “Game Design, Semiotics & Psychophysiology” in 2005. These are the elements from which the project attempts to build a framework for ongoing game research. Here it will be useful to briefly summarize each of the elements of this title. Cognition is the main focus of interest

of the project and especially cognitive psychology as a basis for how to interpret the visual gaze patterns of gameplay. The visual attention of players shows us where they attend to on the computer screen and “Attention theory provides an account of the energetic resources available to cognition, for the distribution of energy (or attention) to the cognitive resources that use (or manifest) it (Lindley, Nacke & Sernersten 2007). There are many variables to account for in a computer game, therefore I have an emphasis in devising studies by deconstructing and building our gameplay understanding from basic variables describing aspects of the game design and exploring the resulting gameplay using principles and theories from cognitive psychology. The research of cognitive psychologists “focuses on describing particular cognitive phenomena, such as how people recognize faces or how they develop expertise” (Sternberg 2003), “…we approach cognitive psychology by asking what it has to tell us about how people carry out everyday activities” (Smyth, Collins, Morris and Levy 1996). Gärdenfors approaches thinking from an evolutionary perspective: “In the light of Darwin’s theory, the question about our origin instead becomes: what is it that has made humans different from other animals?” The answer is “Humans are the only animals who have a symbolic language” (Gärdenfors 2007). “Our species is referred to as homo sapiens” where “the goal of cognitive psychology is to understand the nature of human intelligence and how it works” (Anderson 2000).

Game Design involves a range of different skills from different fields. Game mechanics, programming, graphics/visual arts, audio etc. and all these fields have their own design processes but also have to be synchronized with one another. To be a game designer means that all these skills have to be understood “While good game designers can design good games based upon experience, design principles and extensive tuning, the relationship between the intentions of the designer, the player experience and the mechanisms learned by players to facilitate play have not yet been studied extensively or in depth in an explicit way.” (Lindley & Senersten 2006). In Lindley & Senersten (2007/8) it is stated that “implicit and unsconscious design cultures on one hand and explicit, self-conscious design cultures on the other provide... a principle for interrelating a variety of game design approaches within a coherent game design meta model specifically concerned with design innovation”. Game play (game-play) (Björk & Holopainen, 2005) is the interactive aspect of the game design (where gameplay is what a player actually does/experiences). Core concepts in game play design include the design of: a 3D virtual environment, objects in that environment (which may change state), and “rules governing changes of state of objects, such as position, in response to the state of other objects and/or decisions made by the player and the reward and punishments given to the player as a result of changes to the state of the game.”3 In Lindley & Senersten (2006) it is stated that “Game design features include built in game tasks [that may be] modeled by a process of game task analysis, that ground the in-game meaning of design features”. This same paper takes the concept of a nomology as a coherent integration of theory and empirical methods that constitutes a scientific approach, and proposes that the abstract concept of a nomology “...can be instantiated

in a more specific form based upon attention theory and schema theory, together with a methodology of cognitive task analysis” in order to provide a theoretical and empirical framework for studying gameplay based upon cognitive science. Using a textual metaphor we can say that game play design involves the creation of a variety of verbs, that these verbs form adjectives in the player, and that objects in the game world are nouns. Some examples of verbs include shooting, hiding, trading, learning, etc. These are general elements that should be integrated within a nomology for scientifically understanding gameplay.

_Semiotics_ is a very important tool in the arts (e.g. painting and film) where studies focus on objects (signs and symbols) and the meaning of them individually or in their interrelations. “The semiotics of computer game design is fundamental to design practice, but is generally involved in an unselfconscious way by designers by adoption of existing design conventions.” (Lindley & Sørensen 2007/8, see also: Lindley, Nacke & Sørensen 2007). In games, the same kind of objects can be used several times but can have different meanings (e.g. a door can be a non functional door or lead somewhere else, certain items can be taken or moved in some contexts but not in others) and these different meanings also create different affordances or disadvantages for the player in their context. Objects can therefore shape different associations and motivations. In visual semiotics there is a difference between what you see and what meanings or associations the perception gives. A _denotative_ description or analysis of a view in a 3D virtual world is the description of the iconography, syntax and outer context of that field of view. The _connotative_ description of a view could be its “cultural associations” (Cornell et al. 1999), and within the outer circle of this we can add our private associations that constitute our idiosyncratic subjective relation to the sign. These semiotic elements create different meanings and interaction patterns in the form of gameplay and provide terms for the systematic analytical description of the object of investigation.

_Psychophysiology_ or _Psychophysiological Science_ is, according to Cacioppo, Tassinary & Bernston (2000), “the scientific study of social, psychological, and behavioral phenomena as related to and revealed through physiological principles and events in functional organisms”. It can also be explained as “the branch of psychology that is concerned with the physiological bases of psychological processes”\(^4\). The operation of the eye and its objects of gaze are measurable using eye gaze tracking, one of many physiological measurement technologies that can be used to study game interaction. In addition, Brain activation can be measured with Electroencephalography and Functional magnetic resonance imaging (EEG, fMRI), Heart rate variability (HRV) can be measured with Heart Rate Watch, galvanic skin response (GSR), with sensors and face muscles activation with electrodes called Electromyography (EMG). However, the project reported here is based upon eyetracking.

The study of gameplay requires many different theories serving different purposes. Our interest within the Game and Media Arts Laboratory is to better understand how game

designs motivate and affect players, especially in terms of cognitive outcomes of gameplay and for inspiring the development of innovative designs. Cognitive accounts of gameplay are most relevant since they focus on skills learned via gameplay, which may transfer to other contexts. Visual semiotics is of interest for making the meanings produced by interacting with a 3D game world explicit.

There is a specific problem in research when we use textual descriptions to account for interaction with graphics and sound in these contexts. “There is a critical interrelationship between available taxonomies and the ontologies that they embody on one hand, and what is perceived among classifiable phenomena in the world on the other, as captured in the well known Sapir-Whorf hypothesis that our language conditions our thoughts, and hence the reality in which we live (Lindley, Nacke & Sennersen 2007).” We have all been brought up, generation after generation, with text as the most important language to learn for communicating to express ourselves. A linguistic focus in our lives tends to block our abilities to express ourselves in other ways than through words. The linguistically determined categorization of things perceived is there from when we are born. However, there are things perceived beyond what is determined linguistically, and these have meanings that do not conform to linguistic description (Lindley, Nacke & Sennersen 2007). For example, the visual perception of movement in sweeping a game camera through an arc of 45 degrees conveys far greater information in dynamic perception than a verbal description conveyed in the same amount of time about the changing scene. The linguistic and the visual are different semiotic systems that do not share all of the same meanings. The eyetracking methodology provides a method that may help to bridge this semiotic gap, producing characterizations that may be expressed in text but are based upon detailed observation of visual cognitive behavior.

The First Person Shooter Game

When playing a digital 3D PC game, there is a game running on a PC, a screen where the graphics are displayed, and a player interacting with the game. The interaction is typically through vision, sound and motor input. Gameplay is the activity of a player when interacting with and/or navigating through the in-game 3D objects within the game world. A game consists of different layers: its engine, media assets, logic and rules. The game engine is responsible for rendering the image and sound media as a dynamic process determined by player interaction in relation to the mechanics of the game constituting its logic and the game rules. Game rules and logic implement the state changes for game challenges, such as combating opponents, solving puzzles, finding items, etc., that are at the core of the gameplay experience for the player.

The licentiate addresses behaviors in a single-player First Person Shooter (FPS) environment. For baseline cognitive studies the single-player FPS is a good starting point for three reasons:
1. A concentration on ongoing tasks, otherwise the player (character) risks being killed.
2. High pace of action, few long pauses if any.
3. No social phenomena to complicate ongoing tasks.

The stimulus games used in the study are limited to combat and navigation/movement challenges. Additional game challenges may be added in ongoing studies so that changes in gameplay behaviors due to the additional challenges can be compared with this baseline.

The pace of gameplay is especially important in gameplay in action games, which have rapidly changing camera angles and views that may capture a variety of objects within a short time. There is no text, not even transcription, which can catch up to the pace of information in a way that can give it a comprehensive rendering.

**Theoretical Framework**

The main aim of the project described here is to understand externally measurable actions in relation to visual cognition and correlated with aspects of the subjective experience of the player, and to develop plausible models of the cognitive processes and structures underlying these phenomena. This immediately raises the questions of what it means to ‘understand’ and what represents a plausible model. In our articles “What does it mean to understand gameplay?” (Lindley, Nacke & Sennersten, 2007), “An Innovation-Oriented Game Design Meta-Model Integrating Industry, Research and Artistic Design Practices” (Lindley & Sennersten, 2007/8), “Game Play Schemas: From Player Analysis to Adaptive Game Mechanics” (Lindley & Sennersten, 2008) and “A Cognitive Framework for the Analysis of Game Play” (Lindley & Sennersten, 2006) we discuss how to frame gameplay related knowledge and how epistemology, cognition, and learning processes can provide a foundation for the investigation of different aspects of gameplay and their interrelationships. The aim is to have a critical and reflexive view upon the epistemological questions raised in these kinds of studies, and to develop a position that is not too naïve while nevertheless making sufficient explicit assumptions and adopting premises that allow empirical research to be conducted. The section will by no means attempt to go into depth in this epistemological jungle, but will make a brief overview to contextualize the theory and method, especially in relation to the constitution of objects, which are in central focus in the methods and theory used.

**Object(ive) Knowledge**

Epistemology is by definition “the part of philosophy that is about the study of how we know things”\(^5\). The nature of knowledge, how it is acquired, and its relation to concepts such as truth, belief and justification, are central (and still unresolved) questions in the long history of philosophy.

Virtual game objects are a core foundation of the knowledge learned by the player in virtual game worlds, because graphically represented virtual objects are the primary focus of player attention and the carriers of essential meaning in gameplay, as objects of perception, as action affordances, as the objects of action, and as game pieces in the definition and execution of game challenges. “Games designs include specifications of game objects that are implemented as mathematical object models, as media representations, and, on this basis, presented as objects of perception for a game player. In a sense this is an engineering epistemology, and this must confront the epistemic processes of the player. That is, the game object is a point of intersection between an externally perceived and constructed world (that of the engineering process of the game as understood by its developers) and the subjective experience of the player.” (Lindley, 2008, private communication).

An object by definition is “A noun 1. Something perceptible by one or more of the senses, especially by vision or touch; a material thing. 2. A focus of attention, feeling, thought, or action: an object of contempt. 3. The purpose, aim, or goal of a specific action or effort: the object of the game. 4. Grammar a. A noun, pronoun, or noun phrase that receives or is affected by the action of a verb within a sentence. b. A noun or substantive governed by a preposition. 5. Philosophy Something intelligible or perceptible by the mind. 6. Computer Science A discrete item that can be selected and maneuvered, such as an onscreen graphic. In object-oriented programming, objects include data and the procedures necessary to operate on that data.” An object can in a more summary form be identified as “A thing, an entity, or a being”7. To consider the role of objects in how knowledge is shaped, what knowledge we acquire while playing and what is perceived as being knowledge, it is useful to start out with a short historical review of the concept of an object and how this is fundamentally different in different epistemologies.

Metaphysical realism is “the belief in a reality that is completely ontologically independent of our conceptual schemes, linguistic practices, beliefs etc.”8 where truth in this case consists in a belief’s correspondence to reality. The questions of the nature of beliefs and of what (objects) exist in the world belong to the field of ontology. In metaphysics ontology is a central branch of investigation concerning what types of things there are in the world and what relations these things have to one another. Regarding objectivity and subjectivity, the metaphysician tries to clarify how people understand the world through concepts in existence, objecthood, property, space, time, causality and possibility. Ontology in metaphysics is “studies of being and existence that include the definition and classification of entities, physical or mental, the nature of their properties, and the nature of change”. Metaphysics is by definition “the branch of philosophy that investigates principles of reality transcending those of any particular science, traditionally, cosmology and ontology. It is

---


also concerned with explaining the ultimate nature of being and the world". There are now philosophical alternatives to this approach (and its associated method), some of which will be reviewed below. Computer games allow for any theory of ontology or reality to be simulated, and in a virtual game reality or a virtual game world the game designer shapes the relations between the ‘types of things’ (in-game objects).

Any theory of ontology presupposes an epistemology, since it is impossible to have anything to say about objects unless one has a background theory of how it is possible to know anything at all. Around the time of Renaissance, there were two main epistemological positions that dominated philosophy: Descarte’s rationalism and Hume’s empiricism. Empiricism says mainly that all our knowledge is based upon experience, while rationalism is formed upon rational insight (to ‘see’ something implicitly being true, i.e. based upon intuition and/or deduction, and innate knowledge). Descartes’ view was “experience served as a necessary function in activating our ability to think. But this is very different from claiming that our clear and distinct ideas could be given by experience. The latter is impossible, he thought, because our clear and distinct ideas are far richer in content than our fragmentary experience” (Aune, 1970). Hume on the other hand postulated that we gain knowledge from raw experience via perception through a general class named impressions, which includes feelings, emotions, volitions and desires. The other class is ideas formed by thoughts which were shaped from those impressions. Hume spoke of house or tree perceptions and meant that we all identify images and ideas by reference to the public objects they represent, independently if we see (perceived object) or may seem to be seeing something (mental object).

In the 1970’s, contemporary empiricism was further developed by Mill and Wittgenstein. The new empiricists introduced the verification principle to clarify subjective knowledge and understanding by relating it to what is publicly observable. These empiricists views became known as philosophical behaviourism (Aune, 1970). This is related to phenomenalism, which does not include physical objects as existing things but as phenomena or sensory stimuli. These epistemologies suggest two interesting views of game objects: the concept of an object shaped internally, in a person, isolated from an external world at the moment and the comprehension of an object, perceived by a person, shaped as experiencing it in the external world, the virtual game world in this case. The question is that of the degree to which we create the object mentally, as a mental simulation, or construct a representation of the object while interacting with it, an external virtual simulation, or if we do both, then how are these cross correlated?

There are many arguments for and against the roles of mental derivation and empirical observation, and the interdependencies between these. Pragmatism developed perhaps in frustration with the difficulty of trying to resolve these viewpoints analytically. Pragmatism is by definition “1. Philosophy A movement consisting of varying but associated theories, originally developed by Charles S. Peirce and William James and distinguished by the doctrine that the meaning of an idea or a proposition lies in its observable practical

11 http://plato.stanford.edu/entries/epistemology/
consequences. 2. A practical, matter-of-fact way of approaching or assessing situations or of solving problems.12. Pierce, founder of pragmatism, takes us further to assume that knowledge and action are two separate spheres in what “might be termed an ecological account of knowledge: inquiry is construed as a means by which organisms can get a grip on their environment.”13. What does real mean? The labels ‘Real’ and ‘True’ are labels having a function in inquiry. This is contrast with realists, ideal platonists, who acknowledge an external world which must be dealt with where objectivity or objective truth is more or less defined as “mind-independent” or “true in certain times, places or people.”15.

I see subjective and objective truth to be valid, where objectivity is equated with measurability, and my research aims to understand how these different kinds of knowledge hang together, and whether and how one aspect dominates the other. Pierce saw logic as a branch of semiotics, that he was a founder of, and defined an object as “anything that we can think, i.e. anything we can talk about.”16. In my research I am mainly concerned with visual objects as nouns. There are people playing, players, with individual minds interacting and acknowledging an external digital game world. Understanding the object and conceptualizing it in ways that account for subjectivity and objectivity is at the heart of many philosophical questions, and is a major question for cognitive science, raising difficult questions about the ways that we construct concepts of external reality and subjectivity from within subjectivity and without any external reference point from which to observe being. Heidegger questioned the whole of western philosophy all the way back to Plato, arguing that former philosophers had misunderstood what it meant for something to be. Heidegger wanted to retrace the meaning of being. Heidegger shaped a new concept that brought together the subject and the object in “being-in-the-world”, this concept being a replacement for the terms subject, object, consciousness and world.

The heritage of ideas from phenomenology and Heidegger to critical theory and structuralism (not discussed in detail here) are reactions to the modern spread of positivism “an essentially antiphilosophical philosophy based on empirical methods of the natural sciences” (Kearney, 1986). To take other methods into account we can gain a fuller understanding: “Instead of having one-dimensional surface observation, these modern methods open up other dimensions of meaning concealed behind the empirical

manifestation of things: the intentional activities of consciousness (phenomenology); the historical strategies of domination and liberation (critical theory); the unconscious codes of language (structuralism)” (Kearney, 1986). The player and the meaning of in-game things/objects and challenges to the player can have a diversity of different “ways of meaning”. A challenge for phenomenology, structuralism and critical theory is their lack of empirical methodology. These perspectives add dimensions of meaning to the interpretation of phenomena, but their claims cannot be tested for validity, reliability or the ability to make predictions. Adopting positivist methods adds the ability to generate predictive instrumental knowledge in the context of these approaches. Phenomenology, structuralism and critical theory are of great value for understanding the meaning of positivistic results. The virtual game world is an abstract mathematical model of an ideological physical world “The focus in an ideological world is on individual values viewed by groups.”17. Digital democracy aims to bring people and ideas together and these abstract mathematical models of worlds do not exclude any critique. In virtual game worlds we have propositional knowledge which includes both “knowledge-that as opposed to knowledge-how”.18 This section has described the epistemological position upon which this research is based. However, these philosophical epistemologies do not address all of the knowledge involved in gameplay. Epistemology has traditionally been more interested in “knowledge-that” than “knowledge-how” and this becomes tricky when both forms of knowledge are included in gameplay. Polanyi, a polymath19, started in the 30’s to articulate his opposition towards positivism and argued that it failed in recognizing the part which had to do with personal commitment and tacit knowing, also called personal knowledge that plays an important part in science but also in gameplay. Tacit knowledge has been described as “knowledge-how” (as opposed to “know-what” [facts], “know-why” [science] and “know-who” [networking]). It involves learning and skills but not in a way that can be written down. The knowledge of how to ride a bike is an example: one cannot learn to ride a bike by reading a textbook; it takes personal experimentation and practice to gain the necessary skills”20. In the case of visual semiotics and gameplay “Tacit knowledge is hard to articulate with formal language (hard, but not necessarily) impossible). It contains subjective insights, intuitions, and hunches. Before tacit knowledge can be communicated, it must be converted into words, models, or numbers that can be understood”21. A game has to be played and not read about and all tacit knowledge young people have today about playing and gameplay is more or less informal knowledge. Tacit knowledge can be explored using methods from cognitive science and psychology. Hence a cognitive approach provides positivistic methods in a comprehensive form highly suited to studying gameplay. The next section elaborates upon the cognitive perspective that forms the theoretical framework for this research.

Cognition, Visual Attention and Schema Theory

Cognitive Science seeks its knowledge and understanding of cognitive functions based as far as possible upon scientific methods. Scientific methods are based on observation to find empirical and measurable evidence that is subject to specific principles of reasoning that can be repeated and show a sustainable pattern over time with different people or phenomena. For cognitive science, scientific methods are not shaped in isolation, but "many other methodologies also contribute. A hallmark of cognitive science is its interdisciplinary approach. It results from the efforts of researchers working in a wide array of fields. These include philosophy, psychology, linguistics, artificial intelligence, robotics, and neuroscience" (Friedenberg & Silverman, 2006). Hence in cognitive science there are different explanations for the same phenomena depending on the specific approach.

This study has mainly a cognitive and cognitive psychological approach. Psychology arose in the late 19th century and cognitive science in the 1950’s. Psychology is defined as “the scientific study of mind and behavior. Psychology uses the scientific method as a means of gaining valid knowledge” (Friedenberg & Silverman, 2006). The studies in psychology focus on internal mental events such as perception, memory, attention, problem solving, reasoning, language, and visual imagery, as well as higher level observed behaviors like walking, talking and running. A central question in psychology is ‘How does the mind explain what we do?’ and a scientific answer can infer how the mind processes information the way it does based upon observations of the action performed. The term “cognition” is also used in a wider sense to mean the act of knowing or knowledge, and may be interpreted in a social or cultural sense to describe the development of knowledge and concepts within a group that culminate in both thought and action. Before the 1950’s, behaviorism (positivism) had been dominant in psychology and behaviorists ignored mental phenomena because they thought it was too difficult to define or measure. Hence they observed only external behaviors. The behaviorist approach has been much criticized, especially in the case of more extreme in form stimuli-response experiments. After the 1950’s, cognitive psychology included and studied knowledge representation and use in human beings, a subject omitted from earlier behaviourism. Cognitive psychology wanted and wants to understand how we represent, process and store information. The conceptualization of the human information processes are important and abstract process models used to map out how these processes occur; the models are tools to gain understanding. These process models have also been criticized by later connectionists; they stress lower level learning in their models, which are based upon the functioning of neural networks. Both of these approaches have been incorporated in the field of artificial intelligence (AI), which focuses on using these models to implement intelligence in machines. The intelligence in machines aims to reach or exceed human capacity in reasoning, knowledge, planning, learning, communication, and perception. The AI in NPC’s, which are machine played characters in a game, are also intended to achieve believability in certain human and dramatic ways.

A critical issue in computer game design is the design of the interaction functions for players, an aspect of interaction design which is frequently based upon principles from cognitive
science. Interaction design concerns “designing interactive products to support the way people communicate and interact in their everyday and working lives” (Sharp, Rogers & Preece, 2007). To consider how interaction design principles might apply to game design, it is useful to first summarise some ideas about game and game interaction design. Adams and Rollings (2007) describe how “Game designer Sid Meier’s famous definition is ‘a series of interesting choices’. Another designer, Dino Dini, defines [game design] as ‘interaction that entertains’. Our definition hinges on challenges and actions ... The challenges are any task set that is nontrivial to accomplish and the rules specify what actions the player might take to overcome the challenges and achieve the goal of the game”.

Applying concepts from cognition-based interaction design to the design of gameplay, Lindley and Sennersten (2006) suggest that “Game design features include built in game tasks, [that may be] modelled by a process of game task analysis, that ground the in-game meaning of design features”. Models in interaction design “are typically abstracted from a theory coming from a contributing discipline, e.g. psychology, that can be directly applied to interaction design. For example, Norman (1988) developed a number of models of user interaction based on theories of cognitive processing arising out of cognitive science, that were intended to explain the way users interacted with interactive technologies. These include his cyclical seven stages of action model—that describes how users move from their plans to executing physical actions they need to perform to achieve them, to evaluating the outcome of their actions with respect to their goals” (Sharp, Rogers & Preece, 2007).

The theory of action by Norman (1986) describes what users do step-by-step while interacting with an interface. The steps are: 1. Establish a goal, 2. form an intention, 3. specify an action sequence, 4. execute an action, 5. perceive the system state, 6. interpret the state and 7. evaluate the system state with respect to current goals and intentions. Of course, a system user or a game player does not necessarily carry out these 7 stages in a strict sequence while playing, but such a model can be built based upon the player’s retrospective comments on how they played, and perhaps in more detail by an external observer collecting data during play. Lindley & Sennersten (2008) adapt this kind of model to describe the decision processes underlying gameplay in a form that may be called a gameplay competence model: “1. decode the audiovisual sensory and perceptual information delivered by the game media (e.g. the computer screen and speakers) into the apprehension of a local situation within the synthesized game world (or game space), 2. Evaluate this understanding of the local in-game situation in terms of the overall objectives of play, current goals and tasks, the state of the player character within the game (e.g. capabilities, health and other statistics), and anticipation of various rewards of playing the game, 3. To make decisions about which in-game tactics and action(s) to perform next, based upon the perceived situation and its evaluation, and 4. Perform action(s) based upon competence in interaction mechanics and semantics”. This sequence, which is an established model for robotic action control, may be summarised as: sense > model > evaluate > plan > act, or the SMPA model.
Creating a model of cognitive task performance is may be referred to as Cognitive Task Analysis (CTA) (Lindley and Sennersten, 2006). The resulting task model is essentially the same as a model of a cognitive schema, where a schema is an algorithm for completing a particular goal or subgoal (Schank and Abelson, 1977). Mandler (1984) asks what schemas for stories, scripts, and scenes have in common: “Superficially they concern widely different aspects of our experience. Stories are literary expressions that we read or hear; they often refer to times long past or to imaginary worlds. Scripts, in psychological parlance, represent the familiar, everyday events which fill our daily lives: trips to the grocery store, getting up in the morning, going to work - the routines of our workaday world. Scenes ... represent places, the rooms and streets and buildings in which our daily routines take place. In spite of the variety of experiences that stories, scripts, and scenes represent, when we examine the structure of these domains we find that they have much in common and result in common types of psychological processing. The commonalities reside in the fact that all are represented in the human mind by related schematic forms of organization” (Mandler, 1984). Schemas are an abstraction representing some of our knowledge about objects as organized into categorical (taxonomic) systems. They also include knowledge of how to act to achieve specific outcomes, based upon the presence and properties of objects.

Applying schemas to the study of gameplay, “Analysis is conducted using a theoretical framework providing an account the features of game designs, of player cognition, and of play interaction relating design features to cognitive processes. This comparative analysis is used as a basis for validating or adapting models of hypothetical cognitive schemas underlying gameplay, or to modify the overall theoretical framework if those modifications are found to better account for observed play patterns” (Lindley & Sennersten, 2006). Mandler describes different kinds of schema structures; matrix structure, serial structure, schematic structure (event schemas and scene schemas) and story structure.

Lindley, Nacke & Sennersten (2007) discuss gameplay schemas, where “A gameplay schema is ... the structure and algorithm determining the management of attentional and other cognitive, perceptual and motor resources required to realize the task involved in gameplay”. Cognitive schemas underlying gameplay are “cognitive structures that link declarative (or factual) and procedural (or performative) knowledge together with other cognitive resources (such as memory, attention, perception, etc.) in patterns that facilitate the manifestation of appropriate actions within a context” (Lindley & Sennersten, 2006). From a game designer's perspective “Schema descriptions can also be used to explore the effectiveness of a game design in realizing designers’ intentions” (Lindley & Sennersten 2008).

This licentiate does not present any specific schema models. However, it is concerned with collecting observational data from gameplay, specifically addressing visual attention, as a foundation for later schema modeling (in ongoing research) that will associate perceptual criteria with action decisions.
Perception is "the process by which we gather information from the outside world via the senses and interpret that information" (Friedenberg & Silverman, 2006). Visual attention is a cognitive process selectively concentrating on a certain aspect in the environment while ignoring others. When playing a game, the player distributes their gaze while playing to attend certain objects while performing in-game tasks. The gaze distribution is task related. In the 1920's, von Helmholtz, also called Mr. Where, suggested that eye movements or visual attention had a tendency to wander to new objects as directed by will, or, as he wrote in his Treatise on Physiological Optics, “We let our eyes roam continually over the visual field, because that is the only way we can see as distinctly as possible all the individual parts of the field in turn” (Duchowski, 2007). James, a later colleague of Helmholtz, believed attention was directed upon what the object was in itself, the meaning of the object or the expectation of it. A classic study by Yarbus (1967) showed sequential viewing patterns over particular areas in an image which were significantly different depending on what particular question (and hence visual search task) a participant was asked before viewing the images. Tasks are therefore of great concern when interpreting sequential viewing patterns. Noton and Stark (1971) showed that predictable viewing patterns remain even if participants were not asked any questions at all.

In play interaction with a 3D game environment, cognitive processes are operating constantly and in relation to the overall task and game challenges. Visual attention directs the player to interact with both active and passive objects. Investigating visual attention during gameplay may reveal how players perform actions and how they solve problems in the game context.

Lindley & Sennersten (2008) describe play experience as a learning process and how the learning can be divided into three phases “1. learning interaction mechanics, that is, the basic motor operations required to operate (for example) a keyboard and mouse in a largely unconscious way, 2. learning interaction semantics, that is, the simple associative mappings from keyboard and mouse operations to in-game actions (and meta-game actions, such as setting play options, or loading and saving game states), and 3. learning gameplay competence, that is, how to select and perform in-game actions in the context of a current game state in a way that supports progress within a game.” The lowest common denominator in the game world is the object, the external object, and how subjectivity and objectivity correlate to this object.

The game object is a virtual one presented to the player visually. As such, the research project described in this licentiate takes the understanding of visual cognition within gameplay as its primary task.
Methodology development

From Heidegger’s concept of ‘being-in-the-world’ applied to phenomenology in the physical world, I want to be able to say how human beings relate to a digital game world during gameplay, especially using methods from cognitive science and with a focus on visual cognition. The first question that arises is how we can catch somebody’s visual attention while playing a 3D computer game. Traditionally this has been possible with eye tracking, but this has primarily been used with 2D stimuli. When considering a dynamic 3D digital environment the player can move the camera, potentially up to 360 degrees around the three orthogonal axes of a 3D world. Sennersten (2004) accomplished this by retrospectively watching captured video files of the game sessions with superimposed gaze data (see also Sennersten and Lindley, 2006). Established methods for statistically analyzing 2D gaze data do not apply for dynamic images such as those of a 3D computer game, so if 2D tools are used, the objects of gaze must be transcribed manually frame-by-frame, which is a very time-consuming process. Instead, we want to be able to log objects falling under the gaze in each frame automatically and in real time.

Real time logging of gaze objects in gameplay represents a new method that has several advantages, including saving analysis time and thereby facilitating studies involving many players that would be impractical if based upon manual transcription (Sennersten and Lindley, 2006). Statistically significant data sets allow for more reliable discrimination of correlations between the self-reported subjective states and experiences of players and externally collected data, e.g. about gaze behavior and the objects of visual attention.

The solution developed to achieve real time automated gaze object logging has been to connect an eyetracker with a 3D game engine (Sennersten et al. 2007, Sennersten & Lindley 2008). The game engine used to develop the initial integrated system is the HiFi engine developed by FOI, Linköping, Sweden. This engine was chosen for the full availability of its software code. The interconnection means that the eyetracker and the game engine can communicate every update or simulation tick of the engine while the game is played. The gaze target on the screen is represented by an x and y coordinate captured by the eyetracker every 20ms. The gaze coordinate is sent to the game engine and from this an inverted ray is traced into the 3D world. The first collision box of a virtual game object intercepted by the traced ray is logged by the game engine. This function works the same as processing a mouse click on a graphical object.

The eyetracker uses data from two eyes but generates an average gaze target from these, so the game engine uses the joint (x,y) coordinate determine the object of sight. The study uses a stationary eyetracker. In Virtual Reality (VR) systems, stereoscopic vision is created by two individual pictures, one for each eye, each having a different perspective and therefore slightly different image of the virtual world. VR generally uses goggles or glasses when simulating reality based on stereoscopic vision. To understand the affects of playing mass market games, and for studying play behavior and learning outcomes of these games, means studying game interaction using stationary computers or consoles and televisions, with interaction via conventional mouse and keyboard, or a game console, respectively.
The integrated eyetracker/game engine system needed to be verified in order to establish not only that the basic software worked correctly, but also to gain an understanding of its accuracy in space and time (Sennersten et al. 2007, Sennersten & Lindley, 2008). The verification study verified the object logging by comparison to a second source, the captured video screen file. The verification study was successfully completed before any ongoing experiments were carried out.

Validation of the system goes beyond verification. While verification seeks to establish the bounds of correct operation for the system, validation demonstrates its usefulness in its intended application domain. In this case, validation means showing that the system and its associated methodology can be used to demonstrate interesting results in the study of visual attention in computer gameplay. The validation experiment is described in the next section.

The overall methodology for experiments is based upon methods from cognitive psychology. Specific details of the experimental procedure are documented in Sennersten and Lindley (2008) and Sennersten et al (2008). As described by Cronbach and Meehl (1955), cognition experiments are designed to address requirements of:

*Construct validity:* how can we claim that a measurement is valid? Construct validity refers to the degree to which experimental variables can match a certain theory and the distinctions within the theory in order to form a coherent whole.

*Reliability:* the quality and repeatability of measurements. Repeatability is assessed among different groups independently of each other.

*Predictability:* the ability to forecast a system state and how it operates quantitatively or qualitatively

The empirical goal of the current study is to be able to measure gameplay activity in order to validate theories of how gameplay is generated in relation to specific tasks and settings and how design choices make a difference to the player and what these differences mean. In pinning down how a game task is performed I seek to understand the formal, informal and skill-based knowledge involved and then correlate this with responses to questions and interviews to identify if any patterns occur or if there are any detectable tendencies of particular kinds. Cognitive theories of computer gameplay are currently under development, so our research is focused upon basic questions of finding effective theoretical constructs and empirical methods. Hence the research focuses upon construct validation.

**Empirical Experiments & Results**

The purpose of developing an empirical data acquisition system for player gaze distribution is to readily obtain quantitative data that can be cross-correlated with qualitative data
obtained by questionnaires and interviews to see if there are any interesting relations (i.e. correlations, differences and/or similarities) between observed visual attention behavior and subjectively reported experience during gameplay. Testing specific hypotheses regarding visual attention and behavior require the definition of relevant variables that provide the basis for the implementation of a purpose-specific game level to function as an experimental stimulus. Real time logging is then run during game sessions while participants are playing, and the resulting log data is analysed to determine whether or not there is, or the degree of, support for the experimental hypotheses.

It is not necessarily the case that quantitative and qualitative data reveal the same behaviors, as discovered by Sennersten (2004). In that study, players said in the questionnaire that their visual attention was directed towards the representation of the face region of NPCs (non-player characters) when entering a room to rescue a hostage, but the measured data revealed that 90% of the participants actually directed their gaze towards the hands. A plausible hypothesis for why they looked more at the hands is that most threats are executed from that region.

In the experiment reported in this thesis (Sennersten et al, 2008), real time gaze data in a 3D setting was collected and both pre- and post-questionnaires were completed before and after each game session. Three hypotheses were investigated in this study:

**H1:** the visually static diagonal shape and position of the gun graphic, which does not have a function in aiming the weapon, cues and directs visual attention within the screen, preceding and/or independently of the attentional demands of the player’s in-game task(s).

**H2:** gaze is firstly directed upon an opponent prior to shooting the opponent.

**H3:** for otherwise equivalent opponents, the closest opponent will be targeted first.

The binary variables in the experiment refer to the position of each enemy when encountered in the stimulus level, including are side (left or right), distance (near or far) and gun position (left, right). The data reveal that hypotheses H1 is not supported, H2 is correct approximately 90% of the time: players do look first on their opponent and then shoot, while only 10% shoot first and look afterwards. The latter case occurs in particular when the workload on the player is higher (due to higher perceptual loading), suggesting a greater cognitive demand upon maintaining situation awareness in relation to the relatively clear cognitive task of shooting an enemy. Further experiments will be required to determine in more detail how (and if) this apparent use of peripheral vision changes systematically in relation to changing perceptual loading. H3 is found to be correct in most enemy encounters, with an average of 77% of cases where the closest opponent is be looked at first, irrespectively of the left or right gun graphic position.

To be able to say anything about the hypotheses in this experiment, it was necessary to start out quite simply without too many variables. To understand a game design requires
first minimising complexity and establishing a baseline. A baseline is by definition “a basic standard or level; guideline; to establish a baseline for future studies”\textsuperscript{22}. This baseline is important in all experiments so that it is possible to determine the effects of ongoing systematic changes in design features. This research also starts out with single player game modes because multiplayer games are too complex for the present state of the research.

**Practical notes on the licentiate format**

Part I of this licentiate presents a summary overview of the research project up to the completion of this licentiate thesis. Part II of the thesis, “Theory, context & method”, presents the papers concerning the theoretical framework for this study. This section includes both the relevant theories for interpreting gameplay and also the game design process as a context for the research. Part III of the thesis, “Methodology Development”, describes via the papers the main contribution of this licentiate in developing a technological method for supporting the scientific methodology for investigating visual attention in gameplay. This part describes the integration of an eyetracker with a computer game engine. It also describes in detail the procedure and results of the verification study of the integrated system. Finally, Part IV of the thesis, “Empirical Experiments and Results”, describes the first experiment conducted using the methodology and underlying system. This section represents the validation of the integrated system and its associated research methodology by demonstrating the use of the system to empirically investigate specific experimental hypotheses. The content of all parts except Part I is presented in the form of compiled research publications. Finally, an Epilog presents some concluding comments to the thesis.

\textsuperscript{22} [http://dictionary.reference.com/browse/baseline](http://dictionary.reference.com/browse/baseline), accessed 2008-06-03
PART II
THEORY, CONTEXT & METHOD


What does it mean to understand gameplay?\textsuperscript{2,3}

Craig A. Lindley, Lennart Nacke and Charlotte Sennersten
Blekinge Institute of Technology, Karlshamn, Sweden
craig.lindley@bth.se, lennart.nacke@bth.se, charlotte.sennersten@hgo.se

Abstract
Understanding gameplay requires a consideration of basic epistemological questions about the nature of understanding. Grounded in a tradition of philosophical hermeneutics, it is possible to approach the understanding of gameplay as a matter of generating mappings to explanatory frameworks in alternative interpretation paradigms. It is especially useful to consider gameplay from perspectives of cognitive science, semiotics, consciousness studies and aesthetics. Each of these approaches provides a different but compatible perspective on understanding play. Integrating these perspectives without losing their differences provides a comprehensive theoretical framework for play analysis.

*Key words:* Gameplay, epistemology, cognition, consciousness, semiotics, aesthetics.

1. Introduction
The gameplay experience is at the centre of interest for gamers, and whatever the other qualities of a computer game, its commercial success will ultimately be founded in the quality of its gameplay. Interest in studying play within game research has increased dramatically over recent years, resulting in an increasing diversity in approaches to research on gameplay, with little consideration to date of how this diversity may be unified. To this end, this paper outlines a general epistemological stance from which it is possible to pursue research into play in a way that illuminates play in diverse but coherently integrated ways. We do not present a detailed philosophical analysis, but summarise the foundations of our epistemological position. We then go on to describe a number of orientations towards understanding play, based upon this epistemological perspective. These orientations generally amount to broad interpretation paradigms, based upon cognitive science, semiotics, consciousness studies and aesthetics. Each of these perspectives is summarized, considering the forms of understanding that they may provide, together with a consideration of their complementarity and the different but compatible dimensions of play and games that they can inform.

2. A Poststructural Epistemology for Understanding Gameplay
Central questions leading to an understanding of gameplay include: i) why do players play, ii) how do players play, iii) what are the effects of game play upon players, and how do the answers to these questions depend upon both iv) game design features and v) identifiable discriminations between different players? Each of these central questions may be regarded

\textsuperscript{2,3} Published in the proceedings of *Ludic Engagement Designs for All (LEDA)*, 28-30 November 2007, Aalborg University Esbjerg, Denmark.
as requiring a taxonomy for articulating any kind of answers, taxonomies describing respectively: i) player motivations and rewards, ii) play behaviour, iii) the consequences and affects of play, iv) game design features and v) player types. Taxonomies provide named conceptual (and linguistic) categories providing meaningful differentiations within and among play motivations/rewards, play behavior, consequences/affects, game designs, and players.

In effect, these five conceptual/linguistic taxonomical areas map out the conceptual space to be understood. Any specific taxonomy represents an ontological theory, i.e. a theory of what may exist. E.g. the widely used taxonomy dividing player types into achievers, socialisers, explorers and grievers [1] embodies a theory that these categories represent actual groups of different types of players.

*Describing* any particular example of gameplay or its components amounts to categorizing it according to available (and selected) taxonomies. There is a critical interrelationship between available taxonomies and the ontologies that they embody on one hand, and what is perceived among classifiable phenomena in the world on the other, as captured in the well known *Sapir-Whorf hypothesis* that our language conditions our thoughts, and hence the reality in which we live [34]. The Whorf hypothesis has tended to prioritise (verbal and textual) linguistic forms over other carriers of meaning, but here the point is extended to consider non-linguistic systems of signification and symbolization, such as sounds/music, visual creations (pictures, films), games, etc.. In this perspective, our designed artifacts condition our perceptions of reality, indeed largely construct our reality beyond the bounds of those artifacts. We also consider the Lacanian point [15], that linguistic structures encode power relationships, to be extended to non-linguistic semiotic and symbolic systems; an important function of many contemporary artistic practices is to foreground these relationships.

Alternative taxonomies from this perspective provide mechanisms for restructuring reality. Hence *understanding* a phenomenon is not merely wrapping it up within some descriptive system, since different descriptive systems encode different realities. This does not, however, mean that reality can be arbitrarily changed simply by adopting a different description scheme. We take the ontological foundational state to be a highly structured phenomenology within which linguistic constructs such as subject/object, self/other, nature/culture, personal/social, etc. are heavily involved with all other elements of that phenomenology[4], the whole evolving in what may internally be formed as a continuous movement towards increasing coherence. Arbitrary adoption of descriptive schemes may promote incoherence, in this sense being incompatible with other layers and complexities of the teleological project of phenomenology. Principled adoption of descriptive schemes

---

[24] The following discussion makes many references to higher structured elements of phenomenology, the other being a prime example. Without further justification we take these constructs to serve a higher coherence within phenomenology, as opposed, for instance, to solipsism.
can be regarded as a systematic evolution of phenomenology towards greater coherence, even if sometimes at the expense of a short term reduction in coherence as conceptual structures must be broken in order to be superceded by new concepts within a more coherent overall structure.

Within this view, it may be asked what constitutes a good explanation, or a better explanation, or any kind of explanation at all. In general it may be said that an explanation consists of an expression of one or more associations among the terms (and hence concepts) of one or more taxonomies or their instances. A theory may be understood as “a coherent group of general propositions used as principles of explanation for a class of phenomena” [25]. The propositions of a theory are expressed as associations among taxonomical categories. Hence a theory can provide an explanation for the observed behavior, features, etc. of instances of taxonomical categories.

The associations between the terms of a theory or an explanation may take many different forms, including mathematical functions, existential dependencies, causal relationships, and logical relationships (e.g. inferences). Scientific theories and explanations rely upon these kinds of associations. Many other kinds of explanations and theories are possible, e.g. of personal preference or disposition (“I went to the bar because I didn’t feel like being alone”), appeal to authority (“don’t do x because the priest says it is wrong”), assumed teleology (“go the the secret door under the bridge to find an artifact” that will help you to defeat the level boss, explaining how to finish a level), etc.. Associations between categories and/or their instances may also serve functions other than explanations, such as metaphor, poetry and analogy.

A theory is valuable for what it may explain and value is relevant to purposes. The nature of scientific theories has been a topic of extensive philosophical investigation. Perhaps the most widely known and used characterizations of science are those of Karl Popper [24] and Thomas Kuhn [14]. Popper stipulates that a good scientific theory should be falsifiable, i.e. it should have implications that in principle may be disconfirmed by observations; if not, then it is no better than any alternative unfalsifiable theory, providing no basis for choice and having no explanatory power. Conversely, a falsifiable theory that has not been falsified is never proven to be absolutely true, since it is always in principle possible that an observation may one day falsify it.

Popper’s criteria for scientific validity are widely accepted, but suffer from a number of fundamental problems. Firstly, falsification makes sense as a criterion, but by no means has this principle been strictly followed in the history of science. Many major theoretical changes have occurred despite their initial falsification at their time of formulation; e.g. the many steps from an Earth-centred view of the universe to our contemporary cosmology (e.g. see [8]). Based upon these historical considerations, Kuhn developed a conception of science founded upon the concept of paradigms, where a scientific paradigm is a body of theory based upon a set of foundational assumptions. In Kuhn’s model, science proceeds
in three phases: a pre-paradigmatic phase, where the foundational assumptions are yet to be articulated and established, a normal science phase, where scientists operate in terms of a set of established assumptions (e.g. the laws of Newtonian physics), developing details but without questioning the foundational assumptions, and a revolutionary phase, where a set of new assumptions takes over from the established canon (e.g. relativistic physics superseding Newtonian physics). A revolution occurs in response to the accumulation of falsifications within the old paradigm, representing a transition to a new set of foundational assumptions in a new paradigm that better accounts for observations. This does not in itself diminish the importance of falsifiability in science, but acknowledges a kind of cultural inertia in research communities that may lead to them being slow to respond to falsification. As Kuhn notes, some established scientists may never change the paradigms in which they work, new paradigms often being most actively pursued by new researchers.

A new paradigm has its own language and conceptual categories, representing a move into a new and different ontology. A new paradigm also typically involves new methods of investigation, and may be formed as a consequence of falsifying data obtained by new observational technologies, which in turn require new experimental methodologies. Bearing in mind the discussion above about the role of language in our perception and construction of reality, this model of scientific progress needs to be understood not as a process of gaining better knowledge of an absolutely knowable objective reality, but as an evolution in the conceptualization of the world in interrelationship with methods of action (research methodologies) and observation technologies, moving towards a greater overall coherence in how the world is conceptualized. Part of this coherence is achieved by the scientific criteria of reliability and predictability. *Reliability* increases the value of a theory in that its predictions are preserved across certain systematic changes of context (a study replicated in another place and time should have comparable results, if aspects of the structure of the study are preserved). *Predictability* increases the instrumental value of a theory in that it not only successfully abstracts from observed phenomena but can be used to predict the behavior of future phenomena under specifiable conditions.

These factors are described by Cronbach and Meehl [5], who also characterize the overall working coherence of a body of theory together with its methods of empirical investigation as construct validity. Construct validity can be described as the degree to which inferences can legitimately be made from the operationalizations (i.e. practical experiments) in a study to the theoretical constructs on which those operationalizations are based. Cronbach and Meehl describe how a nomological network provides construct validity, where a nomological network includes “the theoretical framework for what you are trying to measure, an empirical framework for how you are going to measure it, and specification of the linkages among and between these two frameworks”. It is a necessary condition for a construct (i.e. a distinction, concept or theory) to be scientifically admissible that it occur within a nomological (or ‘lawful’) net having at least some laws involving observables. We can extend this point with the hermeneutic observation that the conceptual framework and language of a nomology include its empirical and observational concepts, constructs and procedures.
A scientific approach can therefore be regarded as one seeking nomologies having maximum coherence among abstract theories, observations and actions in the world, leading to construct validity, predictive validity and reliability of knowledge embodied in the nomology. This characterization does not rely upon any kind of absolutist or positivistic interpretations of knowledge, and also does not imply arbitrary relativism. All conscious elements of our experience are interpreted, and as such form a generalized “text” of experience. Then, as stated by Paul Ricouer [26], “If it is true that there is always more than one way of construing a text, it is not true that all interpretations are equal.... The text is a limited field of possible constructions. The logic of validation allows us to move between the two limits of dogmatism and skepticism. It is always possible to argue against an interpretation, to confront interpretations, to arbitrate between them and to seek for an agreement, even if this agreement remains beyond our reach”.

There are, of course, paradigms of interpretation that are not scientific, and are unscientific to the degree to which they are incoherent, or lack falsifiability, construct validity, predictive validity and/or reliability. Science is fundamentally concerned with instrumental effectiveness, the development of theories that provide foundations for maximizing the predictability of future observations and the outcomes of action. Scientific approaches may be severely limited in making aspects of phenomenology coherent that are not concerned with optimizing instrumental action. Even in the sphere of action, science may provide the best understood ways of acting efficiently, but generally provides no guidelines for why to act or what to create when we are involved in the manufacture of our reality by our actions and their consequences.

3. Interpretation Paradigms for Understanding Gameplay

Against this very high level background, we now consider how we approach the task of understanding gameplay. The first implication of the epistemological ideas above is that to understand a phenomenon is to map it within an interpretation paradigm, and that such a paradigm will shift both due to its own internal dynamics and its interrelationships with other interpretation paradigms with which it may be compatible to varying degrees. Clearly there is no single dominant paradigm in contemporary studies of gameplay. Nor should there be: the crucial meta-theoretical task is to map out different approaches and develop a high level perspective for interrelating them as a higher level, coherent body of work, each specific approach formulating different kinds of questions and developing its own nomologies for creating answers. Overall coherence lies in recognizing the different purposes and forms of knowledge that the various paradigms involve.

The question then is how to structure such a map, a question that will no doubt have many exploratory answers over time, with no final answer as long as the study of gameplay continues to develop. Interpretation paradigms are also highly complex, often nested, being vaguely bounded, with numerous overlaps, congruencies and contradictions, and change over time. There is an ambiguity too in relation to fields of enquiry, where different fields may be identified by their own typical motivating questions, foundational theories, histories and methods, but where within a field there may be competing paradigms of
interpretation of their subjects of interest. For example, gameplay may be investigated within fields including: i) cultural studies of games (i.e. game studies), ii) research in the technical platforms and infrastructure, facilitating technologies (e.g. new rendering techniques, AI, physics simulation, network technologies, etc.) and production tools for games, iii) design research into new forms of games and play, iv) scientific studies of games, play and affects upon players, v) research in the philosophy of games and play, and vi) social scientific studies of players, communities and the role and nature of games and play in society.

Our own interest is in research leading to a better understanding of how game designs motivate and affect players (especially in terms of cognitive outcomes), the development of innovative designs and understanding their affects. This is a matter of increasing the self-consciousness of design processes (explained at length in [19]) with a view to facilitating game design in general and enhancing innovation in design concepts. In developing a theoretical framework to support these goals, we are focusing upon the integration of what we regard to be among the most interesting and/or promising approaches or paradigms for understanding gameplay, with a view to creating an overall coherent paradigm addressing different levels of gameplay functionality. The approaches that we are seeking to integrate include: cognitive science, semiotics/symbology, aesthetics and consciousness studies. In many ways cognitive science can include all of the other approaches, but there are limitations in the cognitive perspective. This is elaborated in the sections below.

4. Cognitive Science

Cognitive science seeks an understanding of cognitive functions based as far as possible upon scientific methods. A more scientific understanding of the motivations, rewards and affects of gameplay can be a valuable contribution to understanding how design features function, and also lead to clearer ways of creating designs that more readily achieve specific targeted affects (e.g. for training, therapy, or emphasizing specific elements of entertainment experiences). Gameplay is, amongst other things, a cognitive skill acquisition and performance process, whether played for entertainment or other purposes.

Cognitive science is a broad, multidisciplinary field, characterized by Thagard [33] as: “the interdisciplinary study of mind and intelligence, embracing philosophy, psychology, artificial intelligence, neuroscience, linguistics, and anthropology”. Within this broad scope, it is possible to seek those approaches and theories that appear to have the most to say about gameplay. In this spirit, Lindley and Sennersten [18] present a theory of the underlying cognitive systems involved in gameplay based upon schema theory and attention theory. Schemas are cognitive structures that link declarative (or factual) and procedural (or performative) knowledge together in patterns that facilitate comprehension and the manifestation of appropriate actions within a context. While the taxonomical structures of semantic or declarative memory can be modeled in terms of object classes together with associated features and arranged in subclass/superclass hierarchies, the elements of schemas are associated by observed contiguity, sequencing and grouping in space and/or time ([20].
Schemas can refer to declarative knowledge and taxonomical types with their features and relationships, and integrate these with decision processes. Schemas include scripts for the understanding and enacting of behavioural patterns and routines, a classic example being Shank and Abelson's [27] example of the restaurant script that includes a structure of elements for entering a restaurant, sitting down, ordering food, eating, conversing, paying the bill and leaving (etc.). Scripts, as structures used for both comprehension and behaviour generation, represent a structure of cognitive functions that may include cognitive resources, perceptual interpretations and preconditions, decision processes, attention management and responsive motor actions. Story schemas, are patterns representing a structure of understandable elements that must occur to make stories comprehensible. The presence of story schemas in the cognitive systems of storytellers, listeners, readers or viewers of stories allow stories to be told and to be comprehended, including the inference of missing information. If a story deviates too far from a known schema, it will not be perceived as a coherent story. Script and story schemas are concerned with structures of both space and time, while scenes are schemas representing spatial structures, such as the layout of a house, a picture or an area of a city.

While schemas have been interpreted in many different ways, here a game play schema is understood as a cognitive structure for orchestrating the various cognitive resources required to generate motor outputs of gameplay in response to the ongoing perception of an unfolding game. A gameplay schema is therefore the structure and algorithm determining the management of attentional and other cognitive, perceptual and motor resources required to realise the tasks involved in game play. Examples of types of gameplay schemas described by Lindley and Sennersten [18] include story scripts for understanding high level narrative structures designed into games, and scripts for the combative engagement of an enemy, exploring a labyrinth, interacting with a trader non-player character, and negotiating and carrying out quests.

Attention theory provides an account of the energetic resources available to cognition, together with principles for the distribution of energy (or attention) to the cognitive resources that use (or manifest) it. Attention theory addresses issues of attentional focus, management of attention (including attentional selection), and the allocation of cognitive resources to cognitive tasks. Ongoing research is addressing the question of the detailed operation of attentional mechanisms, including questions such as the degree to which attentional capacity is specific to specific cognitive resources (or modes) or sharable among resources according to demand, and the stage of processing of perceptual information at which perceptual information is selected for attentional priority. Schemas can be regarded as mechanisms or algorithms that, amongst other functions, determine the allocation of attention to cognitive tasks.

In the context of gameplay, attention and the operation of gameplay schemas is driven by hierarchical goals that set tasks for a player. Goals include those intended by designers and those created by players as allowed by a game design. A hierarchical decomposition of game
play goals might at a high level include the completion of a game, which decomposes into the subgoals of finishing each of its levels, each of which in turn decomposes into goals of completing a series of game challenges.

We hypothesise that this hierarchical goal structure is mirrored in a hierarchical structure of schemas within a player’s cognitive system, where a schema is an algorithm for completing a particular goal or subgoal. This schema structure is fundamental to many aspects of the pleasures and motivating factors behind play [18]. These include pleasures of:

- **effectance**, which is a basic feeling of empowerment created when an action of a player results in a response from the game system [13]. The cause-effect relationships underlying effectance are a fundamental premise of goal-oriented schemas for action.

- **closures** at different hierarchical levels (as described by Holopainen and Meyers, [10]), where a closure is interpreted here as the completion of the algorithm constituted by a play schema. Closures may involve completion of expected outcomes and resolution of dramatic tensions, corresponding to the completion of cycles of suspense and relief identified by Klimmt [13]. A distinction must be made here between the intrinsic pleasures of schema completion and more complex emotional experiences and rewards due to fictional identification within the game world (see the point below regarding episodes).

- **achievement** of in-game tasks, which is rewarding due to the displacement of a player’s identity into their character [10], this being a matter of *imaginative immersion* as described by Ermi and Mäyrä [6]. Achievement-oriented reward is a more specific form of reward than mere closure, since it is associated with the completion of schemas by the achievement of specific goals.

- more complex forms of enjoyment in game tasks regarded as *episodes* [13] following from imaginative displacement into the game world. Enjoyment within episodes may include the excitement of possible action, pleasures of curiosity and discovery, the pleasures of experiencing negative emotions of suspense followed by the transference of arousal to an ecstatic experience when the challenge creating the anxiety of suspense is overcome, and enhanced self-esteem. Schemas offer greater discrimination of the pleasures involved in episodes by allowing different forms of episodes to be modelled as different schema patterns having a complex substructure with corresponding emotional effects (e.g. different scripts for solving mysteries, combat, exploration, trading and quest negotiation).

- **escape** to an alternative reality provided by the fictional world represented by a game [13] and facilitated by imaginative displacement. Players have the pleasure of being able to experience new objects, actions, social interactions and experiences at no
risk. These vicarious experiences can help players to cope with felt frustrations and deficiencies in their everyday lives, a process both of catharsis and of perception and feelings of increased competence and relevance. Schemas for stories facilitate displacement, while many additional schema forms provide the foundations for comprehension of the events within the fictional world and provide mechanisms for projection of the player’s sense of self into the fiction.

– achievement of a sense of flow [4] in gameplay, this being a state at the boundaries between engagement and immersion, of being totally absorbed in meeting a constantly unfolding challenge. We hypothesise that the flow state is associated with attentional demand, in particular occurring when schema execution demands attentional resources above a level that would result in player boredom and below a level that would result in excessive difficulty and consequent frustration.

Schema theory therefore has the potential to provide both an explanation of the decision and operational processes underlying gameplay and an explanation of the detailed reward and motivation factors behind play. Validating this potential requires detailed study of play resulting in the development of empirically validated hypotheses about the detailed structure and functionality of gameplay schemas, for individual players and across groups of players.

5. Semiotics/Symbology

Semiotics is the study of signs, signification and sign systems, including how meaning is constructed and understood. The semiotics of computer game design is fundamental to design practice, but is generally involved in an unsconscious way by designers by the adoption of existing design conventions. In this context it refers to the way that perceptual and active elements are organized to convey meanings (e.g. see [21], [7], [17]).

Schema models provide very useful abstractions representing the cognitive ‘algorithms’ involved in gameplay, irrespectively of the more difficult philosophical question of how a schema model is to be interpreted ontologically. Hence schema models can be used to analyse gameplay without resolving the linguistically conditioned questions of whether or how the abstraction of a schema model exists as mind and/or matter within the cognitive system of a player, or whether the mind/matter distinction is an unnecessary construct. A schema model is an abstract representation of mappings from perception through decision processes to action. However, such a mapping is far from exhausting the possible meanings of perceptual and action phenomena. While meanings can be mapped from many perspectives, we might say that a specifically semiotic perspective considers the internal logic of signifying systems, independently of the underlying cognitive mechanisms of signification (which would provide a more specific cognitive semiotics); it may be that a variety of different underlying cognitive architectures could operate the same given semiotic system. Moreover, within human cognitive systems, very much is a matter of the internal structure of knowledge, beliefs, concepts, etc., for which cognitive mechanisms are
very generic, and most of what is interesting is within the abstract relational structures of knowledge and concepts. Semiotic analysis focuses upon this level of structure. From this perspective it is possible to analyse the semiotic processes involved in gameplay in terms of a hierarchy of semiotic and symbolic complexity, for example and in order of increasing symbolic level and complexity:

- basic cognitive functions and emotional rewards associated with their operation
- task-oriented cognitive mechanisms
- semiotic and symbolic constitution of the self
- semiotic and symbolic constitution of immediate social relationships
- semiotic and symbolic constitution of subcultural contexts
- semiotic and symbolic constitution of general cultural contexts

Cognition provides the foundations for semiosis, while semiosis provides the rich content of dynamic conscious experience.

6. Consciousness

As discussed earlier in this paper, structured phenomenological consciousness provides the foundations for all epistemological projects. For the understanding of gameplay, the conscious experience of play provides the connecting nexus between observed behavior and reported experience. A scientific nomology for the explanation of play should ideally include methods and taxonomies for interrelating observed play, psychophysiological measurements, inferred schemas, and subjective reports of the conscious experience of play.

The subjective experience of gameplay, however, is not straightforward to describe. Many studies on games have developed taxonomical terms for different modalities of gameplay experience, aside from general emotional experiences (e.g. see [16]), including: emotional affection [31], immersion ([2], [6]), frustration [9], presence ([32], [11]), and flow ([12], [3], [30], [29], [23]). Concepts such as immersion, presence, engagement and flow are often unclear, with no consensus about their precise meaning or how to clearly distinguish one state from another under all circumstances: Novak et al. [22] describe 16 studies conducted between 1977 and 1996 based upon different concepts of flow.

This diversity in the use of basic terms reflects the fundamental difficulty of assessing states of consciousness. Ideally a nomology should provide empirical measures by which different measurement results can be correlated with differences between states of consciousness captured by descriptive taxonomical categories. However, it is intrinsically difficult to devise taxonomies for reporting subjective states, since validation must rely upon verbal reports, the quality of subjective experience being externally inaccessible to any kind of direct observation. Consciousness cannot, however, be neglected, since it is the foundation of gameplay experiences and epistemic projects that seek to understand them.
7. Aesthetics

The final approach that we sketch here is that which we refer to as aesthetics. Here aesthetics is understood to mean a concern with design form from the perspective of design tradition, formal innovation and the perspectives of contemporary arts practice (this is not aesthetics in a classical sense of a 'love of beauty'). Aesthetics in this sense may be regarded from the perspective of cognition or of semiotics, but we separate the approach as one that is best pursued by understanding a design in relation to other contemporary and historical designs. This is intended to capture the sense of direct theory as described by Small [28] in the case of avant garde cinema. In the case of gameplay, this means the understanding of a game obtained via playing it and in relation to the experience of playing other games, with their similarities and differences. This is the situated understanding of the meaning of a game, manifest in the experience of play and experienced as part of an ongoing discourse of play, a discourse about play conducted via play.

This approach to understanding play does not rely upon cognitive or semiotic analysis, but it does rely upon creating and playing games. Including such a perspective is critical in order not to lose sight of the constructed nature of other interpretation frames, and the alternative possibility of understanding play in the language of play itself as a unique mode of apprehension facilitated by games as a (reality constructing) form. This is to return to the foundations of our epistemological position: all epistemology begins within a highly structured phenomenology. The starting point for understanding gameplay is the experience of the highly structured phenomenology of the play experience (or of the observation of others playing), before its mappings to the other structures in phenomenology that provide paradigms for understanding via interpretative mapping. Our knowledge and understanding of play is grounded in this direct experience, which ultimately may not need to be mapped to any other interpretation frame in order to be understood. Art speaks of art by art.

8. Conclusion

The framework for approaching the analysis of gameplay presented in this paper seeks to understand games from several complementary perspectives within an overall epistemology founded upon philosophical hermeneutics. Cognitive science incorporates empirical, scientific methods for studying gameplay, including consideration of neurophysiological foundations and providing explanations in terms of cognitive operations, functions, structures and affects. Many higher level cognitive functions, however, are highly distributed and represent complex learned functionality from a cognitive perspective. In these areas, it is more productive to examine abstract conceptual and linguistic operations than to focus upon more basic cognitive capabilities. This leads to a consideration of knowledge in itself, independently of its cognitive substrates. Semiotics then takes over, examining the meanings of perceptual and cognitive experiences as more abstract systems. Signifiers are understood as belonging to systems of signification, with the ability to express different and new meanings by combinatorial operations upon signifiers, e.g. the production of linguistic
statements by the applications of rules and terms from languages, or the assembly of the
signifiers of interaction affordances within computer games into a coherent play concept.
Semiotic principles apply to non-linguistic signifiers, such as any meaningful elements of
perception and encompassing designed form in all media. The study of meaning from a
semiotic perspective does not need to consider underlying cognitive foundations, and in
many (or even most) cases there are few adequate comprehensive cognitive theories of
semiosis.

Studies from either a cognitive or semiotic perspective tend to be articulated in verbal and
written languages. However, non-linguistic media function in direct semiotic ways that
may not always be accountable in cognitive and/or linguistic semiotic terms without the
loss of some meaning. Undertaking a discourse aimed at creating deeper understandings
of media (in the current case, of games and gameplay) without the loss of media-specific
meaning must then be conducted in the languages of non-linguistic media themselves, e.g.
by exploring game forms that provide a commentary upon game form itself. In this process,
a process that we here characterize as an active project of creative aesthetics, discourse
may also proceed via cross-media production, e.g. making games that comment upon
dance, dance performances that comment upon games, etc.. Aesthetic discourse may be
examined from cognitive, semiotic and linguistic perspectives, but cannot be exhaustively
appropriated into these modes of interpretation without potential loss of meaning. A game
that is played has meanings that escape descriptions of play, just as a dance that is danced
always has meanings different from descriptions of the dance. Finally, consciousness studies
is concerned with the ground of all other experiences, the starting point and the end point
of all projects of understanding. All of the other interpretation paradigms serve, articulate
and (in-)form consciousness. There are no vantage points outside of consciousness from
which it can directly be compared with its models, and every new model or variation of a
model changes consciousness itself. Hence the process of knowledge generation must be
regarded as a process of expanding consciousness, the various interpretation paradigms
representing alternate and complementary dimensions of expansion.

Acknowledgements
This paper has been written in the context of the FUGA (FuN of GAming) EU research
project. We thank our FUGA colleagues for many stimulating inputs to this work.
References

An Innovation-Oriented Game Design Meta-Model
Integrating Industry, Research and Artistic Design Practices

Craig A. Lindley and Charlotte C. Sennersen
Game Design, Cognition and Artificial Intelligence Research Group,
Gotland University College
Cramergatan 3, SE-621 57, Sweden
+46 498 299 765
(craig.lindley.charlotte.sennersen)@hgo.se

Abstract

The distinction between implicit and unselfconscious design cultures on one hand and explicit, self-conscious design cultures on the other provides a principle for interrelating a variety of game design approaches within a coherent game design meta-model. The design approaches in order of increasing design self-consciousness include implicit design, ‘cookbook’ design methods, taxonomy and ontology-based game design, theory-driven design and formalist reflexive design. Implicit design proceeds by copying existing examples of game designs, while ‘cookbook’ methods generalize from examples to create lists of design heuristics. Taxonomy and ontology-based game design is based upon more systematic models of the types, features, elements, structure and properties of games. The theory-driven level involves the design of game systems to facilitate gameplay motivated by cognitive, scientific and/or rhetorical theories of game affect and functionality, or incorporating technical innovations providing the basis for new game mechanics and experiences. The formalist level represents the application of reflexive contemporary artistic perspectives to games, resulting in games that reflect upon, question or reveal game form. In placing these different approaches within a hierarchy of increasing self-consciousness of design practices, the meta-model provides a clear account of the roles of research and artistic methods in game design and innovation, providing a foundation for more explicit design decision making and game education curriculum development integrated with higher-level research.

Keywords: game design, methodology, pedagogy, innovation, research, art.

Introduction

This paper presents a meta-model describing and interrelating different approaches to and methodologies for game design. Motivations and questions behind the development of this meta-model include the need for more systematic, advanced pedagogical methods for teaching game design within specialized game education programs. A good pedagogical framework must be able to relate games to the history of other media, to be able to

account for the relationships between viewing games as an industrial design activity on one hand, and as a contemporary artistic medium on the other. Games can be designed not only for entertainment or artistic purposes, but also for specific rhetorical purposes (e.g. advergaming), or to embody specific theoretical principles aimed at achieving particular affects within players (e.g. for therapy or to facilitate targeted modes of immersion). A high level view of game design needs to integrate these different design contexts and motivations. It is also necessary, specifically from a pedagogical perspective, to develop approaches to game design that facilitate the evolution of game forms beyond games that are currently available, in order to create new modes of experience, to address new markets and applications, and to deepen our cultural understanding of game form and function. A pedagogical framework for game design education must also foster creativity, leading students to be able to think ‘outside the box’, as well as integrating education, industrial design practice and formal research as it relates to design.

The meta-model presented in this paper is proposed as one way of meeting these requirements. The paper first presents a foundational distinction articulated by Alexander (1970) between self-conscious and unselconscious design cultures. Based upon this distinction in mind, we then present the overall meta-model that integrates implicit game design methods, with what we call ‘cook-book’ design approaches, game design patterns and game ontologies, theory-driven design and formalist design. Each of these approaches is then described in more detail, including discussion of its relationships with the other design methods. The meta-model has been used as the foundation for an advanced game design course, and some of the resulting design concepts are described. The purpose of the meta-model is certainly not to provide any kind of substitute for the creativity of designers. Rather, it is a tool for facilitating, opening up and perhaps amplifying that creativity, based upon the general principle that representations provide amplification of human cognitive capacities, as described by Harth (1999). While the model is being used to facilitate pedagogical processes that encourage more creative design by novices and to speed up the development of design competence, it also clarifies the relationships between industrial game design practice and different forms of research, contributing to ongoing discussions about the relationship between research and industrial practice in game development.

Game Design
Before going into the detailed discussion of design methodology, it is useful to present a preliminary representation of the general objects, or outcomes, of game design, as shown in Figure 1, based upon the driving concept of the gameplay experience, a consideration of what remains the same when a game is realized in different ways, and what design elements change in different implementations of ‘the same’ game. These are the various elements of form representing the final outcome of design and that shape and constitute the designed artifact. In this model, game play is at the center since this represents the core and overall goal of game design, being the design of the space of possible interactive experiences for players. This may be more or less open, from restricting the player to very
limited possibilities (e.g. in a simple game like Tic-Tac-Toe) to very open games having a lot of scope for players themselves to shape their own experiences (e.g. live-action role-playing games, or larps). In all cases, the scope for players to vary their game play within the constraints of a particular game system is always at least implicitly a design decision. Of course, players may use a game system in ways that do not constitute playing within the system (e.g. a game to see who can throw a computer game CD into a hat!), but the game design itself includes, implicitly or explicitly, a scope beyond which play no longer takes place within the designed game. It is the scope of play intended by designers that drives the design process.

Driven by the target game play, the next priority in game design is the design of a logical game system and elements needed to support a space of designed play experiences. Hence the target game play provides a requirement specification driving design of the logical game system and elements. The logical game system and elements include:

- game rules that specify legal moves that players may make, the consequences of moves, win/lose criteria, etc.

- game objects are the things within the game that are referred to by the rules and may be manipulated by the player and/or game system; objects may be active or passive, and their specification can include attributes relevant to game play and referred to by the rules and game system

![Diagram of Elements of Designed Game Form](image)

Figure 1. Elements of Designed Game Form.

- a game space, also referred to by the rules and defining a logical space within which play takes place

51
the game system which specifies how all of these elements are orchestrated
together to constitute a complete game; the game system includes procedures for
the execution of game moves and the manipulation of game objects, according to
the rules, within the logical game space, and in terms of the media used to realize
a game.

The game system might include media-specific procedures, but the rest of the logical game
system and elements will often be transferable across different media. For example, sports
games specify particular rules, game objects (such as bats and balls), player roles and a
game space (such as courts or fields); however, there are many computer versions of sports
games where these elements are intact, although the system of play and the nature of the
play experience are different. To the extent that the system and the play experiences differ
across different media, there are examples of different games, but to the extent that the
game rules, objects, player roles and game space are the same, then they are the same game:
the identity of a game follows from the (variable) scope of elements taken as constituting
that identity.

Once the logical game system and elements are specified, it is possible to undertake the
design of the game media components. This may include 2D and 3D graphics, animations,
video, audio, lighting, costumes, sets or stages, interfaces, technology and infrastructure. For
computer games, costumes, sets or stages are virtual, and the game space may be organised
into game levels within an overall virtual game world. For physically staged games, this
will be physical elements, such as the costumes of larpers or the uniforms of sports players.
Within this layer design techniques from established design fields may be applied, but
always in terms of meeting the gameplay-driven requirements of the game system. Hence
established methods address the design of media elements, while game design as such is
concerned with the inner core of gameplay and the design of the logical game elements
and system required to facilitate gameplay.

Outside the areas of artifact design, game design also has a bearing upon the context
of play. For example, a board game designed for a context such as a family home makes
assumptions about what is possible within that context (e.g. a clear table around which
six adults may sit); if the context does not accommodate those assumptions (e.g. no room
for a larger table) then the context must be modified if the game is to be played (e.g. other
furniture is moved out). Hence the game design implies or specifies requirements for
features within the context of play, amounting to a degree of context design that may be
satisfied either by selecting a suitable context or modifying a context to render it suitable.
Contextual requirements are well understood for computer games and actively analysed by
the designers of console games. Contextual factors are a significant challenge to overcome
for the widespread commercialization of some new game forms, such as augmented reality
games or technology enhanced games; barriers here include cost, an unprepared market,
and the need for some kind of bootstrapping process by which increasing markets can drive
costs down. For this reason, contextual design can have a much greater impact for new game forms having poorly established or supported context requirements.

Self-Conscious and Unselfconscious Cultures

The American architect Christopher Alexander (1970) makes a useful distinction between implicit design within an unselfconscious design culture and explicit design within a self-conscious design culture. While Alexander is specifically interested in architecture, these distinctions will be applied here to processes of game design. This may be regarded by some as a controversial application of what may be seen as a rather dated distinction. Our answers to this are firstly, that the distinction provides a useful heuristic for interrelating different approaches to game design; as a heuristic it is a simplification, but one that we have found to be effective in practice for stimulating more creative design outcomes. Alexander’s distinction is a simple binary one. We have taken it further to distinguish degrees of self-consciousness in order to organize and interrelate what may be regarded as different approaches to design. This organization is just that, a way of approaching and regarding design perspectives. We do not claim that it is more or less correct than other ways of organizing and interrelating approaches to design may be, although we have not yet seen many other high level models, and we argue for the usefulness of this particular model. Secondly, in Architecture the academic discourse around design is well established and mature. In game design, it is not. It should not therefore be surprising that a metadiscussion about game design approaches should perhaps look back within the history of other design disciplines for distinctions that might be used in the early stages of a discourse that will later gain comparable sophistication.

Alexander characterizes an unselfconscious culture as one in which there is little thought about design as such and there are no general principles of design; rather, there is a tradition of right and wrong ways of doing things and practitioners learn to imitate by practice, the same form being learned over and over again. Creation involves the repetition of patterns of tradition because those are the only ones known. There is no particular interest in new or individual ideas, and there are no written records. Concepts and the language of self-criticism are too poorly developed within an implicit design culture to make significant critical discussion possible. A novice learns by very gradual exposure to the craft, being guided by sanctions, penalties, reinforcing smiles and frowns, etc.. Creation is based upon implicit (unmentioned) and specific principles of shape; unspoken rules, of high complexity, are not made explicit, but revealed through the correction of mistakes. This mode of creation is very typical of longstanding creative practices, such as those within traditional cultures for building houses or making artifacts of different kinds. Alexander (1970) characterizes the implicit design methods of an unselfconscious culture as methods that result in highly successful forms, but only if the rate of change of the functional context of creation is comparatively slow. Designs are then adapted to slowly changing contexts by a series of very small scale changes.

53
In many ways, at least until very recently, the commercial game industry has shown many of the features of an unselfconscious design culture as described by Alexander. This is especially the case for games having stable feature sets, comprising standard design features within game genres such as strategy games, first-person shooters and role-playing games. Would-be designers of such games have been faced with a bottom-up model of the road to professional design that begins with hard-core gaming. The gamer might then move on to modding and scripting as an indicator of commitment and nascent design talent. The entry point for a would-be designer within a game company might then be as a tester. After demonstrating some talent for testing, it might be possible to gain a position as a level designer. The career path then goes from level design to game designer within a team to becoming a lead designer for new games. All along the way expertise is developed largely by imitation, trial, correction and experience. There is little innovation involved within design practices throughout this process and the road to becoming a fully credible design specialist may take very many years to travel.

This model of breaking into the game industry applies not only to design roles, but also the development and producing roles (e.g. see http://archive.gamespy.com/articles/january03/education/day2/). While the model may work for comparatively stable game genres, it is not suitable under conditions where design demands/functions are evolving quickly, or when higher levels of innovation are required, such as when the market is bored with established forms, when a company or publisher wants to explore uncharted territory, or to keep up with and take advantage of changing technologies. It is also unsuitable when the training of designers must be accelerated, e.g. to keep up with the demands of an expanding industry. It is therefore not surprising that, especially over the past decade, there has been an increasing development of self-consciousness in game design.

Alexander (1970) characterizes a self-conscious culture as one in which form-making is undertaken by explicit and general (academic) rules and principles. Education is formalised, based upon instructions and teachers who train pupils, and novices learn much more rapidly based upon general principles. Teachers engage in a general process of trying to make design rules explicit, condensing knowledge that was once laboriously acquired through experience. Self-conscious cultures arise in circumstances where new purposes occur all the time and it is not enough to copy old patterns. In this situation design education is based upon explicit general principles of function, facilitating innovations and modifications as required, although the dynamic nature of the design context means that self-conscious design tends to lead overall to less good fit.

Alexander (1970) describes how as a self-conscious design culture develops further, change for its own sake becomes acceptable. Culture changes too rapidly for adaptation to keep up with it and factors sustaining equilibrium drop away. The master craftsman takes over the process of form-making and inventiveness becomes valued as a way of distinguishing craftsmen/artists, leading to the cultural perception of the designer as a star. Specialisation underlies the establishment of design academies, and the academies make principles explicit,
making them available for criticism and debate. Debate requires justification, leading to the formulation of general theories, principles and rules. Questioning leads to unrest, which leads to formal innovation and further self-consciousness.

Self-conscious design culture is concerned with both the design education of novices and explicit, self-conscious debate among established and experienced designers. One of the distinctions of experienced and expert designers (as with all forms of expertise) is an increasing implicitness of knowledge, with ongoing analytical processes oriented towards making that implicit knowledge more explicit. Hence explicit design knowledge accelerates and facilitates the ongoing development of expertise, but it is always very far from fully representing that expertise.

**Game Design Methods and Degrees of Self-Consciousness**

The distinction between unselfconscious or implicit design cultures and self-conscious or explicit design cultures provides a foundation for interrelating different methods of game design. Different methodological perspectives or approaches are described in terms of their degree of self-consciousness the following subsections.

**Implicit Game Design**

Game design within an unselfconscious design culture proceeds primarily by copying. As Alexander (1970) notes, this really amounts to selection rather than design. Highly conservative development cultures fall largely into this mode of operation. Within this culture, a design might be developed based upon a set of known examples, where the game design document, necessary as a social record of design decisions, really amounts to a list of features selected from a range of possibilities understood from past games within the tradition of the genre. For example, if a developer wishes to make a ‘fantasy RPG’, there are highly conventionalized precedents for combat, magic and trading (inventory) systems. A conventional combat system may provide precedents for character features, hit points, armor and attack values, together with rules for how these parameters are interrelated to generate outcomes from combat interactions. The fictional genres of fantasy, in literature, cinema and games, can provide predefined character archetypes, races, functions (fighter, magician, cleric, etc.) from which selections can also be made. Alternatively, new fictional elements may be introduced, such as a unique world with its own kinds of races and character classes, with the game mechanics being nevertheless selected from game genre conventions. In this case innovation is very much in the level of small scale but perhaps extensive features, such as the design of visual styles and graphics, design of specific weapons and armor, or particular novel character classes or races having different combinations or parameterizations of standard features and/or capabilities. Higher levels of innovation created by genre crossover still amount to a selection of features from established designs within genres.

The persistent popularity of genre productions makes implicit design within genre traditions a viable commercial option. The primary requirements for innovation include
the need to keep up with increasing technological capacities in target machines, although
the impact of this is most directly felt in the nature and requirements placed upon game
media assets (animation sequences, mesh models, textures, etc.). What the implicit culture
is not so good at dealing with are the rapid education of designers (it takes time to develop
an extensive experience of playing and then designing games within a genre), to create
new modes of experience within genres for perhaps an aging player base that is becoming
restless with the same modes of play, and for creating innovations in the basic form of game
mechanics for the sake of attracting new and different kinds of players.

‘Cook Book’ Game Design
Design ‘cookbooks’ are compilations of design ‘recipes’ consisting of rules, principles and
heuristics. Cookbook approaches represent the first step in making design knowledge
explicit and in making the design process self-conscious. A good example is Barwood’s
‘400 design rules’ project (Barwood, 2001, and Barwood and Falstein 2002; see also http:
//www.theinspiracy.com/400_so_far.htm). Examples of rules from Barwood’s collection
include: Maintain Level of Abstraction, Make Subgames, Let the Player Turn the Game
Off, Maintain Suspension of Disbelief, Differentiate Interactivity from Non-Interactivity,
Make the Game Fun for the Player, not the Designer or Computer, Provide an Enticing
Long Term Goal, etc.. Cookbook elements are a substantial part of many game design
publications (e.g. Rollings and Adams, 2003, Oxland, 2004, Salen and Zimmerman, 2004,

Cookbook approaches abstract from many specific examples of games to compile a superset
of design features, options and principles. Cookbook design principles may be used as a
foundation upon which more self-conscious approaches are founded, and many game
design handbooks present more theoretical material as a context or justification of basic
cookbook principles (a notable example being Salen and Zimmerman, 2004). What design
cookbooks do not address in any depth are questions such as why certain design rules work,
what it means to for them to work, what the inner motivations and rewards of gameplay
might be, or how to design games for which there are not well understood games that can
function as models to base design upon.

Game Taxonomies and Ontologies
The development of clear taxonomies and ontologies of game elements constitutes
another step in rigour, clarity and comprehensiveness in the next step in making
language clear the process of making game design self-conscious. A taxonomy can be
understood as a system of named and defined classes or categories and their subclass/
superclass relationships. An ontology can be understood as a taxonomy with the
addition of class properties and relations between classes. In this discussion the terms
tend to be used interchangeably, although in general an ontology provides a more
detailed description of the conceptual structure of a domain than a taxonomy does.
An ontology might be represented using i) a vocabulary of terms denoting ontological
concepts, ii) definitions of those terms, that may provide criteria of their applicability, and iii) a specification of how concepts are related, imposing structure on the domain and constraining the meanings of terms.

Within the general development of game design theory, increasing self-consciousness requires the development of game ontologies for discussing the forms and elements of games and raising the structure of the conceptual domain of games into greater awareness. Numerous proposals have been made for this, including the high level taxonomy proposed by Lindley (2003, 2005) that identifies basic distinctions between simulations, games and narratives as alternate formal systems being associated with respectively increasing time scales in the design process; simulations are concerned with modeling tick by tick (or frame-by-frame) changes, games with modeling player-controlled actions at intermediate time scales, and narratives being concerned with the largest scales of time structure. Lindley (2003) further distinguishes the orthogonal classification dimensions of fact/fiction representational functions and physical/virtual staging strategies for games. Aarseth et al (2003) propose a taxonomy based upon a variety of formal (i.e. non-narrative and non-representational) characteristics covering space (perspective, topography, environment), time (pace, representation, teleology), player structure, control (mutability, savability, determinism), and rules (topological, time-based, objective-based). Klabbers (2003) presents a taxonomy of game pragmatics, i.e. a taxonomy of the external functional application domains of games, game form and simulation, including business, administration, education, environment, health care, human services, international relations, military, religion, technology, human settlements and imaginary worlds. Foci of interest (including theory and methodology, instrumental design, research, training and education, and entertainment) are then broken down in a different dimension and themes (including competence, communication, knowledge/skills, management/organization, policy and fun) in another.

Björk and Holopainen (2005) present a taxonomy of the high level aspects of games, presented as a game component framework that includes: Holistic Components dealing with aspects of the game regarded as a whole (game instances, game sessions, and play sessions), Structural Components that are the basic parts of the game manipulated by the players and the system (including an interface, game elements, players, a game facilitator and game time), Boundary Components that limit the activities of a player of a game either by only allowing certain actions or by making some actions more rewarding than others (including rules, modes of play, goals and subgoals), and Temporal Components that describe the time flow of a game (including actions, events, closures and subclosures, end conditions and evaluation functions).

The classic work of Caillou (1958) presents a taxonomy of forms of play based upon an analysis of Latin terminology, including agon, based upon competition, alea, based upon chance, mimicry, based upon simulation and the kind of play associated with acting a role in a theatre production, and ilinx; based upon vertigo, “an attempt to momentarily destroy
the stability of perception and inflict a kind of voluptuous panic upon an otherwise lucid mind”. Caillios (1958) also discusses the distinction between *Paidia* as uncontrolled, free, improvised and ecstatic play, and *ludus*, which is play tightly bound up with arbitrary, mandatory and often tedious rules and conventions. Between paida and ludus there is a continuum between which there are degrees of variation, from total freedom to heavy but arbitrary constraint.

Game design patterns are another form of game ontology. The concept of game design patterns has been developed by Kreimeier (2002), Björk and Holopainen (2005) and Kirk (2005). Björk and Holopainen (2005) define game design patterns as “semiformal interdependent descriptions of commonly reoccurring parts of the design of a game that concern gameplay”. Game design patterns are essentially higher-level structures of game elements that might be described by component-oriented taxonomies, together with interaction patterns. Björk and Holopainen (2005) present 200 game design patterns, including the familiar patterns of Paper-Rock-Scissors, Save-Load Cycles, Enemies, Game World and Combat. The balance between interaction structure and other contents of game design patterns, as they have been articulated to date, varies. It is an issue of ongoing concern to validate the usefulness of the existing patterns within different contexts, and to further refine them or specify new patterns for purposes for which the currently identified patterns are not adequate.

Game taxonomies and ontologies are useful for *describing* game elements and design concepts in a way that is more systematic and comprehensive that cookbook compilations of design knowledge, also addressing structural features missing from cookbook approaches. However, they are not in themselves adequate for *explaining, justifying or motivating* design decisions. The next level in design self-consciousness must address these issues of *why* specific design choices are made. This requires the development of empirically validated theoretical perspectives by which game designs expressed according to suitable taxonomies and ontologies can be interpreted and/or motivated.

**Theory-Driven Game Design**

Theory-driven design refers to the next stage in the explicit and self-conscious development of designs in which design elements, principles and/or patterns are selected according to clear and conscious criteria. Those criteria are regarded as constituting or deriving from some form of motivational or interpretative theory. Relevant theories may include *rhetorical theories* of how a design can achieve changes in the beliefs, behaviors, consciousness or ways of perceiving of players, *scientific theories* about player motivations, the function of games and affects of play upon players, and *general theories*, which may be theories about any aspect of the form, structure, history, purpose or meaning of the world or things within the world.
Rhettoral Game Design

Rhettoral theories amount to theories about how a design can produce specific attitudinal, epistemic, behavioral or perceptual changes in players. Games designed from a rhetorical perspective include so-called ‘serious games’ (e.g. see http://www.seriousgames.org/) or third-party games, i.e. games designed to achieve purposes for some agency other than the players and the developers. Third party games include games that function as advertising, political games, social games, educational games and ideological games. There are many examples of rhetorical games. America’s Army (http://www.americasarmy.com/) is essentially an advergame commissioned by the US defence department and designed to convince players to join the US army. Howard Dean for Iowa (http://www.deanfor americagame.com/) is a political game supporting Howard Dean’s US presidential bid. Foreign Ground (http://www.gamespot.com/news/6137237.html) is an educational game designed for training defense personnel on peacekeeping missions. Many games have been developed for health education (http://www.gamesforhealth.org/) and for making various kinds of political points or statements (see, for example, http://www.watercoolergames.org/archives/cat_political_games.shtml).

Although many rhetorical games have been developed, the theory behind the rhetorical function of games is not yet very advanced. A deeper understanding of the rhetorical functions of games and game play requires deeper and more scientific or empirical study of game functions and player affects.

Scientific Game Design

Scientific theories may address many aspects and levels of game function and affects. This category is distinguished from the previous category of rhetorically motivated design in the adoption of scientific methodology in understanding game affects. Of course, rhetorical game designs could also base their design principles upon scientifically studied design effects, in which case rhetorical design and scientifically motivated design are the same thing. A number of studies of gameplay have investigated emotive issues such as game addiction (Fischer, 1994, Griffiths and Hunt, 1998, Salguero and Morán, 2002) and correlations between computer game play and violent behavior (Ballard and Weist, 1996, Griffiths, 1999, Anderson, 2004, and Smith, Lachlan and Tamborini, 2003). In order to more fully understand how gameplay can change players, and to support much more specifically targeted game design in terms of player affect, more detailed, fine-grained studies of psychophysiological and neurological responses to gameplay are required (e.g. Ravaja et al, 2005, Mathiak and Weber, 2005). The high level context for scientific studies might be regarded as the question of how player characteristics (personality, aptitudes, motivations) together with specific design features and play circumstances result in measurable and identifiable psychological and physiological changes during and perhaps following gameplay, where those changes might vary from very temporary changes to permanent changes. Cognitive, psychophysiological and neurological studies of gameplay hold the potential to reveal the details of cognitive and emotional processing that lay behind player engagement and immersion in gameplay, and unravel the uninformative concept of ‘fun’ into much
more specific factors of motivation, attention and cognitive task performance in relation to different patterns and characteristics of game design features.

Scientific theories of game engagement and affect can provide deeper foundations for designing the rhetorical functions of games. They can also allow games to be designed for various other targeted effects. For example, games have been found to function effectively in therapeutical applications, such as the treatment of phobias (Robillard et al, 2003). A deep understanding of the effects of gameplay upon players holds the potential for the design of games that achieve particular effects of cognitive reprogramming. Of course, there are ethical considerations in this. However, implicit design or design with limited self-consciousness holds the danger of achieving these kinds of effects in a completely unconscious way on the part of designers and players. Articulating a well-developed science of gaming moves gameplay effects into the foreground of consciousness for explicit critical analysis of game functions. This certainly does not mean that scientific theories of game design should only be used for third party or rhetorical functions, since those theories can also support more informed design of the principles and affects of entertainment products.

**Time Frames of Scientific and Technical Research**

A notable aspect of scientific research is that it can also be regarded in terms of levels of innovation, as depicted in Figure 2 (focused upon industrial and technical research in the case of the bottom two levels), analogous to the levels of self-consciousness involved in design innovation. Within this model:

*Basic research*, or blue-sky research, is the pursuit of new knowledge without any assumptions about what it might lead to. This is knowledge for its own sake. In general (but not always) basic research can be expected to have a long time frame to the development of clear results, e.g. 10+ years, with even longer times being required to generate practical applications founded upon these results. Research into the molecular foundations of neural processes falls into this category.

![Figure 2. Levels of Scientific Research and Development Innovation.](image-url)
Strategic research is the pursuit of new knowledge that might in principle have practical applications but without a precise view of the time scale or nature of the application. Strategic research will generally have a mid- to long-term time frame to the development of clear results or practical applications, e.g. 5 to 10 years. A project developing non-invasive methods for detecting brain states might fall into this category.

Applied research is knowledge developed with a specific objective in mind, particularly the conversion of existing knowledge into products, processes and technologies. Applied research will generally have a mid-term time frame to the development of practical applications, eg. 2 to 5 years. A project aiming to create a prototype system that detects player-controllable brain waves and feeds them into a game engine as an interface device might fall into this category (e.g. http://www.heroicsalmonleap.net/mlc/mindbalance/index.html).

Experimental development is work undertaken for the purpose of achieving technological advancement for the purpose of creating new, or improving existing, materials, devices, products or processes, including incremental improvements to these (http://www.ccra-adrc.gc.ca/taxcredit/sred/publications/recognizing-e.html). Experimental development projects are relatively short, typically being completed within one or two years or less. An example of this would be a project to develop a console controller interface to a skull cap containing electrodes as a product to integrate mind control with video gameplay.

Standard development is development by selection of standard solutions to well understood problems, requiring little to no innovation. This is the level not only of traditional crafts, but also of routine industrial production. Standard commercial game development falls within this category.

Within this model of forms of research, long term, basic research asks more fundamental questions, involves more risk and has potentially very high payoff, in some cases generating results that totally transform the basic assumptions of a scientific field. At the other extreme, standard industrial production operates at a level of highly standardized practice, involves little to no innovation, and incurs minimum risk. Applying this model to game research shows that potentially long time scales may be involved (e.g. ten years or more) before more significant research results are generated and fed into industrial game design practice.

Scientific theories can be understood to include technological research, such as research within computing and communications technologies. In this case technological innovations may support new modes of gameplay. Examples here are numerous, including games based upon mixed and augmented reality technologies (eg. Szalavári et al. 1998, Björk et al. 2001, Pickarski and Thomas, 2002, Magekuth et al, 2003, Magekuth et al, 2004), and games based upon modified game play due to the development of artificial intelligence methods.
for more effective characterization, dramatic interaction and emergent story construction (e.g. Cavazza and Charles, 2005, Mateas and Stern, 2002).

**Game Research in Relation to Autonomous Research Disciplines**
In considering the relationship of research to game design, distinctions may be made between research specifically directed at understanding game form, research within autonomous disciplines that is directed at games as an application area, and research within autonomous disciplines that is not specifically concerned with games, as depicted on Figure 3.

![Figure 3. Disciplinary structure of game research.](image)

Autonomous disciplines, such as the examples used in Figure 3 including human-computer interaction (HCI), artificial intelligence (AI), cultural studies, cognitive science, computer graphics, pedagogy and software engineering, have subcultures and processes that have no intrinsic dependency upon or interest in games. However, research within these disciplines may turn to games as an application area or object of study. In this case the specific methodologies and knowledge of those fields is applied to various questions arising from gaming and game design and development. Core game research, however, is concerned with game form as its first priority. These different areas interact. Core game research may derive models and principles from applied research from other disciplines, while those applied disciplines benefit from the deeper analysis of game form undertaken by core game research. Hence the input from research into game design may be highly indirect, generating results firstly within autonomous discipline areas that are then fed into research applied to games, which then feeds into the central analysis and articulation of game form. Also, autonomous research does not need to feed into industrial game design and development via core game research, but may flow directly into industrial game development. In fact, all computer games are based upon research in this way, using research results that provided the foundations for the technologies and communications infrastructure with which computer
games are implemented. At the time of writing, core game research is too young as an academic field to have had time to have much impact upon industrial game development, although this is likely to change as the field matures.

*General Theories Motivating Game Design*

General theories represent interpretation paradigms, or sets of basic assumptions about aspects of the world from which many other understandings may follow. An explanation that maps a phenomenon back to one or more of the basic assumptions or their implications within an interpretation paradigm constitutes an explanation within that paradigm. For example, identification of a player with a player character while playing a first-person shooter might be interpreted within a Freudian interpretation paradigm as the expression of unresolved Oedipal conflicts in which enemy monsters represent threatening aspects of the player’s (primal, symbolic understanding of their) father; then the Freudian-inspired designer might seek a game design that substitutes simplistic victory over the game boss with a more subtle process of identification with the father figure and transformation to a post-Oedipal psychodynamic motivation. Of course, the same phenomena can be explained by different interpretation paradigms in totally different ways, mapping them back to completely different foundational ideas. The interesting outcome may not be the theoretical justification in itself (which could be quite idiosyncratic), but the novel game concepts that result from it.

*Formal Reflexive Game Design*

Formal reflexive design focusses upon fundamental questions about the form of a creative medium; it is concerned with the production of artefacts that are about the medium within which they are produced, including conventions of production and interpretation for the medium in question. This approach has been a significant facet of modernist investigations of different media. Modernism has been a prominent cultural movement, particularly in the twentieth century, representing a radical change in the way different creative fields approach their work. This change has occurred within all established art forms and media, including painting and visual art, sculpture, literature, poetry, music, theatre, cinema, dance, and architecture. When cultural movements go through revolutions, genres tend to be attacked and mixed up, new genres are generated and old ones fade. These changes are often reactions against the prior cultural form, which typically has grown stale and repetitive. While the history of modernism is very complex, here it is possible to refer to a number of strong tendencies relevant to the game design meta-model:

- maximisation of self-consciousness and reflexivity in relation to a particular medium.

- questioning all aspects of the form and content of a medium.

- there is a movement away from representation, and away from or to disrupt conventional codes of representation.
- there is a strong focus upon pure form itself (e.g. line, colour, texture, material of the medium), frequently with a concern with the emotional, conscious and/or affective states induced by form (rather than by any denoted object). In the case of games this means a focus upon the essential nature of a game, and the relationships between the core game system and the media used to realise a game.

- a lot of modernist work has sought to answer the fundamental structuralist question: what is the medium? Since games have a tendency to be trans-medial (i.e. a particular game may be able to be realised using quite different media, e.g. as a board game, as a computer game, or as a game staged by people), the reflexive questions may be asked as to whether games can really be considered to be a medium, what may be gained or lost by considering them to be a medium, and if they are not a medium, then what are they?

Modernist work that disrupts conventions and expectations functions to make those conventions and expectations, and their consequences as media functions, become explicit rather than implicit. In a sense this amounts to the deliberate production by modernist artists of what might be seen as ‘misfit variables’ in Alexander’s (1970) terms in relation to the preceding media culture and how it expects art (or a medium) to function. Just as design misfits make design features visible, disrupted expectations make media form and function visible. Modernist works therefore function as explicit statements of abstracted media form and affect.

This paper will not dwell in any detail upon the vast and complex field of modernism. However, formal reflexive ideas of a kind that have been strongly demonstrated within modernism are here regarded as informing a large-scale cultural project of maximizing self-consciousness within various media. One key issue here, however, is that, as Small (1994) points out in the case of avant garde cinema, the high levels of self-consciousness and reflexivity involved in much formally reflexive practice do not rely upon textual or verbal representations. Rather, a great deal of work is self-consciously articulated within the tradition of its medium, in a form referred to by Small (1994) as direct theory. Writing may surround, refer to, critique and analyse works in other (than literary) media, but those ‘other’ media also have an autonomous discourse. This is the whole point of reflexive design, to create work that reflects upon its own form.

These considerations apply within all media. It may also be observed that formal reflexive or programs within a medium tend to ‘burn out’ or become exhausted after a period of intensive investigation. Avant garde artworks thereby function to map out a space of design possibilities from which ongoing work may draw without in itself being avant garde. In other words, avant gardes define a self-conscious design space in terms of which ongoing design may locate and define itself. Further innovations within these (‘modernistically’) exhausted media may proceed via postmodern strategies, often involving hybrid media forms, reiteration of past styles and forms, new combinations of formal elements, and self-
conscious production of pre-modern or naive forms in the form of self-conscious kitsch. This tends to occur together with a devaluation of the academised values of modernism and formal reflexiveness (which may be regarded as being ‘too sterile’), and the distinctions between ‘high’ culture, popular culture and commercial culture break down. Meanwhile, if new media are developed, the modernist impulse becomes relevant again in application to those media, to push their limits and expose their form, meaning and functions.

Applying the formal reflexive perspective to game design asks fundamental questions about the nature of games and play, leading to experiments intended to stress our understanding of and assumptions about games and play. Examples of games manifesting this perspective include the genres of alternate reality games (see http://www.argn.com/) and larp's played in games spaces where people may be interacting with gamers in a game but without any consciousness that they are part of the game. These examples raise questions about whether unwitting participants are players or not, and even of whether the game is a game or not, when it has extra-game consequences for (possibly unknowing) participants.

Formally reflexive game design constitutes a kind of avant garde game design practice having philosophical and operational similarities to other avande garde practices in the history of the arts. This does not mean that formalist work exhausts the possibilities of dealing with games from a contemporary artistic perspective. For example, artistic projects concerned with games may also be driven by theoretical considerations, or focus upon various rhetorical possibilities of game form. However, it is formally reflexive game design concepts that have the greatest self-consciousness about the fundamental nature of games and play.

An Integrated Meta-model

The design approaches described above represent an increasing level of explicitness and self-consciousness about game design, leading to the integrated design meta-model illustrated in Figure 4. The triangle form is used to suggest the metaphor of a pyramid, with higher levels being at least conceptually built upon lower levels, and involving a hierarchy of increasing abstraction and conceptual essentialism.

![Game Design Meta-Model](image)

Figure 4. Game Design Meta-Model.
The pyramid from the base to the apex also represents an increasing scope of novelty and innovation. At the implicit and cookbook levels novelty tends to be limited to representational content and small-scale details of formal design. Representational innovation refers to the development of new and novel types and instances of game stories, scenarios, characters and game objects, especially as conveyed by graphics and sound design. Small scale details in the game form itself might include elements such as parameter ranges for the features of game characters (e.g. attribute statistics, like strength, intelligence and dexterity), inventory items (values, damage points), etc. At the level of design patterns, novelty may be achieved by new combinations of patterns being realized together within a specific game. At the level of theoretically motivated design, theoretical motivations may lead to, facilitate or require novel game mechanics, or mechanics designed to frustrate player expectations in order to make a specific point or to serve particular rhetorical aims. At the formal reflexive design level novelty may occur in the most fundamental aspects of a game design, leading to totally new kinds of games and play experiences, or even to questioning and redefinition of the players’ understanding of the very nature of a game.

Results of Applying the Design Meta-model

The design meta-model presented here has been applied in a game design education context. The model was initially devised to address the concerns listed in the introduction, and particularly with a view to structuring design activities aimed at achieving higher than usual levels of innovation. ‘Higher than usual’ here refers to a history of design workshops within undergraduate game development programs and other professional and semi-professional design workshop contexts.

Statistical empirical testing of this kind of meta-model is impractical in the short term, the best global measure of its usefulness being the degree to which it may be referred to and/or used over time by professional game designers or game design educators. In terms of our immediate experiences in applying the model within undergraduate game design education programs, the following anecdotal evaluation information is offered:

– the meta-model has been used as the foundation for an advanced game design course in which students study the various levels of the proposed hierarchy. Within the course, games are brainstormed and developed to a playable hierarchy. Within the course, games are developed to a playable stage and evaluated in a sequence corresponding to a movement up through the hierarchy, beginning in this course with the design patterns and taxonomies level, moving through theory-driven designs and ending with formally reflexive game design.

– the meta-model appears to be understandable to students (although not all levels are initially obvious). Students in the third year of a game development education program expressed the view that they really would have benefited from having this framework presented to them much earlier in their studies, since it provides a framework and language for talking about game design that they had lacked and would have greatly benefited from.
– the higher levels of the design pyramid represent areas that could be developed in endless ongoing detail. The framework therefore appears to represent a very convenient conceptual model for integrating ongoing research and development activities in game design, game aesthetics and related fields. Within the environment of the authors, the framework is very appealing as a high level map clearly interrelating the content of undergraduate game education programs and higher-level game research activities.

– the design workshops at the game pattern, theory-driven and modernist levels have resulted in examples of games having relatively high levels of interest and novelty compared with the typical results from game concept workshops in our experience. One simple example is the formally reflexive computer game Sumo, designed by Kajfa Tam. Sumo is a two-player game in which each player must place their fingers on specific keys on the keyboard. They must then try to use their respective hands to push their opponent’s hand so that at least one finger is pushed off its assigned key, without stopping pushing down on their own assigned keys. The first person to take a finger off a key is the loser. Sumo is a very simple game that nevertheless completely violates our normal expectations about computer game play and interaction.

– in many cases the initial reaction of students to the design assignments based upon the meta-model has been trepidation if not outright fear of entering design spaces having few if any exemplified precedents. Despite this, most of the resulting designs are successful in achieving fresh results and often highly entertaining gameplay.

The meta-model has also been very useful in clarifying our own thinking about design processes and methods, their interrelationships and the role of different kinds of research.

Conclusion

The game design meta-model presented in this paper is a principled heuristic framework interrelating a variety of design approaches, including implicit design (by copy), cookbook design methods, taxonomy and ontology-based game design, theory-driven design and formal reflexive design. The theory-driven level inspires new game and play concepts based upon technical, scientific and theoretical innovations, while the formal reflexive level represents the application of contemporary artistic perspectives to games. The meta-model provides a clear account of the nature and place of research both for motivating design decisions and for game design innovation, and provides a foundation that can be used for game education curriculum development integrated with higher-level research. We do not claim that the model is absolute; for instance, the boundaries between levels could be drawn differently; they represent tendencies rather than precisely definable distinctions. However, our experience indicates that the meta-model is effective in opening up new ways of thinking about, talking about and practicing game design, leading to fresh and innovative gameplay concepts.
Acknowledgements

We express our thanks to Staffan Björk and Jussi Holopainen for holding a week long game workshop on design patterns for us, and for many fruitful and interesting discussions about game design patterns and related issues.
References


Kreimeier B. ”The Case For Game Design Patterns”, Gamasutra, December 12, 2002.


69


A Cognitive Framework for the Analysis of Game Play: Tasks, Schemas and Attention Theory

Craig A. Lindley and Charlotte C. Sennersten
Game Design, Cognition and Artificial Intelligence Research Group,
Department of Game Development, Narrative and Time-Based Media,
Gotland University College
Cramergatan 3, SE-621 57, Sweden
+46 498 299 765
(craig.lindley.charlotte.sennersten)@hgo.se

Abstract
A fully developed nomology for the study of games requires the development of explanatory theoretical constructs associated with validating observational techniques. Drawing from cognition sciences, a framework is proposed based upon the integration of schema theory with attention theory. Cognitive task analysis provides a foundation for preliminary schema descriptions, which can then be elaborated according to more detailed models of cognitive and attentional processes. The resulting theory provides a rich explanatory framework for the cognitive processes underlying gameplay, as well as detailed hypotheses for the hierarchical structure of pleasures and rewards motivating players. Game engagement is accounted for as a process of schema selection or development, while immersion is explained in terms of schema execution. This framework is being developed not only to explain the substructures of gameplay, but also to provide schema models that may inform game design processes and provide detailed criteria for the design of patterns of game features for entertainment, pedagogical and therapeutic purposes.

Introduction
This paper refers to examples of computer role-playing games (RPGs) and first-person shooters (FPSs). In reading this, readers familiar with computer games will have expectations about the design features in question and the nature of the experience of playing these genres of games. Game designers similarly have expectations about what kinds of play styles and experiences may be supported by particular game designs, based upon their own experiences as both players and designers. While good game designers can design good games based upon experience, design principles and extensive tuning by play testing, the relationship between the intentions of the designer, the player experience and the mechanisms learned by players to facilitate play have not yet been studied extensively or in depth in an explicit way. A more scientific understanding of the relationship between design features, player cognitive processes and the experience of play can provide clearer guidelines for design and the realisation of design intentions, as well as providing a more

rigorous foundation for the application of games in fields where the cognitive changes involved in learning how to play a game are required to have specific effects for purposes such as therapy and education.

The scientific study of gameplay requires the establishment of construct validity for the theoretical frameworks and practical methodologies used. Construct validity can be described as the degree to which inferences can legitimately be made from the operationalizations (i.e. practical experiments) in a study to the theoretical constructs on which those operationalizations are based (Cronbach and Meehl, 1955). Cronbach and Meehl describe how a nomological network provides construct validity, where a nomological network includes “the theoretical framework for what you are trying to measure, an empirical framework for how you are going to measure it, and specification of the linkages among and between these two frameworks”. It is a necessary condition for a construct (i.e. a distinction, concept or theory) to be scientifically admissible that it occur within a nomological (or ‘lawful’) net having at least some laws involving observables.

Figure 1.

Gameplay is a cognitively complex activity for which a suitable nomology is needed in order to understand play on a scientific basis. A nomology for game research must provide laws correlating player features, processes and effects with game play patterns in relation to game design features. In this paper it is proposed that the theories and methods of cognition sciences provide a strong foundation for a game research nomology. This nomology has

27 A more complex nomology might include theories and methods for investigating factors of designer intention, but once a game is shipped this is lost and the player’s experience is a function of what the player brings to the released artefact of the game software and the unfolding experience of interaction with that artefact.
a high level structure as depicted in Figure 1. Here methods are used to observe a player and their interactions with a specific game design. The interaction patterns, referred to as "gameplay gestals," are compared with previously observed patterns based upon different game stimuli designs. Analysis is conducted using a theoretical framework providing an account of the features of game designs, of player cognition, and of play interaction relating design features to cognitive processes. This comparative analysis is used as a basis for validating or adapting models of hypothetical cognitive schemas underlying game play, or to modify the overall theoretical framework if those modifications are found to better account for observed play patterns. It is also possible to adopt a computational cognitive science methodology, leading to the computational implementation of the hypothetical schemas of players, followed by comparison of the play results with those of human players as a method of refining and validating the schema models.

In this paper it is proposed that the abstract nomology above can be instantiated in a more specific form based upon attention theory and schema theory, together with a methodology of cognitive task analysis, as shown on figure 2. The resulting framework is discussed in some detail. Hierarchical schema structures are then considered that reflect hierarchical time structuring principles designed into games. The paper then goes on to consider how current theories of game enjoyment, immersion and engagement can be mapped onto a schema-based theory of the cognitive processes underlying play in relation to design features. Schema and attention theory provide foundations for a detailed account of the nature of game engagement and immersion, as well as providing explicit models of the cognitive learning outcomes of play. This theory promises a more explicit mapping from designs to cognitive affects and play patterns than current design practice allows for, providing clearer principles for design decision making and the translation of design intentions into patterns of features.

A Cognitive Theory of Gameplay: Tasks, Attention and Schemas

Analysis of gameplay must begin with a consideration of game features that motivate, facilitate and constrain player action. Game design features include built in game tasks, modelled by a process of "game task analysis", that ground the in-game meaning of design features. Game tasks include configurational tasks such as trading and inventory management, and teleological and planning tasks such as quest management, found in RPGs, and game challenges (Rollings and Adams, 2003) such as coordination, reflex and reaction time challenges typical of both RPGs and action games. The result of a "game task analysis" is the representation of a game task model, which is a representation of the goal environment by which play is made meaningful: the primary driver of play is the motivation to achieve goals within the game task environment. Goals may include things like defeat autonomous enemies (‘mobs’), find a key, open a door, finish a level, etc.. The game task environment of an FPS or RPG is likely to be hierarchical, the high level goal being to finish the game, which can be decomposed into subgoals of finishing each level, which in turn may be subdivided into completing significant scenes, areas or quests within each level, which may also be subdivided into significant challenges along the way, etc..
A player plays a game via an interface. During this process, the interface and a pattern of interaction are learned that result in progress being made within the terms of progress designed into the game and measured by factors such as (in the case of FPSs, and RPGs) the defeat of enemies, progressing through the game world settings and locations, gathering wealth and items, increasing in skills, experience points and capabilities. A repetitive pattern of game interaction, observed by an analyst, may be referred to as a *gameplay gestalt* (Lindley, 2002), so called because it is a grouping of observed interaction features recognised as a repeating pattern.

A cognitive explanation of gameplay must provide an account of the underlying structures and processes that result in a player manifesting a gameplay gestalt. Here it is proposed that such an explanation may be based upon a combination of concepts from attention theory and schema theory, hypothetical schemas being developed on the basis of gameplay analysis conducted by the application of Cognitive Task Analysis (CTA).

*Attention theory* addresses issues of attentional focus, management of attention (including attentional selection), and the allocation of cognitive resources to cognitive tasks. Ongoing research is addressing the question of the detailed operation of attentional mechanisms, including questions such as the degree to which attentional capacity is specific to specific cognitive resources (or modes) or sharable among resources according to demand, and the stage of processing of perceptual information at which perceptual information is selected for attentional priority.

![Figure 2. Proposed framework for the cognitive analysis of game play.](image)

*Cognitive Task Analysis* (CTA) refers to a variety of methods used to analyse and represent the knowledge and cognitive activities that people utilize to perform complex tasks in a work domain (CTA methods and related techniques are reviewed at [http://mentalmodels.mitre.org/cog_eng/ce_methods_1.htm](http://mentalmodels.mitre.org/cog_eng/ce_methods_1.htm)). Different CTA methodologies
emphasise different aspects of tasks and their context and use their own particular task definition constructs. For example, Critical Decision Method (CDM) uses probes, which are targeted analysis questions addressing Decision Point Options, Cues, Causal Factors, Goal Shifts, Analogues, Errors, Hypotheticals, Missing Data, Imagery and Task Analysis. Particular probes in this method are cross-referenced to specific forms of knowledge, including Structure, Perceptual, Conceptual, Analogues and Prototypes (see Klein, 1996). Variants of CTA may address issues such as the identification of knowledge resources within organisations and the contextual functions of cognitive decision processes within broader operational systems. These contextual issues are of limited relevance to the cognitive task analysis of entertainment gameplay, although player subcultures and the contexts of play may have a strong influence on functions, expectations and motivations of play for individual players. Furthermore, verbal methods, such as structured interviews and think-aloud protocols, relied upon heavily in many CTA methodologies, are of limited effectiveness for analysing less conscious or heavily automatised aspects of gameplay, or for observing small scale details of cognitive processes and their interaction with limbic (i.e. emotional) processes. Hence verbal techniques need to be complemented with sociological/ethnographic methods, together with interaction logging, psychophysiological methods, eye-tracking, brain scanning, etc. GOMS (standing for Goals, Operators, Methods and Selection; see John and Kieras, 1994) is a promising foundation for detailed player cognitive task modelling, with the advantage of having variants that can be computationally executed.

Attention theory provides a model of the energetic resources available to cognition, together with an account of principles for the distribution of energy (or attention) to the cognitive resources that use (or manifest) it (see, for example, Johnson and Proctor, 2004). CTA provides a representation of cognitive tasks involved in gameplay at a level of description that makes sense to an observer and conforms to an observed (set of) gameplay gestalt(s). A cognitive task (CT) model can be used for workload analyses that relate subjective experiences of cognitive workload to contextual demands, psychophysiological indicators, situation awareness and performance (see, for e.g., GARTEUR, 2003). However, the framework proposed here is intended to provide foundations for a detailed account of the cognitive processes and structures involved in play that underlay and explain task performance as addressed by workload analysis. This kind of cognitive account must include the detailed mechanisms by which attention is orchestrated together with cognitive resources, process and skills, in relation to perceptual inputs and in order to generate the motor outputs involved in gameplay. Hence CTA and workload analysis provide ‘top down’ routes towards understanding the cognitive functions and structures of gameplay. GOMS is a more detailed CTA method, but further detail is required in order to address learning, fatigue and mapping to neurophysiological mechanisms. Furthermore, task models do not address declarative understandings or structures facilitating comprehension and task performance.
It is proposed here that a sufficiently detailed explanation of gameplay can be provided by developing an account of the cognitive schemas underlying gameplay, which can be referred to as *gameplay schemas*. Schemas are cognitive structures that link declarative (or factual) and procedural (or performative) knowledge together with other cognitive resources (such as memory, attention, perception, etc.) in patterns that facilitate the manifestation of appropriate actions within a context. While the taxonomical structures of semantic or declarative memory are comprised of object classes together with associated features and arranged in subclass/superclass hierarchies, the elements of schemas are associated by observed contiguity, sequencing and grouping in space and/or time (Mandler, 1984). Schemas can refer to declarative knowledge and taxonomical types with their features and relationships, and integrate these with decision processes. Schemas include scripts for the understanding and enacting of behavioural patterns and routines, a classic example being Shank and Abelson’s (1977) example of the *restaurant script* that includes a structure of elements for entering a restaurant, sitting down, ordering food, eating, conversing, paying the bill and leaving (etc.). Scripts, as structures used for both comprehension and behaviour generation, represent a structure of cognitive functions that may include cognitive resources, perceptual interpretations and preconditions, decision processes, attention management and responsive motor actions. *Story schemas*, are patterns representing understandable elements that must occur to make stories comprehensible. The presence of story schemas in the cognitive systems of storytellers, listeners, readers or viewers of stories allow stories to be told and to be comprehended, including the inference of missing information. If a story deviates too far from a known schema, it will not be perceived as a coherent story. Script and story schemas are concerned with structures of both space and time, while *scenes* are schemas representing spatial structures, such as the layout of a house, a picture or an area of a city.

While schemas have been interpreted in many different ways, here a *gameplay schema* can be understood as a cognitive structure for orchestrating the various cognitive resources required to generate motor outputs of gameplay in response to the ongoing perception of an unfolding game. A gameplay schema is therefore the structure and algorithm determining the management of attentional and other cognitive, perceptual and motor resources required to realise the tasks identified by a CTA of game play. A CTA itself provides a first approximation description of a gameplay schema, but a CTA is also heavily determined by the language and cultural constructs of the observer. The phenomenologically meaningful terms of a CTA may have to be further analysed to account for the ways in which those high level constructs are actually realised by underlying neurophysiological constructs, and this mapping could involve different parsings of functional units at the CTA and neurophysiological levels. Hence a gameplay schema might be described at different levels of abstraction, some being meaningful in terms of the subjective languages of task performance (e.g. the terms of self-reported task performance) or CTA and others in terms of implementational neurophysiology that may have a very different structure and functional decomposition than that of more linguistically conditioned accounts. Different levels of abstraction in the description of gameplay schemas are equally valid,
neurophysiological descriptions providing an implementational explanation of verbally expressed task-oriented descriptions.

Hierarchical Time and Schema Structure in Games

Games are designed with time structures that can be considered in terms of a hierarchy of time scales (Lindley 2005a,b). At the smallest scale of time units the game simulates continuous actions and processes, such as character, non-player character (NPC) and mob movements, weapon and projectile behaviour, flowing water, rockfalls, the movement of abstract geometrical objects, doors opening, explosions, etc. These simulations proceed according to a sequence of discrete simulation time steps, or ticks, that may (but do not necessarily) correspond with the rate of update of graphical display frames. Perhaps with the exception of movement control (e.g. by pressing ‘a’, ‘w’, ‘s’ and ‘d’ keys to move left, forwards, backwards or right over extended simulation ticks) player actions are usually not tracked as extended sequences over ticks at this scale; more typically player actions occur at discrete points within the time sequence, functioning as triggers initiated by a key press or release, a mouse button press or release, etc. Those player actions may be made meaningful at a scale above that of single simulation ticks, initiating the automated simulation of synthetic actions over an extended number of ticks, such as slashing with a sword or firing a missile. These larger scale units of meaning may be said to belong to the game scale, since they correspond to what may be called game moves in more traditional (non-computational) game forms. In general, the available (legal) types of games moves in a particular game are the types of actions specified in game rules, while events at the simulation scale in a computer game implement game moves.

Game moves are primitive player actions that can be interpreted from several perspectives. The game perspective grounds their significance in the competitive, rule constrained form of a game. However, moves can be regarded from other perspectives, such as socialisation, construction, trading and dramatic performance. As noted in Lindley (2005a, 2005b), game moves regarded from the perspective of narrative or dramatic performance are equivalent to what Mackay (2001) refers to as fictive blocks, basic fragments or units of fictional/narrative significance that may be strung together to form a larger scale narrative. Mackay takes fictive blocks divorced from their original context to be equivalent to Schechner’s strips of imaginary behaviour, patterns that constitute a repertoire of potential behaviours that are performed by an actor in new arrangements in ways that may appear spontaneous and unrehearsed during improvisational performances.

This kind of ‘perspective’ for interpreting and providing significance to game moves is not merely a matter of arbitrary interpretation, but a fundamental condition of making the play experience comprehensible. A computer game typically provides a player with a limited number of commands to perform, including both the initiation or performance of game moves and metagame commands such as saving the game state or quitting the game. Game moves may be moves such as ‘move left’, ‘move forward’, ‘move right’, ‘fire’, ‘toggle run/walk’, ‘cycle weapon’, ‘gesture’, ‘tilt’, etc.. Suppose there are a total of 30 such
move types to choose from while playing a specific computer RPG (this is an easy figure to reach considering basic movement controls and the possibility of having multiple hotkey assignments to each function key, e.g. for selecting weapons, health boosts and spells). Considering a sequence of 10 moves made by the player, the number of combinations of 10 selections from 30 available moves is $30^{10}$, or $59,049 \times 10^{10}$ move sequences. Since a typical single player RPG may be played for a hundred hours or so, often with the player selecting several moves per second, the combinatorial space of possible move sequences over the play time required to finish the game is extremely large (in excess of $10^{5319}$).

By far the major part of this vast space of possible move sequences would make no sense at all to play. ‘Making sense’ here is a matter of conforming to the pattern of expectations of a higher level play structure, which is a temporal game play schema, and in particular a script, providing at least an implicit perspective for the interpretation of the meaning of moves. Many kinds of scripts might apply to making sense of a player’s selection of game moves in a play sequence. Some examples of these are:

– a script for the combative engagement of an enemy. This script might begin with the detection of an enemy and then proceed through selecting a configuration of armour, weapons and magic according to the enemy type, deciding upon a tactic for approaching and engaging the enemy, configuring companions if the player controls a party of NPCs, selecting a formation and/or angle of attack, launching the attack, perhaps with a specific sequence for using different weapons upon approach, cycling between attacking, parrying and taking health boosts, deciding to retreat or carrying on to death or victory

– a script for exploring a labyrinth, e.g. the ‘left hand rule’ of following the left wall and occasionally checking to detect traps or secret doors, until the character arrives back at the entrance

– a script for interacting with a trader, including evaluating items, selling unwanted and unnecessary items, upgrading useful items and replenishing diminishable stocks

– a script for negotiating quests, e.g. exploring background information, considering alignment implications, accepting or rejecting a quest, returning to a quest giver when the quest is accomplished, accepting or rejecting rewards or punishments, deciding upon follow on quests, etc.

Many other schemas are possible, including those for interacting with other players for accomplishing various group tasks or activities in multiplayer games.

The designs of FPS and RPG computer games frequently embed the moves of players within higher level pre-designed time structures, including structures of levels and high level narratives with key structural elements (such as introducing a central conflict, major
plot points, and final closure after defeating a game boss) being delivered as non-interactive cut scenes (Lindley, 2005a, 2005b). Schemas for stories (Mandler, 1984) provide structures by which players may anticipate and comprehend high-level story structure within which their gameplay via gameplay schemas are embedded. Game structure schemas based upon levels do not need to present any more specific narrative structure, but include expectations of thematic changes and changes in the difficulty of play, increases in player character experience, and increases in the statistics of items such as found weapons and artefacts. Story schemas for computer games can include transitions between story sections that also function as level transitions with their associated expectations and features.

The function of player schemas in making a gameplay experience comprehensible raises a critical question of how design intentions are translated into game design features with which the player interacts. Design decisions are also made according to schemas, this time in the cognitive systems of designers. These latter schemas amount to the designers’ understanding of how players will make design features meaningful. Once a design is released, however, only the design features remain and it is up to the players to find ways of creating their meaning. Shared schemas among designers and players can facilitate effective game design. Designers’ conceptions of play schemas may take the form of high level ideas about broad styles of play, but are otherwise implicit and learned from the designers’ own playing and design experience. Undesirable consequences may follow from this. For example, players may play a new game according to their established schemas and thereby miss new possibilities of play afforded by the design (which are thereby rendered pointless). To overcome this, specific design techniques may be required to disrupt established player schemas and initiate a fresh gameplay schema learning process. High level story structures, which have been an ongoing issue of contention in game design, are problematic not only due to problems of requiring cognitive resources to be directed away from gameplay for full comprehension (Lindley, 2002) or not appealing to taste variations within players (Lindley, 2005), but also due to potential mismatch between schema-driven story construction on the part of players on one hand with pre-designed story elements created according to the story schemas of designers on the other. Since players are active and highly immersed participants in creating the details of the history of the game experience, their own story schemas can be used to create their own retrospective story of the play experience. These schemas may be different from those of the designers, or may cast or conceive of the player character and its experiences in a role suiting the players’ own tastes, needs, motivations and desires. Infrequent pre-defined story elements, which are not immersive in the same way as the interactive play experience, may not figure strongly in these player constructed stories, or may even clash with them, resulting in player distaste for those design elements. This might be avoided by design techniques based upon heavy infiltration of the play experience with pre-designed story content (as in the Final Fantasy series of games), forcing players to ‘play the designers’ story’, or by designing to facilitate player story construction, with little or no pre-designed story content.
The question of differences between player scripts and designers’ expectations of player scripts is not always problematic, and can in fact be highly desired by designers. This amounts to the design of games facilitating emergent gameplay (discussed, for example, by Salen and Zimmerman, 2004), where complex player behaviours are derived during play in interaction with a game system based upon comparatively simple game rules. Design principles for emergent gameplay are not yet well developed and a schema-based theory, derived from observation of gameplay schemas and taking into account schema learning processes, could provide a foundation for developing stronger emergent gameplay design principles.

Schemas, Immersion and the Pleasures of Play

Gameplay schemas provide a foundation for analysing and understanding the various motivational factors that keep players engaged and immersed in gameplay. Holopainen and Meyers (2000) suggest that players are driven to continue in gameplay due to the desire for both predictive and dramatic closure, where predictive closure is a desire to complete a mental model, while dramatic closure is the resolution of the tension driving a story structure. Holopainen and Meyers suggest that game play consists of the performance of typically repetitive actions, driven by the expectation of closure where in a well-designed game there are multiple hierarchical levels of subclosures. Holopainen and Meyers also suggest an important role for spatial and temporal displacement, which is the tendency of players to project themselves into the representations of the play experience such that rewards defined within the fictional world of a game are experienced as rewards bestowed upon the player. In the terms of the framework described here, temporal closure may be attributed to the operation of schemas having resolutions corresponding with successful executions of the schemas. Hierarchical structures of closures correspond with the kind of hierarchical schema operation evident in gameplay as described above. Displacement can be regarded as a higher level cognitive process by which the representation of agency achieved by game media, together with interaction within the dynamics of the representation via the feedback loop from perception of game events to motor operations enacting game moves, achieves a level of player identification with the player character (‘I killed the goblin’, rather than ‘my character killed the goblin’). A player’s self-identification with their game character links perceptual events attended according to schema priorities with more general criteria of self value and reward built into the player’s cognitive construction of identity.

Klimmt (2003) offers an account of the enjoyment of game play based upon three factors, the experience of effectance, cyclic feelings of suspense and relief, and the fascination of a temporary escape to an alternative reality provided by the fictional world represented by a game. Effectance on the most basic level, corresponding to the atomic level of initiating and experiencing the consequences of game moves as described above, is realised as a series of single loops of player input followed by game system response and output. Effectance provides inherent pleasure in the form of immediate feedback to the player as a causal agent influencing the game world, a satisfying experience in comparison with the often ambiguous or limited influence of actions or intentions in the everyday world. The pleasures
of effectance in computer gameplay may be greatly enhanced, since a small scale action like clicking a mouse button can have large scale in-game consequences, such as mass destruction, landscape alteration or transfers of large amounts of money.

Klimmt describes the next level of organisation in gameplay in terms of episodes consisting of sequences of input/output loops structured by possibilities or a necessity to act followed by action and a result (e.g. overcoming a game challenge) leading to the next episode. Hence an episode corresponds with what Lindley (2005a, 2005b) refers to as a game bout, actually constituting a complete mini-game as a self-contained competitive experience, this corresponding to an action sequence executed according to a gameplay schema as described in this paper. Enjoyment within episodes may include the excitement of possible action, the pleasures of curiosity and discovery, the pleasures of experiencing the negative emotions of suspense followed by the transference of arousal to an ecstatic experience when the challenge creating the anxiety of suspense is overcome, and enhanced self-esteem. Due to the interactive nature of games and the strong ego involvement via identification of players with their game character (i.e the displacement phenomenon described by Holopainen and Meyers, 2000), the pleasures of relief at the resolution of episodes are greatly enhanced in comparison with non-interactive media since they function to boost player self-esteem. Since games offer long sequences of such short episodes, they uniquely offer an ongoing opportunity for the pleasures involved in the transference of arousal from suspense to joyful resolution, driven according to underlying game play schemas.

Finally, Klimmt describes the whole session of playing that includes the sequence of episodes and provides the player with the experience of participating in a narrative. As active participants in the events of the narrative, players have the pleasure of being able to experience new objects, actions, social interactions and experiences at no risk. These vicarious experiences can help players to cope with felt frustrations and deficiencies in their everyday lives, a process both of catharsis and of perception of increased competence and relevance.

Klimmt’s model provides a basis for explaining reward during the detailed execution of gameplay schemas during which each mapping from perception through a decision process to an action is rewarded as the experience of effectance. Schemas as described in this paper offer greater discrimination at the level of episodes by allowing different forms of episodes to be modelled as the result of different schema patterns (e.g. different scripts for combat, exploration, trading and quest negotiation). All of these script types can be rewarded upon completion by the intrinsic rewards of completion described by Holopainen and Meyers (2000), while different types of scripts may have their own specific rewards, such as curiosity and discovery for exploration scripts, the overcoming of suspense for combat scripts, and enhanced self-esteem for trading and questing scripts. Schemas for stories include rewards not just via displacement facilitating vicarious experience, but also through the gratifications of expectations created by the story schema structure being met, a large scale manifestation of the rewards of completion.
It can be seen, then, that schema models may provide a detailed description of the processes underlying gameplay that can accommodate a variety of different forms of player reward. In addition to this, schema theory can provide detailed hypotheses regarding the nature of game engagement and immersion. Douglas and Hargadon (2001) note that it is highly normative schemas that enable readers of narrative texts to ‘lose’ themselves in the text in what they refer to as an immersive affective experience. Douglas and Hargadon contrast this with engaged affective experience where contradictory schemas or elements defying conventional schemas tend to disrupt reader immersion in the text “obliging them to assume an extra-textual perspective on the text itself, as well as on the schemas that have shaped it and the scripts operating within it”. Disruption of expectations requires engagement with an unfamiliar narrative structure.

While Douglas and Hargadon use schemas to explain immersion in and engagement with hypertexts, the same concepts apply to gameplay schemas. As noted above, when a gameplay schema has been learned, a player may attempt to apply it within games that appear to resemble those within which the schema has previously been successful. Hence players may repeat gameplay patterns without exploring new interaction possibilities afforded by new games, as long as the established patterns support progress within the game. A new kind of game in which previously established gameplay schemas do not work might in these terms be said to require players to become involved in the engaged affective experience required for developing new gameplay schemas (or indeed to initially determine which already available schemas apply within a specific game). Once a successful schema is matched or formed, the player shifts from engaged affective experience to immersed affective experience in the performance of the schema.

Extrapolating from Douglas and Hargadon’s (2001) suggestion for the case of interactive hypertexts, and as previously observed by Turkle (1984) in the case of games, gameplay often reaches states of timeless loss of self-consciousness, involving the state of consciousness that Csikszentmihalyi (1991) calls flow, a state at the boundaries between engagement and immersion, of being totally absorbed in meeting a constantly unfolding challenge. Contrary to Douglas and Hargadon’s view, however, this need not have anything to do with narrative. For games it is generally a state of pure gameplay, in the present terms, a state of selecting and performing gameplay schemas to provide an immersion in performance within the internal world of a game, but not (necessarily) in the internal world of a narrative. The performance is itself a process requiring attention, and for successful flow the attentional demands of performing an integrated cognitive and instrumental pattern must fall above thresholds of boredom and within the bounds of what is achievable and sustainable for the player in a flowing state of concentration. Csikszentmihalyi characterizes flow, an experience that is simultaneously challenging and rewarding, as one of the most enjoyable and valuable experiences that a person can have. Careful study of the structure of gameplay schemas, together with an exploration of the attentional mechanisms orchestrated by schema structure, may provide a cognitive explanation of this experience of flow.
It may be asked whether a schema-based explanation is sufficient to explain all forms of immersion in gameplay. Based upon observations of children playing games, Ermi and Mäyrä (2005) have proposed the SCI model in which immersion in gameplay can take three forms. *Sensory* immersion involves an immersion in the audiovisual perceptual qualities of a game. *Challenge-based* immersion involves immersion in the cognitive and motor tasks performed in order to meet the challenges designed into a game. *Imaginative* immersion involves immersion within the represented imaginary world and fantasy of a game. In terms of the framework presented here, imaginative immersion can be regarded as immersion facilitated by the execution of story schemas. Challenge-based immersion can be regarded as immersion facilitated by the execution of gameplay schemas. It is not clear, however, that sensory immersion can be accounted for by a schema theory. Some forms of sensory immersion may be accounted for by schema models, such as immersion in the musical experience of a game, or in visual completions within the visual representational functions of a game (e.g. perceiving an environment, involving scene schemas, or in the rewards of gestalt completion for recognising objects in the game world by cognitively grouping a set of perceptual components). However, immersion in more freely structured perceptual phenomena, involving sounds, colours, lines, textures, movement and forms beyond what they represent, may involve different neural processes, perhaps similar to the kind of selective activation of neural centres to which Ramachandran and Hirstein (1999) attribute the general appreciation of aesthetic value. In this sense sensory immersion may be *extra-schematic*, an immersion in lower level perceptual cognition while perception is not attached to decision processes.

**Conclusion**

This paper has presented a cognitive framework for analysing gameplay based upon game specific schemas and story schemas. This framework supports numerous hypotheses regarding the cognitive processes underlying play, integrating and accounting for the pleasures of play experienced by players. Ongoing research is addressing the development and validation of more detailed schema models derived from observations of players during play. This work requires the validation of cognitive task models, including determination of appropriate levels of abstraction within those models and eventually leading via schemas to a mapping of their terms to models of underlying neurophysiological models. Current experimental methods include the use of questionnaires and speak aloud protocols together with interaction logging and eyetracking studies during gameplay, using both off the shelf and modified commercial games as experimental stimuli.
References


Game Play Schemas: From Player Analysis to Adaptive Game Mechanics

Craig A. Lindley and Charlotte C. Sennersten
Gotland University College
Cramergatan 3
621 57 Visby, Sweden
+46 (0) 498 29 99 00
{craig.lindley, charlotte.sennersten}@hgo.se

Abstract

Schema theory provides a foundation for the analysis of gameplay patterns created by players during their interaction with a game. Schema models derived from the analysis of play provide a rich explanatory framework for the cognitive processes underlying gameplay, as well as detailed hypotheses for the hierarchical structure of pleasures and rewards motivating players. Game engagement is accounted for as a process of schema selection or development, while immersion is explained in terms of levels of attentional demand in schema execution. However, schemas may not only be used to describe play, but might be used actively as cognitive models within a game engine. Pre-designed schema models are knowledge representations constituting anticipated or desired learned cognitive outcomes of play. Automated analysis of player schemas and comparison with predesigned target schemas can provide a foundation for a game engine adapting or tuning game mechanics to achieve specific effects of engagement, immersion and cognitive skill acquisition by players. Hence schema models may enhance the play experience as well as provide a foundation for achieving explicitly represented pedagogical or therapeutic functions of games.

Keywords: Cognitive simulation, schema theory, computer games, gameplay, learning.

Introduction

Computer game genres, such as role-playing games (RPGs) and first-person shooters (FPSs), imply particular sets of design features supporting expectations that prospective players have about the nature of the play experience that games support, based upon past experiences with other games in the same genres. When a player first encounters a computer game within an unfamiliar genre, they will, if sufficiently motivated, interact with the game and eventually learn sufficient patterns of interaction to make progress within the game, perhaps eventually completing it. Gameplay is therefore fundamentally a process of players learning, adapting and improving play skills. Since computer games are predominantly played by the use of very generic interaction technologies (e.g. a keyboard and mouse), learning and adaptation in play are, for the most part, processes of developing cognitive skills focused upon the mechanics of a game and its media realization, based upon an existing general skill set for computer use. Keyboard and mouse operations are mapped onto in-game actions in a game world synthesized by the game software. Learning how to
play can therefore be divided into three phases: 1. learning *interaction mechanics*, that is, the basic motor operations required to operate (for example) a keyboard and mouse in a largely unconscious way, 2. learning interaction semantics, that is, the simple associative mappings from keyboard and mouse operations to in-game actions (and meta-game actions, such as setting play options, or loading and saving game states), and 3. learning *gameplay competence*, that is, how to select and perform in-game actions in the context of a current game state in a way that supports progress within a game. Interaction semantics represent a basic level of competence in playing a particular game; these mappings are often carried across different games within a genre and even across genres (e.g., using ‘w’, ‘a’, ‘s’ and ‘d’ keys to move a player character forwards, left, backwards and right, respectively). Learning interaction semantics represents a form of game challenge (in addition to those noted by Rollings and Adams [9]), but once the basic mappings have been learned, they become a largely unconscious foundation for ongoing gameplay. The focus of learning then shifts to the development of gameplay competence, which involves the development of forms of in-game situation awareness and decision making needed to meet the more complex challenges such as those documented by Rollings and Adams [9].

Gameplay competence involves the ability to: 1. decode the audiovisual sensory and perceptual information delivered by the game media (e.g. the computer screen and speakers) into the apprehension of a local situation within the synthesized game world (or game space), 2. evaluate this understanding of the local in-game situation in terms of the overall objectives of play, current goals and tasks, the state of the player character within the game (e.g. capabilities, health and other statistics), and anticipation of various rewards of playing the game, 3. to make decisions about which in-game tactics and action(s) to perform next, based upon the perceived situation and its evaluation, and 4. perform action(s) based upon competence in interaction mechanics and semantics. The details of the cognitive process underlying this repetitive sequence, which could be described as the sense->model->evaluate->plan->act sequence (essentially the same as the sense->model->plan->act structure used to simulate higher level action control in robots and agents within artificial intelligence research; see http://www.csie2006.murdoch.edu.au/game_ai.html for extensive references), are the primary higher level cognitive learning outcome of learning how to play a particular computer game.

The general usefulness of these different aspects of learning in gameplay relates to the degree to which the knowledge or skills learned may transfer to other contexts. Competence in interaction mechanics is very general, transferring to all contexts within which the same interaction technologies are used; however, the contribution of a particular game to the development of this competence is likely to be very limited, and certainly no greater than other applications using the same interface technology. In fact, a game may be less effective than other applications that are more demanding in terms of knowledge, e.g., of keyboard layout, such as word processors. Competence in interaction semantics transfers only to other systems using the same mappings from mechanical interaction operations to in-game actions. This may include many other games, especially those within the same genre but also
across genres, depending upon their adoption of implicit or explicit conventions in game interaction design. However, interaction semantics may be limited in their transferability to other contexts, since contemporary methods of triggering synthetic actions synthesized by a computer game are unlikely to be the same as methods of realising actions that are not synthesized by a computer.

Gameplay competence has similar transferability across computer games to competence in interaction semantics, i.e. high transferability within a genre but decreasing across genres. However, the potential for transfer of gameplay competence to contexts other than computer games may be much greater, since similar cognitive processes implementing a sense->model->evaluate->plan->act sequence could apply within those contexts. For example, a flight simulator based upon accurately modelled flight planning and air traffic control procedures may help players to learn how to manage flight planning and air traffic control operations in a real flying context. The key issue here is whether the particular mechanics and design features of the game lead to the development of cognitive structures that can transfer to other contexts. The effectiveness of computer games as situated learning environments (as characterized by [3]) critically depends upon this issue of transfer.

The nature of the cognitive structures underlying gameplay is not only relevant to knowledge and skill transfer. Those structures are the key to therapeutic applications of gameplay (e.g. [8]), and in fact are the key to the ability of computer games in all contexts to engage and immerse players and motivate ongoing play. This follows since it is the game play schema driving the situated decision process that determines the nature and timing of emotional rewards motivating play. Hence a greater understanding of the cognitive structures underlying gameplay and how motivations and rewards are related to these can aid in better game design in entertainment, pedagogy and therapy. More than this, it is the central claim in this paper that explicitly modeling those cognitive structures and processes within a computer game engine has the potential to greatly enhance design effectiveness by providing the foundations for a game system itself to guide the development of cognitive structures and control the emotional rewards underlying play.

This paper explores this issue by first considering the cognitive framework for analyzing gameplay described by Lindley and Sennnersten [6]. Methods for conducting analyses of play with a view to identifying underlying gameplay schemas are then described. Based upon this, the paper goes on to consider potential methods by which a computer game system might itself form hypotheses about the schemas underlying the play of a particular game. Finally, we consider some ways in which hypothesized gameplay schemas can be used automatically within a computer game system to modify game mechanics as a basis for guiding play and influencing ongoing schema formation and refinement on the part of a player. This work differs significantly from many past projects to create computational players (e.g. see [12]) in that the latter are typically focused on optimal gameplay methods that do not need to use computational techniques based upon human play performance. In the case of the work described here, the particular strength and interest of the method
is the characterization and explicit representation of the specific algorithmic strategies and cognitive processes of human players, both for analytical purposes and as a foundation for adaptive game mechanics.

A Cognitive Theory of Gameplay: Tasks, Attention, Schemas and the Pleasures of Play

Lindley and Sennersen [6] present a theory of the underlying cognitive systems involved in gameplay based upon schema theory and attention theory. Schemas are cognitive structures that link declarative (or factual) and procedural (or performative) knowledge together in patterns that facilitate comprehension and the manifestation of appropriate actions within a context. While the taxonomical structures of semantic or declarative memory are comprised of object classes together with associated features and arranged in subclass/superclass hierarchies, the elements of schemas are associated by observed contiguity, sequencing and grouping in space and/or time [7]. Schemas can refer to declarative knowledge and taxonomical types with their features and relationships, and integrate these with decision processes. Schemas include scripts for the understanding and enacting of behavioural patterns and routines, a classic example being Shank and Abelson’s [11] example of the restaurant script that includes a structure of elements for entering a restaurant, sitting down, ordering food, eating, conversing, paying the bill and leaving (etc.). Scripts, as structures used for both comprehension and behaviour generation, represent a structure of cognitive functions that may include cognitive resources, perceptual interpretations and preconditions, decision processes, attention management and responsive motor actions. Story schemas, are patterns representing a structure of understandable elements that must occur to make stories comprehensible. The presence of story schemas in the cognitive systems of storytellers, listeners, readers or viewers of stories allow stories to be told and to be comprehended, including the inference of missing information. If a story deviates too far from a known schema, it will not be perceived as a coherent story. Script and story schemas are concerned with structures of both space and time, while scenes are schemas representing spatial structures, such as the layout of a house, a picture or an area of a city.

While schemas have been interpreted in many different ways, here a gameplay schema is understood as a cognitive structure for orchestrating the various cognitive resources required to generate motor outputs of gameplay in response to the ongoing perception of an unfolding game. A gameplay schema is therefore the structure and algorithm determining the management of attentional and other cognitive, perceptual and motor resources required to realise the tasks involved in gameplay. Examples of types of gameplay schemas described by Lindley and Sennersen [6] include story scripts for understanding high level narrative structures designed into games, and scripts for the combative engagement of an enemy, exploring a labyrinth, interacting with a trader non-player character, and negotiating and carrying out quests.

Attention theory provides an account of the energetic resources available to cognition, together with principles for the distribution of energy (or attention) to the cognitive
resources that use (or manifest) it. Attention theory addresses issues of attentional focus, management of attention (including attentional selection), and the allocation of cognitive resources to cognitive tasks. Ongoing research is addressing the question of the detailed operation of attentional mechanisms, including questions such as the degree to which attentional capacity is specific to specific cognitive resources (or modes) or sharable among resources according to demand, and the stage of processing of perceptual information at which perceptual information is selected for attentional priority. Schemas can be regarded as mechanisms or algorithms that, amongst other functions, determine the allocation of attention to cognitive tasks.

In the context of gameplay, attention and the operation of gameplay schemas is driven by hierarchical goals that set tasks for a player. Goals include those intended by designers and those created by players as allowed by a game design. A hierarchical decomposition of gameplay goals might at a high level include the completion of a game, which decomposes into the subgoals of finishing each of its levels, each of which in turn decomposes into goals of completing a series of game challenges (and other tasks invented by the player).

We hypothesise that this hierarchical goal structure is mirrored in a hierarchical structure of schemas within a player’s cognitive system, where a schema is an algorithm for completing a particular goal or subgoal. As argued by Lindley and Sennertsten [6], this schema structure is fundamental to many aspects of the pleasures and motivating factors behind play. These include the pleasures of:

- **effectance**, which is a basic feeling of empowerment created when an action of a player results in a response from the game system [5]. The cause-effect relationships underlying effectance are a fundamental premise of goal-oriented schemas for action.

- **closures** at different hierarchical levels (as described by Holopainen and Meyers [4]), where a closure is interpreted here as the completion of the algorithm constituted by a play schema. Closures may involve completion of expected outcomes and resolution of dramatic tensions, corresponding to the completion of cycles of suspense and relief identified by Klimmt [5]. A distinction must be made here between the intrinsic pleasures of schema completion and more complex emotional experience and rewards due to fictional identification within the game world (see the point below regarding episodes).

- **achievement** of in-game tasks, which is rewarding due to the displacement of a player’s identity into their character [4], this being a matter of *imaginative immersion* as described by Ermis and Mäyrä [2]. Achievement oriented reward is a more specific form of reward than mere closure, since it is associated with the completion of schemas by the achievement of specific goals.
– more complex forms of enjoyment in game tasks regarded as episodes [5] following from imaginative displacement into the game world. Enjoyment within episodes may include the excitement of possible action, the pleasures of curiosity and discovery, the pleasures of experiencing negative emotions of suspense followed by the transference of arousal to an ecstatic experience when the challenge creating the anxiety of suspense is overcome, and enhanced self-esteem. Schemas offer greater discrimination of the pleasures involved in episodes by allowing different forms of episodes to be modelled as different schema patterns having a complex substructure with corresponding emotional effects (e.g. different scripts for solving mysteries, combat, exploration, trading and quest negotiation).

– escape to an alternative reality provided by the fictional world represented by a game [5] and facilitated by imaginative displacement. Players have the pleasure of being able to experience new objects, actions, social interactions and experiences at no risk. These vicarious experiences can help players to cope with felt frustrations and deficiencies in their everyday lives, a process both of catharsis and of perception of increased competence and relevance. Schemas for stories facilitate displacement, while many additional schema forms provide the foundations for comprehension of the events within the fictional world and provide mechanisms for projection of the player’s sense of self into the fiction.

– achievement of a sense of flow [1] in gameplay, this being a state at the boundaries between engagement and immersion, of being totally absorbed in meeting a constantly unfolding challenge. We hypothesise that the flow state is associated with attentional demand, in particular occurring when schema execution demands attentional resources above a level that would result in player boredom and below a level that would result in excessive difficulty and consequent frustration.

Schema theory therefore has the potential to provide both an explanation of the decision and operational processes underlying gameplay and an explanation of the detailed reward and motivation factors behind play. Validating this potential requires detailed study of play resulting in the development of empirically validated hypotheses about the detailed structure and functionality of gameplay schemas, for individual players and across groups of players.

Methodologies for Identifying Gameplay Schemas
Identification of gameplay schemas is a knowledge acquisition and representation process. Our current methodology for doing this includes analysis of the design features of test games, logging of player key strokes and mouse movements, recording of the screen history of play, eyetracking data showing the locus and dynamics of player gaze behavior, and think-aloud protocols to gain some insight into the player’s conscious experience of play and its decision processes. Analysis of this data then proceeds by a process of detailed analysis of individual play sessions in order to identify different play modes and abstract hypothetical
underlying gameplay schemas. This in itself is a complex process that may begin with
cognitive task analysis (CTA, see http://mentalmodels.mitre.org/cog_eng/ce_methods_.
html), but must end with a detailed cognitive explanation of the decision processes involved
in terms of basic cognitive functions. Statistical patterns of play interaction (mouse moves,
key strokes and eye movements) that may correlate with the presence and execution of
specific gameplay schemas are then identified. This requires the separation of an analysis
dataset from which schema models and initial statistical distributions are derived from a
test data set that can then be used to validate those schema models. This sequence is iterated
in order to refine the identified schema models.

The design features of the games used within these studies are crucial. Hence an initial
analysis of the selected games must be made in order to identify their general features.
The iterative process of refining and validating hypothetical gameplay schemas must
also involve the creation of purpose-specific test games or levels, this being done by level
editing and modding (i.e. modification of off the shelf games, potentially including their
media content and scripted behaviour). It is also possible to implement a hypothetical
schema to create a computational player and to test the resulting gameplay interactions
with actual player interaction as another method of validating a schema hypothesis. As
noted by Lindley and Sannersten [6], a CTA provides a first approximation description
of a gameplay schema, but a CTA is also heavily determined by the language and cultural
constructs of the observer. The phenomenologically meaningful terms of a CTA may have
to be further analysed to account for the ways that those high level constructs are actually
realised by underlying neurophysiological mechanisms, and this mapping could involve
different parsings of functional units at the CTA and neurophysiological levels. Hence a
game play schema might be described at different levels of abstraction or from different
interpretation perspectives, some being meaningful in terms of the subjective languages of
task performance (e.g. the terms of self-reported task performance) or CTA and others in
terms of implementational neurophysiology that may have a very different structure and
functional decomposition than that of more linguistically conditioned accounts.

The choice of the level of abstraction in gameplay schema descriptions may depend
upon the purpose of the analysis. More importantly, however, it may be that distinctive
statistical profiles can be associated with schema characterizations at an optimal level of
abstraction; more abstract schema descriptions may be too general to have any statistical
discrimination between them, while more detailed descriptions may involve details that
cannot be correlated with statistical groups. Hence an important ongoing task is the
statistical validation of suitable levels of schema description. It is yet to be determined
how consistent the level of description needs to be, across game genres, games within
genres, different kinds of players, different players within those types, and different play
sessions for the same player. It is hoped, however, that applying this methodology will
result in statistical profiles uniquely associating player types and game design feature sets
with distinctive statistical distributions of interaction primitives at the level of interaction
semantics that indicate specific hypothetical gameplay schemas (or sets of schemas) within
the cognitive systems of those players. This is a large undertaking (and in fact endless, as
game design continues to evolve) that must be approached incrementally by focusing upon
specific genres, games and design feature subsets.

Questions of levels of abstraction and also of higher level structures also apply to interaction
primitives. Basic interactions implement game moves at the semantic level. However, the
presence of specific schemas may be indicated by specific sequences or clusters of interaction
semantics, rather than, or in combination with, their frequency. Different play modes, such
as setting options versus gameplay commands directed towards achieving in-game goals (i.e.
game moves) can often be distinguished by specific discrete interaction primitives, such as
hitting an escape key. However, manual interpretation of play and the formation of schema
hypotheses based upon this is crucial for defining criteria for distinguishing between the
presence of different schemas that involve the same or similar interaction primitives.

To illustrate this discussion, a hypothetical schema can be described based upon observations
of play of the roleplaying game *Neverwinter Nights*. *Neverwinter Nights* is a third-person
point of view game in which the player has a primary in-game character and this character
can gain a number of companions in order to form a team, also controlled by the player.
Within the game world there are many underground labyrinths consisting of rooms and
chambers connected by passages. Rooms and passages often have doors and the labyrinths
in general contain threats such as monsters and traps, non-player characters that may be
friendly or hostile depending upon how the player character interacts with them, power-ups
and various treasures. The play patterns observed for this example occurred within a period
of the game during which the player is intended (by the game designers) to be acting to
achieve a number of higher level story goals pre-designed into the game; in particular, the
player is on a quest to find four specific creatures that are the key to creating an antidote for
a plague and each of which is hidden somewhere within its own labyrinth. Each labyrinth
has a similar abstract structure and distribution of game challenges, with differences in
its thematic realization. This leads to a style of gameplay that manifests highly repetitive
patterns of interaction and decision-making. The schema, expressed in this case in a kind
of high level and informal pseudocode, is a hypothesis about (part of) the underlying
algorithm responsible for manifesting these repetitive patterns as the player character and
team move through the labyrinths:

1. Stop at Closed Door
2. Check health of party
   if >1 party member low, then :
       Rest Party
       Resummon summoned creature
   else
   if 1 party member low, then :
       if lots of healing potions, then:
           administer healing
else
    Rest Party
    Resummon summoned creature
3. Enter combat configuration
4. Open door and enter room
5. If there is an enemy
   Select target
   Monitor health of party until enemy defeated
   if > 1 party member has low health, then :
     Run away
     Rest Party
     Resummon summoned creature
     Go back to step 3
   else
     if 1 party member low, then :
     if lots of healing potions, then:
       administer healing
     else
       Run away
       Rest Party
       Resummon summoned creature
       Go back to step 3
6. If enemy remains, go to step 5
7. Check for traps ...
   ... etc. ...

The question of the level of abstraction involved is illustrated by considering a significant number of possible subtasks and additional tasks that are not represented in the above description: Check health bar for 1 .. N characters, Check for treasure/items to pick up, Check item attributes/quest relevance, Select navigation waypoints for movement, Avoid enemies during retreat, Tweak group member positions, Bring back strays, Check status of quests, Talk with NPCs, Accept/reject quests, Check minimap window, Reconfigure inventory, Reconfigure equipped items, Select level up options, etc..

A complete schema description must include all possible subschemas and include a way of representing the operation of simultaneous parallel schemas, their relative priority and the principles for switching from one schema to another. The detail involved can be high. For example, the detailed description of a subtask such as ‘Check map window’ must include an account of exactly what it is that is being looked for in the map window, how the data is to be interpreted, and some kind of representation for the outcome of the minimap check (e.g. a decision about being lost and/or activating a goal-related reorientation subschema).
Automated Identification of Gameplay Schemas and Schema-Based Adaptation of Game Mechanics

The intended outcome of schema analysis over significant numbers of players and play sessions is a probabilistic profile of the frequencies, clusters and/or sequences of semantic interaction primitives (game moves) associated with different types of underlying gameplay schemas for a specific game (i.e. its design features). If such a set of statistical profiles is available, it may be possible to use the profiles for automated identification of the schemas of particular players/play sessions. It is possible to automatically record (or log) interaction semantics for a particular player during a particular play session or across different play sessions. This will result in a count of the absolute frequency of each type of semantic interaction primitive used by that player, which can be turned into a relative frequency by subdivision with the total count of interaction primitives. This might be used (in addition to specific commands that indicate changes in play mode) to match against a database of statistical profiles of different gameplay schemas in order to derive a probabilistic hypothesis about the likelihood that specific known schemas are underlying play. Here we hypothesise that poor overall correlations are likely to indicate the presence of previously uncharacterized schemas, which ideally should be returned to a central schema repository for the game for analysis and new schema description development and distribution.

As described above, gameplay schemas represent the significant learning outcomes of gameplay and also encapsulate various rewards of play. An explicit representation of desired and observed gameplay schemas within a game system constitutes a knowledge base that can potentially be used automatically for a variety of purposes by the game system.

Schema representation and mapping can be used for automated monitoring of a game design as a method of validating the design in terms of player satisfaction. Since a schema includes various points of player reward and presents a time structure for the emotional experience of a game, the schema indicates a hypothesis about the nature of the emotional experience of the player. Different players may prefer different forms of satisfaction. Monitoring schemas and schema execution may indicate which forms of satisfaction particular players are seeking. It may also provide a foundation for determining when a player is not achieving enough satisfaction (based upon criteria that may be derived from a player’s history of play, since different players may have different demands in terms of the nature and intensity of rewards). This may be because a game is too easy or too difficult, in which case the game mechanics and parameters determining difficulty for a specific schema can be modified to achieve a better subjective experience. It may also be that a player has not discovered those elements of a schema needed in order for its execution to result in a satisfying experience, in which case the game system model of the player schema related to a target schema might be used to change the game mechanics, e.g. to dynamically adapt a level design or to introduce instructional material (perhaps by spawning a suitably informed NPC) to lead the player to actions (such as going to a specific training scenario) that result in gaining the appropriate skills. In effect this can amount to more efficient and dynamic use of in-game
tutorials together with an automation of the normal processes of game tuning carried out manually during game testing prior to release, but having the advantage of being tuned to specific players rather than a group of commercial testers.

Schema descriptions can also be used to explore the effectiveness of a game design in realizing designers’ intentions. Simple observations may indicate basic design failures, such as the visual design of interactive elements leading them to be too unobtrusive within the game space for players to notice. However, schema descriptions may show deeper and less obvious problems, such as design features leading too soon to limited modes of play that reward players too much for play patterns that are developed very quickly, discouraging them from exploring a game enough to discover other aspects of its mechanics. As with the other examples presented here, actively using these models within a game system allows the mechanics to be varied for individual players, instead of providing a single solution that is supposed to accommodate everyone.

Monitoring schema formation can also result in automated detection of the degree to which a game design is achieving emergent gameplay (see [10]) where the design rather loosely constrains the nature of the play experience. In this case, poor correlations with known schemas may be a positive indicator of emergent play. Conversely, design features may be selected that are compatible with a broad set of known schemas representing very different play styles, ensuring that a design accommodates a wide variety of play styles, a space within which players have a lot of freedom to create patterns of interaction.

A major use of explicit schema representations may lay in pedagogical or therapeutic functions of gameplay. In this case, target schemas may not be initially derived from gameplay but from the target application domains for learning or training. For example, in a military application, observation of tactical decision-making in the field could support the development of schema descriptions for tactical decision-making. A game for tactical training should then encourage players to preferentially develop the same or functionally similar schemas. The effectiveness of the design of a game intended for tactical training can then be assessed by comparing the schemas of players derived from observation of their play patterns with those of operational tacticians. This may be a great advantage compared to assessing performance outcomes, since performance outcomes alone only indicate how a player has mastered a game system, with no indication of how well the mastery of the game will transfer to an application domain. The schema description is an explicit representation of the cognitive capabilities that facilitate operational competence, thereby having much greater transfer potential from the game to the target application environment. Moreover, aspects of the operational schemas that cannot be facilitated by game design provide an explicit representation of the limits of transfer that may then be used to appropriately focus supplementary training.

Just as in the case of tuning game mechanics for player satisfaction, explicit schema representations and monitoring of player schemas can be used to adapt game mechanics
to achieve pedagogical or therapeutic outcomes. For example, a game designed to train players to achieve better spatial navigation skills might present an initial diagnostic level involving a comparatively complex navigation task based upon a variety of cues, such as verbal descriptions, minimaps, distance cues, and local cues like footprints and vehicle tracks. Based upon which cues players use, ongoing levels can reduce or exclude cues that are already taken into account and emphasise neglected cues to encourage the development of broader attention patterns.

Conclusion
This paper has described an approach to the analysis of gameplay based upon schema theory and attention theory. An empirically based method has been described as a basis for identifying and validating hypothetical gameplay schemas. Automated schema recognition and the potential uses of explicit schema representations within game systems have been explored. This approach provides for explicit modeling of the cognitive systems and processes underlying gameplay, both for analytical studies of play and as a potential implementation mechanism for adaptive games. Work on the analysis of games using this approach is ongoing. It is hoped that the results of this work will provide the foundations for future implementation of schema-based adaptive game systems.

Acknowledgements
This paper has been written in the context of the FUGA (FUn of GAming) EU research project, and a collaboration with the Swedish defence research institute, FOI. We thank our colleagues for many stimulating inputs to this project.
References


PART III
METHODOLOGY DEVELOPMENT


Eye Fixations in an Action Game Tutorial

Charlotte C. Sernessen, Craig A. Lindley
Gotland University College
Cramergatan 3, Visby, SE-62157 Sweden
charlotte.sennersten@hgo.se, craig.lindley@hgo.se

Abstract
Game engagement can be analysed from the perspective of perceptual loading and cognitive capacity. Eyetracking is a technique that can be used within an experimental methodology for systematically studying focus of attention in game play according to eye fixation patterns. Initial results indicate significant differences between fixation behavior and subjective reports of attention.

Author Keywords: Eye-tracking, game design, cognition.

ACM Classification Keywords: H1.2. User Interfaces: Evaluation/methodology.

Introduction
Computer games have a high capacity to engage players and represent perhaps the most successful examples to date of human-computer interface and interaction design. However, the nature of game engagement is poorly understood, with designers referring to a very vague concept of the ‘fun’ of games. A deeper understanding of design principles requires better theories of game engagement, which in turn requires well developed methodologies for the empirical analysis of gameplay. A basic hypothesis of our work is that cognition research is perhaps the most promising field from which to develop well founded theories of gameplay, together with methods for the empirical investigation of gaming at the level of individual players. A particularly interesting approach for the study of engagement is the use and ongoing development of theories of cognitive capacity and perceptual loading. Gameplay is a situation of highly focused attention and it is important to start to understand how game tasks are structured cognitively and how attention patterns contribute to ongoing situation awareness within games of different kinds. Our initial choice of a methodology for investigating these issues involves the use of eyetracking systems, based upon the hypothesis that focalisation is closely involved in attention management.

This paper describes the initial experiments conducted within this framework and the results obtained.

Experimental stimulus
The study reported here was based upon a stimulus environment consisting of tutorial version 1.0 of the game Counterstrike (Valve/Sierra, 1998), a First Person Shooter. The
Counterstrike tutorial was chosen since it includes basic elements the player trains on in a single player mode, off line, providing a simplified scenario compared with on-line multiplayer gaming. The tutorial was also chosen since all subjects must have the same stimulus conditions, so data can be compared across subjects.

A limited version of the tutorial was used in the experiment in order to have simple and strongly goal-oriented player eye behaviors to investigate first. This particular tutorial takes place in a training camp where the player/subject practices different tasks or operative missions, including hostage rescue, bomb/defuse and escape (+ assassination, not included in the test). Two shooting simulation scenarios were also involved in the tests, with 5-6 moving targets in each and a smoke grenade simulation. Between the missions the player is moving in rooms and corridors. Each round of Counterstrike lasts from 3-15 minutes, but in this tutorial the player had unlimited time. Due to unspecified time limitations, the subjects moved quite rapidly.

The goal of the gameplay in Counterstrike is to demolish and to shoot as many terrorists or counter terrorists as possible, resulting in the players getting points or money. Counterstrike is a team game in which the individual goal and the goal of the team is to survive. Time is an important factor, and hence rapidness, together with tactics and strategy. A player cannot win in Counterstrike unless working with or for their team. The experiment, however, involves a single player scenario in order to isolate player goals with in-game tasks without the complexities of social coordination. The tutorial itself has a goal to provide learning and understanding of the environment as a preparation for multiplayer play. Hence it is designed for the creation of the cognitive maps and skills required for full game play, making the tutorial an interesting environment for investigation.

Counterstrike environments have a ‘naturalistic’ look. Players interact in a first person perspective, i.e. they have a ‘point of regard’ instrumentally as the film camera, with a more restricted view angle (the screen) compared to a normal situation. A player never sees their own character in the environment unless the character dies, at which point it is seen in a third person view. What the player sees and how they move in the environment is directed by the keystrokes pushed down on the keyboard, or by the mouse.

The experiments were designed to find out where and what the players direct their attention upon during gameplay. The subjects played individually in a free examination after preliminary six part instructions regarding keyboard use and audio cues that would be encountered within the tutorial. Eyetracking data consisting of fixations and saccades were recorded.

The experimental subjects were male students and gamers. One woman attended and had the same prior experience as the male attendants. Skilled players were preferred to eliminate influences of learning basic gameplay skills and to be able to concentrate experiments on the visual flow patterns of competent players in the game environment. The average age
of the subjects was 25 years. The subjects were told beforehand that the experiment was about gameplay and that the camera used would track their pupils.

**Eyetracking Experiment Design**

To measure the fixations of the subjects, the pupil-centre/corneal reflection method was used. This method determines the orientation of the eye from the relative locations of the pupil and one reflection from an infrared light shining on one of the eyes of the subject. The eye gaze system is designed to determine where a user directs attention on a computer screen; the coordinates the system produces can be interpreted as coordinates on a virtual plane in front of the subject. If a subject looks outside this virtual plane, the eye fixation is said to be ‘plane not found’ by the eyetracker. When the subject’s point of gaze remains inside a small circle (of 20 pixels diameter) on the virtual plane during three frames (taken at 20ms intervals) of the camera, the eye gaze system reported a gaze fixation. A calibration procedure was repeated until the error level was under 3.8mm (equivalent to of 20 pixels diameter) on the screen. Measurements used a head-mounted SMI iView ©eyetracker, which is a monocular 50 Hz pupil and corneal reflex video imaging system. The experimental set-up consisted of one table where the computer screen and the keyboard were placed. The subject was seated in front of the screen and the keyboard. Two speakers were placed on the floor underneath the table. Between the subject’s table and the tracking computer with screen, a partition was placed so as not to disturb the experiment.

**Experiment results**

The experimental results discussed here concern what subjects looked at (fixated upon) in the game environment. The data that was collected from the meeting with a hostage in a room showed 10% fixations upon the facial region and 90% upon the body. Initially subjects tended to look at the hands of the hostage. These results have opposite ratios to gestures in live human interaction and on video. Interestingly, in answering a questionnaire afterwards, 6 out of 8 subjects answered that they had looked only towards the facial region of the hostage.

**Analysis**

There may be a variety of reasons for this disparity between fixations and subjectively understood attention, including: 1) if a subject has played action games before, their attitude and pre-knowledge suggest that nothing will be gained out of the face region in regard to money or points. 2) time pressure combined with keyboard placement can lead to head-down looking tendencies. Since action is carried out from the keyboard, considerations of safety may also draw attention to that region. 3) mental avoidance of surprises may lead the player to look low on the screen for shadows and hazards. 4) eye behaviors that show in data are fixations upon visual information that has been processed in the fovea but not in periphery. Peripheral processing occurs but a fixation is considered to be most important and deepest information processed. Foveal vision is constrained to 2 degrees of the visual field while the peripheral view is 160/130 degrees in a natural
situation. Here it can be considered to be 20 degrees. The theory of visual attention claims that visual information is processed when a fixation occurs. However, the answers of the subjects may rely upon peripheral attention and answer ‘face’ as a reflection of social cultural behaviour reflecting a general belief in the greater importance of faces than the actual fixated objects.

**Conclusions and ongoing work**

This study has been a preliminary investigation of the use of eye-tracking as a methodology for investigating game engagement and the structure of attention during gameplay. The results obtained are interesting in revealing significant disparity between observed behavior and reported behavior. Ongoing work is being undertaken in the context of a more extensive theory of the cognitive foundations of game engagement. This work includes further development of the technology used for experimentation. A first step has been the real-time integration of the eye-tracking software with a game engine (the FOI HIFI engine) in order to automate object identification based upon fixation data, a process that can greatly speed up data analysis.

**Acknowledgements**

We thank Martin Castor, Johan Hedström and Björn Lindahl of FOI for experimental software development and Kenneth Holmqvist of the Cognitive Science Department, Lund University, Sweden for the supervision of the original thesis work reported in this paper.
Verification of an Experimental Platform Integrating a Tobii Eyetracking System with the HiFi Game Engine

Charlotte Sennersten, Jens Alfredson, Martin Castor, Johan Hedström, Björn Lindahl, Craig Lindley, Erland Svensson
### Report title (In translation)
Verification of an Experimental Platform Integrating a Tobii Eyetracking System with the HiFi Game Engine

### Abstract (not more than 200 words)
Playing a commercial PC or console game is a highly visual activity regardless of whether the purpose is entertainment or situated learning as discussed in the Serious Games field. If more information about the visual attention of the player can be recorded and easily analysed, important design information can be extracted. A range of different eyetracking equipment exists on the market and has been used in many studies over the years. However, very few studies describe dynamic stimuli involving the visual interaction of the user/player with a moving 3D object displayed on a computer screen. The reasons for this are that methods and software developed for eyetracking studies of static 2D stimuli are inappropriate for dynamic 3D stimuli, and manual analysis of dynamic 3D visual interaction is extremely time consuming. In order to address this, the authors have developed a software interface between the Tobii™ eyetracking system and the HiFi Game Engine for use in automated logging of dynamic 3D objects of gaze attention. This report describes the verification study performed to assess the performance of this integration between the eyetracker, logging tools and game engine. Detailed analysis shows effective results within the derived accuracy range, which is certainly sufficient for studies from a small scale to large scales necessary for extensive statistical analysis. The work presented in the report has been conducted in collaboration between FOI, Blekinge Institute of Technology and Gotland College.

### Keywords
Eyetracking, game, gameplay, design, visual attention, instrumentation, methodology, human factors, behavioural science, cognition, measurement

<table>
<thead>
<tr>
<th>Field</th>
<th>Value</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Report number, ISRN</strong></td>
<td>FOI-R--2227--SE</td>
</tr>
<tr>
<td><strong>Report type</strong></td>
<td>Methodology report</td>
</tr>
<tr>
<td><strong>Programme Areas</strong></td>
<td>8. Human Systems</td>
</tr>
<tr>
<td><strong>Month year</strong></td>
<td>February 2007</td>
</tr>
<tr>
<td><strong>Project no.</strong></td>
<td>E7096</td>
</tr>
<tr>
<td><strong>Subcategories</strong></td>
<td>81 Human Factors and Physiology</td>
</tr>
<tr>
<td><strong>Subcategories 2</strong></td>
<td></td>
</tr>
<tr>
<td><strong>Author/s (editor/s)</strong></td>
<td>Charlotte Sennersten (BTH/HGO)</td>
</tr>
<tr>
<td></td>
<td>Jens Alfredson</td>
</tr>
<tr>
<td></td>
<td>Martin Castor</td>
</tr>
<tr>
<td></td>
<td>Johan Hedström</td>
</tr>
<tr>
<td></td>
<td>Björn Lindahl</td>
</tr>
<tr>
<td></td>
<td>Craig Lindley (BTH/HGO)</td>
</tr>
<tr>
<td></td>
<td>Erland Svensson</td>
</tr>
<tr>
<td><strong>Project manager</strong></td>
<td>Erland Svensson</td>
</tr>
<tr>
<td><strong>Approved by</strong></td>
<td>Lena Bergvin</td>
</tr>
<tr>
<td><strong>Sponsoring agency</strong></td>
<td>Swedish Armed Forces HQ</td>
</tr>
<tr>
<td><strong>Scientifically and technically responsible</strong></td>
<td>Erland Svensson</td>
</tr>
<tr>
<td><strong>ISSN</strong></td>
<td>1650-1942</td>
</tr>
<tr>
<td><strong>Language</strong></td>
<td>English</td>
</tr>
<tr>
<td><strong>Pages</strong></td>
<td>54 p.</td>
</tr>
</tbody>
</table>
1. Introduction

At the “Serious Games Conference 06” in Washington DC, the content was divided into two major categories: War games and Simulations and ‘Situated Learning’ in Teaching and Training. The games under consideration provide these situated environments with their tasks and goals, which are provided either with instructions or not, and are of major interest for a broader target group than the entertainment industry.

Serious games are of interest for very specific effects and outcomes on their players in the form of various different kinds of learned skills. The ongoing improvement of the design and effectiveness of serious games requires an increasing understanding of the relationship between game design features, player characteristics, rewards of gameplay, gameplay interaction patterns and learning outcomes. Psychophysiological techniques provide a foundation for the detailed collection of data during real time play of games as an empirical foundation for this increasing understanding of play. Psychophysiology is a broad scientific field in which different types of sensors are used to measure the reaction and various states and behaviours of the human body to events in the world (or to experimental stimuli) and use the resulting data to make inferences about the mental state, emotional state and cognitive activity of the human player.

The psychophysiological perspective on game and simulation environments is of great importance in providing empirical foundations for the development of validated theories by which a designer (or an instructor) can make decisions in relation to game play design. Of course there are different purposes of gameplay but the foremost is to engage the participant and create strong motivation to play. Very often people ask the question of how much time players can spend on one game and also the question of why they are so immersed in this kind of a learning environment. Gamers might freak out to hear that their games are learning environments, but if the problem solving involved in game play is considered, we can understand that this is the case. The difference between entertainment games and serious games is that entertainment games are designed to reward people for playing as an end in itself, with no consideration of the transferability of the skills involved in other contexts. For serious games, skill transfer is a critical issue and the whole point behind using games for non-entertainment purposes.

The biometrics obtained from psychophysiological studies of game play can be used to provide important knowledge for many reasons. From a design perspective it helps to answer why to use specific design features and arrangements. The purpose can be to control the player’s physiological interaction to reach a certain interaction goal or to monitor the state of the player for training and instruction. Generally, the cognitive side of these interactions involves the development of specific skills allowing players to overcome game

---

28 Here we use ‘game play’ to designate play forms intended for a game design by its designers, and ‘game play’ to designate the actual play behaviour of players.
challenges that may include problem solving, tactics, spatial navigation, reaction time, logical reasoning and many other functions.

Today when situated learning is widely discussed, not just because of its effectiveness in learning but also because of economic efficiencies, the transferability of skills among virtual game environments, simulators and real life operative environments must be considered. To interact with a 2D display and a common keyboard, in relation to a simulator that emphasizes the real setup of devices with visual displays of different kinds and real world operation, can be questioned due to differences in sight/vision and motor interaction between games/simulators and operational environments. Gameplay biometrics can provide important data contributing to the answers to this question.

Brown\textsuperscript{29} points out three important factors concerning biometrics. First, that gameplay analysis which is based on biometrics can drive new instructor created missions/scenarios. Second, that increased avatar linkage to actual player state is achievable using biometric measurement technologies as feedback to a game system. Thirdly, the real time dynamics of gameplay can be adapted; if the time requirements of game interaction are adaptable system variables, then workload involved in play can be either a fun factor or a more severe one depending on the context.

The project reported in the present study contributes to the more efficient use of the biometric technique of eyetracking for the analysis of gameplay. Previous work by Sennersten (2004) demonstrated the usefulness of eyetracking studies of gameplay, but encountered the problem of the very long analysis times required; four hours of gameplay took three months to analyse manually. If the identification of objects under the point of gaze could be automated, however, this basic aspect of analysis could be achieved far more quickly. While the automated identification of objects by image processing of the resulting screen data together with the eye gaze data would be complex, a much simpler solution is possible, i.e. integrating the eyetracking system with the game engine and having the game engine log the objects under the gaze point during play. The implementation of this is conceptually and technically simple, since an \((x,y)\) coordinate representing a gaze point on the screen can be treated in the same way as an \((x,y)\) cursor position when a mouse button is clicked, ray tracing from the \((x,y)\) point to intercept the closest perceived virtual object within the 3-dimensional world synthesised by the game engine. It is then straightforward to modify the game engine code to ensure that the information about the object under the gaze point (i.e. its identity and position) is entered into a log file. In this way automated identification of game objects under the gaze point can be achieved and recorded in an ongoing log of a play session. Object identification under the gaze point

\textsuperscript{29} Notes from the presentation “Applying Biometrics to Assess Training In-Game and during AARs” by Randy Brown. (AAR stands for After Action Review) at Serious Games Summit Conference 2006, Washington DC, USA
then takes place automatically and in real time during play, taking no extra time at all than the play session itself.

This integration of an eyetracking system with a game engine, with real time logging of the object data, has been achieved by the project as reported in this technical report. The eyetracker used is the Tobii™ 1750 eyetracker owned by HGO, while the game engine is the HiFi Engine developed by the Swedish Defence Research Agency (FOI). The verification study reported in this document describes the system and the analysis process used to verify its correct operation. The results of the verification process are described and analysed, and the resulting precision model for the integrated system is presented.

2 Background

Designing and interacting with 2D worlds has its own rules and possibilities. Moving on to digital 3D worlds adds many questions in how to design for a view frustum (or viewing volume) potentially having 360 degrees of freedom around three orthogonal axes. Virtual worlds can also be arbitrarily large, so how should we design for a player’s attention within a space, which can be experienced as endless and open?

In order to understand how 3D worlds are constructed and perceived, one can refer to existing knowledge from the study of aviation and training simulators. Studies of human factors, situational awareness and workload have revealed a wealth of information on how to optimize and understand the interaction of humans and machines. This approach contrasts with the more prevalent approach to game design from a more general culture studies viewpoint, which has been dominant within international game research for the first few years of the millenium.

For example, in a helicopter there are also 360 degrees of freedom while rotating around the pilot’s own shoulder or around their own axes. The instrumentation in a cockpit is both mechanical and digital. Visual Flying Rules (VFR)* are formalized and trained regulations for a pilot. One can say that the operational 3D environments in these cases are primarily situated outside the helicopter. They are external, outside both the physical human body and the physical body of the machine. The rules within the regulations are guidelines for how to meet this physical/external environment and have to be understood and solved by a

* VFR and IFR are rules for how a pilot can and is allowed to act depending on daylight or darkness, weather conditions, the equipment of the aircraft, ground equipment and certification. Simply expressed, VFR rules indicate how a pilot, with his/her own eyes and the aircraft’s ground equipment, is allowed to perform with the aircraft within areas and aerodromes that have requirements on factors such as least sight distance and distance to clouds and obstacles (sight to ground). The cloud height cannot go under 600ft and sight not under 800m. For IFR, the decision height for “precision flight” is 200ft (60m) or lower for category II or III. The sight circumstance along an IFR path way is allowed to be 0m.
human pilot. Of course, these environments, especially in aerial work, also include surprises of different kinds like terrain conditions, animals, unpredicted winds, changing light, etc. Instrumentation Flying Rules (IFR), on the other hand, are those rules the physical human uses within the cockpit in conditions of poor or non-visibility of the external space outside the machine (e.g. in bad weather). Instrumentation within the helicopter can be considered to be read mostly as 2D graphics, as within a game (e.g. staminas, bars etc.), while the pilot’s motor actions are carried out through a joystick, thrust control and pedals, that could also be used to play a game although it is more typical for the player of a PC game to use a keyboard and mouse. The theory in this field is strongly connected to visual attention theory, association theory and also theories of neuro-psychophysiological demands on the operator or player. Previc (1998) describes this in an integrated theoretical model of how different brain systems mediate our perceptual-motor interactions in peripersonal space ((the region immediately surrounding our bodies) and the three major compartments of extrapersonal space (the focal, action, and ambient realms)).

To operate in a digital 3D space one has to decide where one should be and how to navigate depending on what objects are present and relevant in the digital world. These decision processes are both present in a pilot situation as well for a player in a gameplay session. To connect this awareness of space, either in a physical reality or in a virtual game reality, cognitive functions try to map (associate) and categorize these things depending on the current task. The mind builds these understandings partly in the “Where” and “What” centers in the brain (Nyberg, 2002). Castor et al., 2003 (p.112), state that the operating attribute in Military Aircraft Operations includes Flight and Mission Planning as a template for change. This means that the pilot has to plan the flight regarding distance. A central hypothesis of interest to the primary author of this report is that gameplay, depending on genre, is also carried out with decision-making based upon virtual distance, which could follow from real life conventions or vice versa.

It is against this background of the interrelationships between aviation operations, situational awareness, workload and computer gameplay that psychophysiological investigations of gameplay are of interest, particularly to the primary author of this report.

2.1 Reason for doing this project – the Research Project Context

As noted in the introduction, during a previous study on eye movements in an action game tutorial (Sennersten, 2004) the need for more efficient methods of data analysis became evident. Already at Lund University while working on a master’s thesis in Cognitive Science, Sennersten turned to FOI wondering about the eye-tracking methods used in previous studies. The focus on 2D environments in Lund did not meet Sennersten’s needs in saying anything about player’s interaction patterns in 3D game worlds. The game worlds Sennersten was thinking of were particularly First Person Shooters (FPS) and Role Playing Games (RPG’s). What could help would be to be able to have an automatically obtained relation between x and y gaze coordinates and logging of both objects (identification
and location) and events while playing to improve the efficiency of the analysis process, which otherwise is very time consuming. It is also a problem when making judgements by eye about the objects that are under gaze points identified by eyetracking, subjective interpretation is an unavoidable factor. “Clean” data is needed. The use of eyetracking to study gameplay is part of Sennersten’s ongoing research, which will be greatly facilitated by automating lower level aspects of the analysis process.

The Swedish Defence Research Agency (FOI) has over recent years conducted several studies in game related areas (Hasewinkel & Lindoff, 2002; Kylesten & Söderberg, 2001; Lindoff & Hasewinkel, 2004; Rencrantz, 2003; Wikberg, Hasewinkel, Lindoff, Eriksson, Stjernberger, & Persson, 2003). These studies have mainly studied the usefulness of game consoles and games in training as a basis for choice of learning methods. Another study from FOI by Svensson et al (1997) has shown that the frequencies of shorter fixation times (head-up, HU) and frequencies of longer fixation times (head-down, HD) increased as a function of the average information load on a Tactical Situation Display of two JA37 aircraft simulators. One aspect that could be studied more thoroughly as a foundation for assessing simulation-based training is motor and visual interaction across computer-, simulator- and real life missions: can there be gaze and behavioural reinforcements in between these different environments and what differences do they create in humans when interacting? When bringing in game related environments and a generation with pre-learned individual gameplay patterns, one also has to consider pre-established gameplay behaviours that could be more or less efficient in different contexts.

The foremost reason for doing this project has been to develop the integrated eyetracker/game engine system and to see how accurate this kind of system can be for application within the research in the areas described above. In general, a lot of effort goes into biometric measurement, but our most important organ, in relation to perceiving the environment around us and generating the highest perceptual data input, is the eye, which to date has been underused in biometry. Without distributed visual attention it is impossible to say a great deal about the meaning of other kinds of physiological or psychological measurements in relation to responses to an environment in which visual data is significant. A peak in heart rate variability, for example, with no reference to audiovisual stimulus can be hard, if not impossible, to interpret.

2.2 Review of eyetracking technology
Eyetracking is a method that has been used commonly over the last 40 years and given insight into how (visual) attention is distributed when carrying out different tasks. The technique involves tracking eye movements and recording the distribution of gaze over time. Eyetracking techniques have been used in many fields including neurology, cognition, linguistics, branding, advertisement in newspapers and the like, car design, art, media communication, security of different kinds, etc..
There are four broad categories of eye movement measurement methodologies (Duchowski, 2003): electro-oculography (EOG), scleral contact lens/search coil, photo-oculography (POG) or video-oculography (VOG), and video-based combined pupil and corneal reflection. Today the latter measurement is the most widely used because of ecological validity and ease of use.

EOG relies on measurement of the skin’s electric potential differences, using electrodes placed around the eye. Scleral contact lenses and search coils provide the most precise eye movement measurements. The search coil is embedded in the contact lens so the electromagnetic field frames can be placed directly onto the eye (Eye and Vision at Bernadotte Laboratory at Karolinska Institutet has these devices). POG measures the eyes under rotation/translation (the shape of the pupil) using corneal reflections of a close directed infrared light. The video-based technique with corneal and pupil reflection demands the head to be steady so the eye’s position in relation to the head and the point of regard coincide. This last technique has also been integrated into a variety of devices including the head-mounted eye tracker, which does not just track the screen but also gaze behaviour outside the screen area.

It is important to distinguish between techniques for measuring eye movements and techniques for measuring eye-point-of-gaze (EPOG), also called “point of regard”. Eye movements may be measured to determine where someone is looking. Only if the absolute orientation of the eye in relation to the surrounding objects is known can the EPOG be determined. Eye movements could, however, be measured for other purposes than determining the EPOG. Often it is not only interesting to know where a person is looking, but also how the person is looking. For instance, can eye-movements tell us something about the mental state of an operator, such as the mental workload? For this purpose it is not only interesting to measure the orientation of the eye, but also measures such as blinks and pupil diameter.

A user uses both foveal vision and peripheral vision to gather visual information. With foveal vision, which is about two degrees of visual angle, the user can get a detailed look at a certain point, the eye point of gaze. The movements of the eye that maintain fixations are called fixational eye movements. To make the image of a constant moving object stay in the fovea, pursuit movements are made by the eye. Movements of the eye that are jumping from one point to another are called saccadic movements. Vestibuloocular eye movements are those compensating for the movements of the head. Optokinetic eye movements are used when a large part of the visual field is rotating. Finally, vergence or disjunctive eye movements are used, for example, to follow a target that is getting closer or further away, by moving the eyes in opposite directions to one another.

The time that the eye is not moving, but staying in the same position, is called fixation time or dwell time. Fixation times are used to evaluate computer displays or other interfaces.
However, it is not always obvious how to interpret the meaning of a long fixation time. For example, important information can make the user fixate upon the information for a long time, as well as badly displayed and confusing information. By comparing the fixation times to other data it can sometimes be possible to draw valuable conclusions from the data. In active vision or overt attention, sampling is achieved by a fixation-move-fixation rhythm and this pattern can be found in the vision of humans, most other vertebrates and some invertebrates (Land, 1995; Land and Nilsson, 2002). Passive vision is also referred to as peripheral vision or covert attention, with the meaning of attending without looking. Peripheral vision increases in degree outside the 2-degree high-resolution foveal centre. Findlay and Gilchrist (2003) point out that overt attention, via the fovea, plays the major role in visual attentional selectivity.

It has long been believed that the eyes and eye-movements can tell us some things about the human mind; when looking at a stranger for the first time, it is possible to see whether he or she is frightened, sad, tired etcetera simply by looking at the persons eyes. In recent years, eyetracking technology has been developed to provide competent aids in determining user interaction with technology. Computer interfaces have been evaluated with eye movement data in several studies. Graf and Krueger (1989) used eye movement data as a performance measure for evaluation of an alphanumeric display, as well as a measure of the cognitive load on the user. Lankford, Shannon, Beling, McLaughlin, Israeliski, Ellis, and Hutchinson (1997) have used eye gaze tracking in the design of telecommunications software, arguing that “...graphical interfaces can benefit substantially if eye tracking data on a user’s visual interaction with the software is considered as a part of the design process. This eye tracking data would supplement the more traditional human performance data (e.g. timing and errors gathered during usability testing)”. The combination of eye movement data with other data sources, to get a more complete picture of what is happening, is a promising area deserving much more investigation.

Eye movement measures can be used both in the design of a new system and in the later evaluation of the same system. Thus, the purpose of using eye movement data to evaluate interfaces could be a concern both for the better design of new interfaces and for ensuring that interface designs are really effective by later testing.

2.2.1 Comparison of specific eyetracking systems
Different eyetrackers have different qualities suited to different purposes. The sampling rate, for example, can be a major issue when choosing equipment for a particular study. A relative assessment of various systems was made by researchers at FOI to get an overview of their pros and cons in relation to purpose and goals for different studies. The table below is derived from the technical report presenting the results of this assessment (Alfredson, Nählinder & Castor, 2004), where five different systems/techniques, all of which FOI has experience with using, are compared. In this report the properties of the Tobii eyetracker used in this study are added to the table.
<table>
<thead>
<tr>
<th></th>
<th>GazeTracker</th>
<th>EOG</th>
<th>JAZZ</th>
<th>Smart Eye</th>
<th>Video based</th>
<th>Tobii 1750</th>
</tr>
</thead>
<tbody>
<tr>
<td>Measurement principle</td>
<td>Differential corneal reflection and electromagnetic head-tracking</td>
<td>Corneoretinal electrical potential</td>
<td>IR-pupil tracking</td>
<td>Optical feature recognition</td>
<td>Optical video</td>
<td>Optical feature recognition</td>
</tr>
<tr>
<td>Maturity</td>
<td>High</td>
<td>High</td>
<td>Low</td>
<td>Medium</td>
<td>High</td>
<td>Medium</td>
</tr>
<tr>
<td>Accuracy</td>
<td>High</td>
<td>NA</td>
<td>NA</td>
<td>High</td>
<td>Low-medium</td>
<td>High</td>
</tr>
<tr>
<td>Max range horizontal</td>
<td>Medium. Limited by eye movements</td>
<td>Medium</td>
<td>Medium</td>
<td>High. Limited by head movements</td>
<td>Unlimited for head</td>
<td>High, Limited by head movements</td>
</tr>
<tr>
<td>Max range vertical</td>
<td>Medium. Limited by eye movements</td>
<td>Medium</td>
<td>Medium</td>
<td>Medium. Limited by head movements</td>
<td>Unlimited for head</td>
<td>Medium. Limited by head movements</td>
</tr>
<tr>
<td>Head movements</td>
<td>Yes</td>
<td>No</td>
<td>Yes</td>
<td>Yes</td>
<td>Yes</td>
<td>Yes</td>
</tr>
<tr>
<td>EPOG</td>
<td>Yes</td>
<td>No</td>
<td>No</td>
<td>Yes</td>
<td>Yes</td>
<td>Yes</td>
</tr>
<tr>
<td>Time resolution</td>
<td>60 Hz</td>
<td>Optional</td>
<td>1000 Hz</td>
<td>30-60 Hz</td>
<td>Low</td>
<td>10-50 Hz</td>
</tr>
<tr>
<td>Fixation duration</td>
<td>Yes</td>
<td>Yes</td>
<td>Yes</td>
<td>Yes</td>
<td>No</td>
<td>Yes</td>
</tr>
<tr>
<td>Scan paths</td>
<td>Yes</td>
<td>No</td>
<td>No</td>
<td>Yes</td>
<td>Yes, but very time consuming</td>
<td>Yes</td>
</tr>
<tr>
<td>Saccade velocity</td>
<td>Yes</td>
<td>No</td>
<td>Yes</td>
<td>No, not implemented</td>
<td>No</td>
<td>No</td>
</tr>
<tr>
<td>Eye blink detection</td>
<td>Yes</td>
<td>Yes</td>
<td>Yes</td>
<td>Yes</td>
<td>Yes</td>
<td>Yes</td>
</tr>
<tr>
<td></td>
<td>Yes</td>
<td>No</td>
<td>No</td>
<td>No, not implemented</td>
<td>No</td>
<td>Yes</td>
</tr>
<tr>
<td>--------------------------------</td>
<td>--------------</td>
<td>-------------</td>
<td>-------------</td>
<td>---------------------</td>
<td>-------------</td>
<td>--------------</td>
</tr>
<tr>
<td>Pupil diameter</td>
<td>Yes, optional</td>
<td>No, not implemented</td>
<td>Yes</td>
<td>Yes</td>
<td>No</td>
<td>Yes</td>
</tr>
<tr>
<td>Eye activity (EME)</td>
<td>Yes</td>
<td>Yes</td>
<td>Yes</td>
<td>Yes</td>
<td>Yes</td>
<td>Yes</td>
</tr>
<tr>
<td>Costs</td>
<td>High</td>
<td>Medium</td>
<td>Medium</td>
<td>High</td>
<td>Low, but high analysis costs</td>
<td>Medium</td>
</tr>
<tr>
<td>Reliability</td>
<td>Medium</td>
<td>Medium</td>
<td>Medium</td>
<td>Medium</td>
<td>Medium</td>
<td>Medium</td>
</tr>
<tr>
<td>Intrusiveness</td>
<td>High</td>
<td>Medium</td>
<td>Medium</td>
<td>Low</td>
<td>Low</td>
<td>Low</td>
</tr>
<tr>
<td>Type of data</td>
<td>Eye movement and EPOG</td>
<td>Eye movement (Saccades)</td>
<td>Eye movement and EPOG</td>
<td>EPOG</td>
<td>Eye movement and EPOG</td>
<td></td>
</tr>
<tr>
<td>Interference from surroundings</td>
<td>High (Electromagnetically)</td>
<td>Low (Muscle activity)</td>
<td>Medium (Infra red illuminatio n)</td>
<td>Low (Illumination and picture background)</td>
<td>Medium (Illumination and picture background)</td>
<td></td>
</tr>
<tr>
<td>Ease of use</td>
<td>Low</td>
<td>Medium</td>
<td>High</td>
<td>Medium</td>
<td>High</td>
<td>Medium</td>
</tr>
<tr>
<td>Preparation time/calibration time</td>
<td>High</td>
<td>Low</td>
<td>Low</td>
<td>Low, A couple of hour’s installation if moved to new setting, Low calibration time.</td>
<td>Low</td>
<td>Low</td>
</tr>
<tr>
<td>Analysis time/cost</td>
<td>Medium</td>
<td>Medium</td>
<td>Medium</td>
<td>Medium</td>
<td>Medium</td>
<td>Medium</td>
</tr>
</tbody>
</table>
The “High”, “Low” and “Medium” values in the table do not have any scientific thresholds; they express qualitative judgements based upon the experience of the FOI researchers and authors of this paper, and could possibly be ranked differently by another group. The maturity of experience in different studies nevertheless gives the table a high trust value as overview information.

3 Experimental platform specification
The project that this document describes uses the Tobii eyetracking system since it has EPOG, blink detection, is robust against head movement, tracks both eyes for higher reliability, can be calibrated quickly, has comparatively good spatial accuracy, an adequate sampling rate, and an API for the server that supports convenient integration with the game engine.

The HiFi Engine was selected since the source code is accessible and FOI own the engine and are therefore not constrained in its use. Features and performance of the engine are very good and it is an engine for first-person shooter games, which are of interest to the researchers involved. Since FOI developed the engine, familiarity with the engine also makes it very easy to modify, thus facilitating software integration with the eyetracker.

3.1 Technical architecture - HiFi Engine
The platform used in the experiment is based on the HiFi Engine (HFE), a game/simulation engine developed at Man-System Interaction - Swedish Defense Research Agency (FOI) in Linköping, Sweden.

3.1.1 Open source components
HFE is developed around several Open Source components that take care of specific tasks such as physics, sound, network communication, etc.. Besides wrapping in several open source components, there is a core of functionality that glues everything together and adds some specific functionality to the simulation engine. This is functionality that is specific for the simulation area that HFE is mostly used for at FOI.

HFE itself is developed as a simulation component that could be used for different applications. It could be used for pure simulation purposes like simulating a specific vehicle, or for more game-related simulations such as the context of this report, or in a completely different type of application such as a command and control application. In the latter case HFE can take care of everything from visualization to symbol handling, including move, add and remove symbols, send symbols to other systems, logging etc. (e.g. see the report for the project “Cognitive overview” where CoMap, Cognitive Map is described, which is a C&C application for urban combat, Kylesten, 2004).

With HFE it is possible to run one stand-alone client or to create a bigger simulation with many clients. This could be very useful when looking at, e.g., effects of solving situations in
a cooperative manner. HFE can also enhance simulations with its excellent visualization and in-depth vehicle physics simulations etc.\(^3^0\)

![Diagram of the HFE architectural hierarchy.](image)

**Figure. 1.** The HFE architectural hierarchy. An application is built using HFE, which in turn uses open source components that handle some functionality in the simulation engine.

### 3.1.2 The Test Application
A basic FPS (First Person Shooter) game has been built for use as the verification stimulus game for the present study that utilizes some of the advantages that HFE provides. Some of the things implemented in the FPS are basic weapon handling, picking up ammo boxes and similar triggers to pick up health-increasing objects.

### 3.1.3 Logging
During execution of the experiment, interesting perceptual objects within the game world, as indicated by player gaze direction, are logged in a file. The logging data consists of structures with one entry per gaze sample, where a gaze sample is taken from the eyetracker server with every frame generated by the game engine, corresponding also to a simulation time tick. The data logged includes:

- object identifier, consisting of name and \((x,y)\) position of the object model origin upon the virtual ground
- time of the gaze sample

### 3.2 TET Server and TOBII eyetracker
The Tobii Eye Tracker 1750 runs on a PC with the Windows Operating system (Windows 2000 service pack 3 or Windows XP service pack 1 or 2). The system can either be used in single or double computer setups. In this study a game application is running at the same time as eyetracking, which is most likely too much processing to do on one computer, hence in this study we use the double computer and double screen setup. One computer is running the eyetracking software ClearView (2.6.3) and the TET eyetracking server. The other computer is running the game application. The two are connected via TCP/IP. The TET server gets the data from the eyetracker and sends it to the ClearView software. For

\(^3^0\) For further information about HiFi Engine, contact Johan Hedström, johon@foi.se.
calibration of the eyes of the player, the operator switches the eyetracker display over to the second screen of the operating computer that runs the Clearview software, for display of the calibration pattern (hence a dual head graphics card is needed on the operating computer for the double display setup). The output from the Game Computer goes through a splitter where two video outputs are created, one going to the video capture card on the operator computer, where the captured video data is input to ClearView, while the other output goes to a switch that selects either the game display or the operator computer second monitor display to send to the eyetracker display. The routing of the video display is illustrated in Figure 2.

![Diagram of data interconnection](image)

*Figure 2. Display routing.*

A typical play session using this system involves the game player playing the game in front of the eyetracker. After briefing the player with any instructions relevant to the study, the calibration process is run, calibration patterns being displayed on the eyetracker screen while Clearview uses the known positions of elements in the display to calibrate eyetracking for that specific player. Then the eyetracker screen is switched over to the game display via the routing switch. The operator must start game object logging within the game engine using a command line instruction, at which a synchronisation message is sent by the game engine via the TET server to Clearview to commence eyetracking. Synchronisation is necessary to simplify time correlation of object log entries on the game computer with video frames captured by the operating computer. At an instruction from the operator, the player then commences playing while eyetracking and game engine object logging are running.

During game play, with each simulation tick or game display frame the HiFi (game) engine requests a gaze data sample from the Clearview software via the TET server. The game computer then traces a ray from the \((x,y)\) coordinates of the gaze point to intercept the first collision mesh within the game world under that \((x,y)\) point, corresponding with the object being looked upon by the player. Within the game engine, a log entry is generated
that includes the name of the 3D object model appended with the (x,y) coordinates of the model instance on the horizontal plane of the 3D game world, ensuring that each object instance is uniquely identified, together with a time stamp generated by a trigger signal from the eyetracker collected with a high-resolution timer, based on Windows multimedia timer. The log record is then appended by the game engine to the object log file stored on the gaming computer. When the goals of the session have been met, the operator asks the player to stop playing. A command line instruction on the gaming computer is then used to turn off the object log and eyetracking.

The Tobii eyetracking system detects and collects eye gaze data at a rate of 50 samples per second, with frame grabbing of the corresponding game display at a lower but selectable frame rate. The eyetracking system monitors the eyes of the subject/participant based upon reflected infrared light patterns and calculates gaze positions automatically, with hardware in combination with advanced software algorithms. The system has head motion compensation and low drift effects, with binocular tracking, which yields higher accuracy then just a monocular method. Another advantage of this system is that if the participant happens to move, especially the head, tracking is not lost if only one eye cannot be tracked during the movement. (With other systems, the participants may have helmets, headrests or markers, with different respective freedoms and constraints of movement.) For ecological validity, the eyetracking camera is implemented at the bottom part of a normal sized screen. The Tobii 1750 system can only track what is presented on the screen and not anything outside the screen area. The calibration time is only 3 minutes if no problems occur due to eye disorders, blinking, etc., The recommended eye distance to screen is approximately 60 cm. The accuracy of an eyetracker is measured in degrees. The accuracy or bias error of the Tobii 1750 system has been tested over a set of individuals to 0.5 degrees of visual angle when using standard accuracy measurement principles for eye trackers. One degree of accuracy corresponds to an average error of about 1 cm between the measured and intended gaze point at 60 cm distance from the user to the screen. The tolerance of head movements is about 10 cm sec⁻¹.

The light condition requirements for eyetracking are generally the same as when using an ordinary computer display. Near Infrared light (NIR) such as components of sunlight are not recommended to be present while tracking the eyes, since this can cause disturbances in the tracking. The eyes and especially the pupils are very light sensitive, so it is also best to avoid direct and/or strong light from the study environment.

For both eyes (left and right) the following data is available derived from the eye tracker data processed by the Clearview software and TET Server:

- Time (The timestamp is in microseconds),

- Screen X (horizontal position),
- Screen Y (vertical position),
- Cam X (horizontal location of the pupil in the camera image),
- Cam Y (vertical location of the pupil in the camera image),
- Distance (distance from camera to eye),
- Pupil (Size of pupil in mm) and
- Code (validity of gaze data).

Validity codes ranging from 0-4 and are logged for each eye with every gaze data point. 0 is the highest validity value and means that all relevant data for the particular eye is correct.

The frame rate and timing are crucial when this system is connected with a game engine. The game engine has its own tick-by-tick simulation update rate, the eye tracker has a different sampling rate and the real time video capture has another time rate again. The sampling rate of the 1750 eyetracker, representing the rate of gaze point data collection by the system, is a constant rate of 50Hz. The display frame update rate of the game engine varies depending upon the complexity of the visual scene, generally being between 25-35 ms. Video data capture was set at 10 fps (frames per second) for the present study. These different rates must be correlated during analysis in order to obtain the most accurate correspondences between gaze data, game frames and video frames, although due to variable latencies in the system the correlations are never perfect. This is analysed in detail in later sections of this document. Latency is defined as the time taken from the generation of a signal to its reception, in this case including time taken for eyetracking camera exposure, transfer to the TET server, calculation and delays in the server, transfer to Clearview, request generation from the game engine, processing time in the TET server, and response propagation back to the game engine.

![Sampling time of logging, gaze and video capture](image)

**Figure 3.** The start time from the Game Engine has to be synchronized with the Real Time Video Capture due its delay. The figures above show the internal relationships in this study.
4 Description of the Verification Procedure

The verification study reported in this document concerns verification of the effectiveness and accuracy of the integration of the Tobii 1750 eyetracker with the HiFi game engine (HFE). This concerns the basic success of the software methods used to perform the integration as well as the effects of various potential sources of error in both spatial accuracy and temporal accuracy. The essential method of the verification procedure is to conduct a study using the integrated system, then to conduct an analysis of the objects under the gaze point by manual interpretation of the captured game video data with superimposed gaze data, and then to compare the results of this visual analysis with the object log produced by HFE. Detailed analysis of the comparative data is then used to obtain an accuracy model of the integrated system that provides the basis for interpreting the accuracy of automated log data obtained from ongoing studies based upon the integrated system. The accuracy model is the significant analytical result of this project.

Note that the first verification procedure showed that the integrated system did not work due to faults that were later found in the implementation of ray tracing within HFE. These faults were quickly fixed and the verification procedure was repeated to undertake the analysis presented here.

Various software packages are used for the different stages of the process. They are:

- **BattleCraft™** – Stimulus
- **ClearView™** – Videocapture and Video output generation with superimposed gaze data
- **Virtualdub™** – Analysis/Transcription
- **Excel™** – Statistics/Transcription

The rest of this section presents an overview of the steps of the verification procedure. These steps are then exemplified in the subsequent section, followed by a section containing a detailed discussion of issues observed within the verification process and data. The final section then presents the accuracy analysis and model based upon the data and the verification process.

4.1 Steps in the verification process

A. Prepare Stimulus

The 3D stimulus is created using a commercial, off the shelf game level editor. Most popular 3D games (e.g. HalfLife, Oblivion) have these editing tools to allow players to create and
play their own game levels. ‘Mods’ are also possible, allowing players to create their own game content (media assets), worlds and simulations. Integrating a game engine with an eyetracker as described in this study, however, requires access to the source code of the game engine. Source code is typically not freely available unless it is an open source game engine. For the project reported here, FOI has access to the HFE code and HFE can import game levels from a commercial level editor, making the editor a convenient tool for stimulus development.

B. Select Player Participant(s)

The criteria for selecting players depends upon the study being conducted, the hypotheses of interest, etc. It may be important to consider player characteristics, since variability of play style preferences, tastes, etc. can have a major impact on study results. One early example of a system for categorizing players is that developed by Bartle (1996). Bartle’s four main player categories are: Achievers, Explorers, Socialisers and Killers, each reflecting differences of in-game behaviour. Also there are numerous psychological tests that one can consider for categorising players to find character and personality differences that may correlate with statistical differences in gameplay patterns. The latter tests may have to be carried out by authorized and certified psychologists in some countries if they need to be approved by ethical authorities. Different demographic categories of players could also make a difference to their gaze behaviour during gameplay, including: Age, Gender, Novice/Average/Skilled player, How long the participant has been playing games, What games the participant has been playing, Certain tactics or strategies habitually used by the player, etc. A standardized eye test should preferably also being carried on participants since eye conditions can influence the results in eyetracking.

In a military context, when there are often homogenous trained groups, the emphasis of interaction testing can be on to see how a particular task is carried out or on the usefulness of specific equipment. In a non-homogenous group the categorisation of player types is more important, so a baseline can be created. A baseline is a so-called neutral interaction pattern, which can be related to individual differentiations among players.

Despite these concerns with player characterization in experimental studies, the verification procedure is not concerned with the analysis of play patterns as such. Hence no questionnaires or observations that would be used in a full-scale experiment were used in this verification study. Similarly, biometric measures can say very much about temporal states within a player, one purpose with eyetracking being to relate biometric responses to what is represented on the display, similarly to traditional stimuli-response-result learning (e.g. Pavlov’s saliva test, conditioned reflexes, Eysenck, 2000). However, in this verification study only eyetracking data is considered.
C. Log Data

The sequence of operations for undertaking data logging is as follows:

i) Operator: Start Clearview software
ii) Operator: Start game using HiFi engine.
iii) Operator: Set player start point within game level (HiFi command line instruction).
iv) Operator: Spawn player character within game level (HiFi command line instruction).
v) Operator: Initiate logging by game engine (HiFi then starts eyetracking via message to TET server; HiFi command line instruction)
vi) Player: Start playing, play to objective.
vii) Operator: When player reaches objective, end logging and eyetracking via game engine (HiFi command instruction)

The resulting logs and data include:

- Object log from the game engine
- Clearview eyetracking data
- Captured video of the play session with superimposed gaze data, also generated by Clearview

Figure 4 shows an example of data from the object log created by the HiFi engine. For the verification study, the HiFi engine logged in total 13285 rows in a 5 minutes play session.

![Figure 4. Object log by HiFi Engine.](image)
D. Create Transcription

A transcription here means a textual description of what objects are judged by the analyst as being visually fixated in the captured video of the play session. This requires the analyst going through the video frame by frame and making a transcript record for each frame. The video is displayed using the VirtualDub software, which allows each frame to be accurately examined.

CENSORED

Figure 5. The software VirtualDub version 1.6.15 used for manual transcription.

A transcript record includes:

- Time of video frame
- Object name
- Gaze behaviour (e.g. new fixation, extended fixation)
- Object(s) interpreted by the analyst as being under the gaze point

These records are entered into a spreadsheet to facilitate further analysis. Figure 6 shows an example of the information from the avi file transcribed into the spreadsheet (note that data in column H is not used for the analysis).

<table>
<thead>
<tr>
<th>E</th>
<th>F</th>
<th>G</th>
<th>H</th>
<th>I</th>
</tr>
</thead>
<tbody>
<tr>
<td>1,000 terrainobject</td>
<td>new fixation</td>
<td>2,62</td>
<td></td>
<td></td>
</tr>
<tr>
<td>1,100 terrainobject</td>
<td>fixation grows</td>
<td>2,62 The dot is at SW</td>
<td></td>
<td></td>
</tr>
<tr>
<td>1,200 terrainobject</td>
<td>fixation grows</td>
<td>2,72 The dot is at SW</td>
<td></td>
<td></td>
</tr>
<tr>
<td>1,300 terrainobject</td>
<td>fixation grows</td>
<td>2,88 The dot is at SW</td>
<td></td>
<td></td>
</tr>
<tr>
<td>1,400 terrainobject</td>
<td>fixation grows</td>
<td>2,92 The dot is at SW</td>
<td></td>
<td></td>
</tr>
<tr>
<td>1,500 TR_Amman_829_948</td>
<td>fixation grows</td>
<td>3,07 The dot covers the</td>
<td></td>
<td></td>
</tr>
<tr>
<td>1,600 TR_Amman_829_948</td>
<td>fixation grows</td>
<td>3,12 The dot is covering</td>
<td></td>
<td></td>
</tr>
<tr>
<td>1,700 TR_Amman_829_948</td>
<td>fixation grows</td>
<td>3,23 The dot covers the</td>
<td></td>
<td></td>
</tr>
<tr>
<td>1,800 suburbhouse_2_close_n1.039.977</td>
<td>shift fixation grow</td>
<td>3,31 The dot is mainly</td>
<td></td>
<td></td>
</tr>
<tr>
<td>1,900 suburbhouse_2_close_n1.039.977</td>
<td>new fixation</td>
<td>3,43</td>
<td></td>
<td></td>
</tr>
<tr>
<td>2,000 suburbhouse_2_close_n1.039.977+suburbhouse new fixation</td>
<td>3,52 Hard to determine</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>2,100 suburbhouse_2_close_n1.039.977+suburbhouse fixation grows</td>
<td>3,67</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>2,200 suburbhouse_2_close_n1.039.977+suburbhouse fixation grows</td>
<td>3,86</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>2,300 suburbhouse_2_close_n1.039.977+ground fixation grows</td>
<td>3,83</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>2,400 suburbhouse_2_close_n1.039.977+ground new fixation</td>
<td>3,93</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>2,500 suburbhouse_2_close_n1.039.977+ground fixation grows</td>
<td>4,03</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>2,600 suburbhouse_2_close_n1.039.977+ground+an fixation grows</td>
<td>4,12 Hard to determine</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>2,700 TR_Amman_829_948+suburbhouse_2_close_m1.039.977+fixation grows</td>
<td>4,22 Among cross</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>2,800 TR_Amman_829_948+suburbhouse_2_close_m1.039.977+fixation grows</td>
<td>4,32 Among cross</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>2,900 TR_Amman_829_948+suburbhouse_2_close_m1.039.977+fixation grows</td>
<td>4,42 Among cross</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>3,000 TR_Amman_829_948+suburbhouse_2_close_m1.039.977+fixation grows</td>
<td>4,52 Among cross</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>3,100 TR_Amman_829_948+suburbhouse_2_close_m1.039.977+fixation grows</td>
<td>4,62 Among cross</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>3,200 TR_Amman_829_948+suburbhouse_2_close_m1.039.977+fixation grows</td>
<td>4,72 Among cross</td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Figure 6. An examplification of transcription in the spreadsheet, see column “F.”
Object identifiers are first annotated by the analyst on a printed top-down view of the game level obtained from the level editor or from within the editor on a second screen to help identify objects on the video and to work out the camera orientation within the game world (Figure 7). The identifiers have to be marked out manually because there is no automation of this.

![Image](image.png)

Figure 7. A “print screen” of dual screen mode while transcribing from VirtualDub. The first-person view to the left and the top-down-view to the right. Source: VirtualDub and Editor “BattleCraft”.

**E. Synchronise record times from the Game Engine Object Log and Video Transcription**

The objects in the object log created by the game engine need to be correlated with the corresponding records in the transcribed information from the avi file. The procedure used to do this is:

- Insert the object log entries in the same spreadsheet as the video transcription
- Search through the video to find frames representing distinctive new objects of fixation
- Find the corresponding log entry changes in the object log data
- For those correspondences that are found, look at the time difference for each of those records between the object log and the video transcript
- Average those time differences to obtain a best estimate of the starting time differences between the object logging and the video data capture
- Create a new (hypothetically) synchronized object log time column by modifying the frame times in the original log data by the offset time value found in the previous step
For the verification procedure reported here, 7 match points were selected for obtaining an average value as a basis for time synchronisation. The 7 locations that were picked for the study are objects that the player participant has not looked at before until that point in the game session. The 7 different match points used, with corresponding time differences, are:

<table>
<thead>
<tr>
<th>Match</th>
<th>AviFrame</th>
<th>GameEngine_Start</th>
<th>Time difference</th>
</tr>
</thead>
<tbody>
<tr>
<td>1st</td>
<td>4,000</td>
<td>5,515,037</td>
<td>1,515,037</td>
</tr>
<tr>
<td>2nd</td>
<td>8,100</td>
<td>9,674,996</td>
<td>1,574,996</td>
</tr>
<tr>
<td>3rd</td>
<td>172,500</td>
<td>174,026,64</td>
<td>1,526,642</td>
</tr>
<tr>
<td>4th</td>
<td>234,900</td>
<td>236,446,518</td>
<td>1,546,518</td>
</tr>
<tr>
<td>5th</td>
<td>248,200</td>
<td>248,395,844</td>
<td>0,195,844</td>
</tr>
<tr>
<td>6th</td>
<td>284,400</td>
<td>286,175,049</td>
<td>1,775,049</td>
</tr>
<tr>
<td>7th</td>
<td>360,700</td>
<td>362,234,406</td>
<td>1,534,406</td>
</tr>
</tbody>
</table>
The average value used for synchronization is taken from the 1st-4th match values and the 7th value. The two other values are taken away because the records are far enough away in time from the other values to be regarded as aberrant. The average time difference works out to be 1,539520 s. This positive number represents the time period for which the object log was active before video capture commenced. Hence it must be subtracted from object log times in order to synchronise the object log entries with the video transcript entries.

Once this average is obtained, a new column is created in the spreadsheet with values representing the original object log times modified by this value as the offset representing the difference in start time between object logging and video capture. The new object log values represent the best estimate of times since the same starting time as the video capture. This is illustrated in Figure 8, where column C is the original object log time and column D is the start time synchronised with video capture. Note that the negative values in column D indicate that object logging commenced prior to video capture, indicating a delay from when the start of eyetracking message is issued by the game engine to actual start of eyetracking and video capture by the Tobii system.

![Figure 8. Column C is original time and column D is new time concerning start delay for the video capture.](image)

**F. Run Sorting Macro**

The spreadsheet resulting from step 5 above includes the object log data alongside the video transcript data, including the offset time values for the object log data. For the sake of comparing the object log data with the video transcript data, it is then very convenient to sort these two data sets into a single time ordered list. This can be done automatically by running a macro written for this purpose in the spreadsheet. An example of the result is shown in Figure 9 below. Since the object log samples are taken every 25-30 ms and the video frame rate gives a frame interval of 100 ms, there are typically three object log entries between each video transcript entry. Sorting in the verification study resulted in a total of 17064 rows of object log and transcript information. Since this is the sum of the number of object log rows and video transcript rows (3780 + 13285), it may be inferred that the calculated times never exactly match (i.e. there is no single row, corresponding to a time from start, at which there is both an object log entry and a video transcript entry).
G. Detailed analysis of sorted data

The sorted list is inspected to identify three categories of correlation information of the object log with the video transcript that describe the accuracy of the system. Classifications in these categories are entered into three new columns added to the spreadsheet, as shown in Figure 10. The categories are:

- the number of ambiguous objects under the representation of the eye point of gaze observed in the video transcript (column E on Figure 10). The figures derived from this are used to obtain a probabilistic model of spatial accuracy in the accuracy model developed later in this document.

- number of records apart in time of objects that correlate (column F on Figure 10). Since the object logging time is variable, object log and video transcript records do not occur at exactly correlating times. The analyst therefore records the number of records separating log entries and the apparently corresponding video transcript entries. A value of ‘0’ means that the object is logged both before and after the video transcript record in which it occurs. ‘-1’ means that the object log entry occurs one record later than the transcript entry, ‘+1’ one record earlier than the transcript entry, etc.. This illustrated on Figure 11. The figures derived from this analysis are used to obtain a probabilistic model of timing accuracy in the accuracy model developed later in this document.

- a description of specific circumstances of correlation or reasons for error, for example that a logged object lies under a saccade (valid) or that an object was logged when the raytracing in the game engine intercepted its bounding volume when the bounding volume extended beyond the visual edge of the object (so the actual gaze object was behind this transparent area covered by the bounding volume) (column F on Figure 10).
H. Derive Accuracy Model

The first two categories under point 7 above provide the basis for statistical characterization of the temporal and spatial accuracy of the system. The accuracy model is then the derived accuracy for the Tobii eyetracking system integrated with the HiFi game engine. This is described in detail below, both in terms of frequencies and an abstracted probabilistic model (in the section ‘Analysis: Resulting Accuracy/Precision Model’).

4.2 Summary of Data Quantities in the Verification Procedure

The verification study documented here covers logged data from a player in the ‘SW’, ‘NW’ and minor parts of the ‘NE’ zones of the 3D environment stimulus described below. The generated text data from the Game Engine is logged every 25-30 ms. The real-time captured AVI (Audio Video Interleaved) video file including superimposed gaze data has a frame rate of 10 fps corresponding to an inter-frame time of 100 ms. So 3-4 rows of game engine log data correlates on average with one row of AVI information in this case. The total object log entries generated by the game engine for the session used for the verification study is 13,285 rows, while the AVI file includes 3780 video frames resulting in 3780 rows of transcribed AVI information.

4.3 Detailed Description of the Stimulus Used in the Verification Procedure

The first author of this report developed the stimulus (i.e. the game world) using the game level editing environment BattleCraft®. BattleCraft is a level editor for the game "Battlefield
1942”. Before starting level editing, the scenario, created by one of the coauthors at FOI, was described as a foundation for the layout of the game world. The description of the stimulus is as follows.

4.4 Scenario for the Verification of Automated Gaze Object Logging

The player plays a soldier alone in a city environment. The task is to navigate through a city labyrinth to find the rest of their group. The level of challenge might be increased by increasing the time pressure to find the group, e.g. using the scenario that an enemy artillery attack is expected within x minutes.

In order to be able to find the group, the player can use a mini map; the use of the minimap (i.e. how often and how long the player looks at it) is an example where eyetracking can provide useful design information.

In the game environment several “navigational clues” can be provided as to where the player’s group can be found (e.g. tracks in the terrain, road signs, wheel tracks from their tank, “red ribbons around branches”, lit torches, signals/arrows painted on walls, dead bodies, bullet holes, exploded holes in walls, blown up cars, discarded weapons, etc.). It is also possible to have allied soldiers in the town that the player can run up to and gain clues from (e.g. by dialog text such as “I saw them by the church”).

Threats/objects to avoid in the terrain include remotely deployed antipersonnel mines (“fjärrutlagda” mines, FASCAM mines that are deployed by artillery, which means that they are openly exposed on the ground and easy to see), improvised explosive devices (IEDs, booby-traps, which are hard to see), enemy vehicles and soldiers that patrol the area. Game testing will be needed to tune the severity of damage from these threats.

Figure 12. An overview of the level with 4 starting points, SW (South West), NW, NE and SE.
A number of "power-ups" and objects to collect are located in the game world. The player has a secondary task to collect ammunition for the machine gun since the group is low on ammunition.

All of the above mentioned objects and mines have to be graphically large enough to be certain of being able to detect gazes upon them from the eyetracker. With the following scenario in mind, the first author designed the maps presented in figure 12 to 16, which are images from the level editor (these high level displays are not available to the player within the game itself). The map shall be possible to copy so it is possible to play the same map several times but entering it from a different direction. This creates a labyrinth of four quadrants entered from each compass direction and having approximately the same difficulty.

The stimulus has met the requirements mentioned above except for the provision of clues. This is a major element for game play but because of time limitations this was not implemented for the verification study, where the logging of the objects is the priority, and not the game play itself.

Figure 13. The objective of the player is to reach the church. The player is told that their group can be found near the square. The number of opponents increases as the player gets closer to the square and to the church.
Figure 14. A top-down-view of the environment (North West corner), the game objects and their positions.

Figure 15. The BattleCraft "Camera Fly Mode" and the "Material Mapper" view.
4.5 Discussion of Observations Made During the Verification Process

This section presents commentary and examples of accuracies and inaccuracies observed in the object logging while creating the video transcript and cross-reference to the HiFi object log. Undertaking this analysis requires detailed investigation of the relationships within and between close rows through all 17,064 of the sorted data table, with reference also to the display of the video frames in the video file.

Note that some instances are discussed primarily in the body of the text while others are only discussed within figure captions. Examples of observed errors and issues of interpretation are given in this section, while error frequencies are presented in the section 'Analysis: Resulting Accuracy and Precision Model' below.

4.5.1 Picking and raytrace errors

Picking errors occur when the engine is not logging the raytraced object correctly within the virtual world. The raytrace is a linear pointer and the target object has to be hit by that line. Raytrace errors happen when a raytrace goes through the 'correct' object and the engine picks out another object behind it. An exemplification of a raytrace error follows below.
Figure 17. On row 3,674 one can see the transcribed object on the right side and on the next row (3,675) shows the logged object by the HiFi Engine.

Figure 18. Player’s gaze is on corner of building citymesh4_closed_m1 [976_1042]. In frame 819 the avi and the transcription state that the gaze-point-dot-representation represents a gaze point on citymesh_4_976_1042 (row 3,674), but the object log says citymesh_4_1083_1054 (row 3,675). See Figure 17,18 and 19.

Figure 19. The player is on the left side watching building citymesh4_closed_m1 [976_1042] but the Object Log (HiFi Engine) logs citymesh4_closed_m1 [1083_1054].
In row 3,679, it could be that the house on the other side of the square is correctly picked. Here is an ambiguity though, see figure 17 and 19. Between two houses the german_soldier_[965_1052] stands and behind him there is a ruin with open windows (figure 20). Row 3,679 is either a raytrace error through the german soldier or the visual interpretation is wrong. The fixation marker suggests that it is a visual error.

Figure 20. “Print Screen” from situation row 3,679. Soldier is behind blue marker.

Figure 21. Both highlighted soldiers were fixated many times, but only three log entries for the right hand soldier were generated.

Neither of the soldiers in figure 21 has been logged even if the fixation has been placed on central parts of the body representations. The soldier on the right has had only three of its fixations logged, while the left soldier appears not to be logged at all. However, the wdfence in the background is logged all the time.
Below is an example of when the player is inside citymesh2_m1_interior and the raytrace goes through the wall.

![Image](image_url)

Figure 23. The player is inside the building and fixations are distributed mainly on the indoor walls. This happens from row 11,859-11,864 /frame 2605-7.

In figure 23 the blue marker is on the wall but the raytrace seems to go through the wall to pick eu_strltlight instead. The eu_strltlight and citymesh4_1081_1254 are outside in the same direction and behind the blue marker and the wall. The house though is more likely to be picked up because it is a bigger object and could be logged while player is sweeping their view across the window. Another example of raytrace error is presented in figure 25.
Figure 25. Here the ray-trace goes through the finger graphic and picks the house object in the prolonged direction behind it [row 14,144 / frame 3,129].

Figure 26. The same problem occurs but the ray-trace hits the ground instead of the weapon or the right hand [row 14,211 / frame 3,144].

Figure 27. The engine seems to have difficulty logging the soldiers. Here is an obvious case where a soldier should be logged and it is [row 14,085 / frame 3,116] but most of the time the building behind is logged. See also Figure 28.
In figure 28 we can without doubt see that the fixation is directed at the centre of soldier 's chest, but most of the time the Object Log records the house in the back-ground. Raytracing goes through the object in these cases.

4.5.2 Saccades and logging

Saccades and fixation histories over the past 100 ms are represented on the AVI file, to ensure no data loss since 100 ms is the AVI inter-frame time. Sometimes, however, the saccades do not show.

Figure 29. Saccades can cover 2 objects and more, so what is under a saccade is tracked and logged, but maybe not shown as a fixation. See row 1,304.

Figure 30. The spreadsheet entry corresponding with the image in Figure 29 shows how the house is logged instead of the actual fixated object as shown in the AVI file. The fixation itself is logged later (with a -4 row lag).
4.5.3 Rendering problems

Figure 31. Here a problem with the logging occurs and it seems to be a rendering problem. A house of this type has a “window-door-window-door” sequence at the bottom level. The black and white part of the same house is logged here as two different objects. The left black part of the house, which has not been rendered correctly, seems to be logged as the house on its left side, Row 5,119. When the fixation has been on the white side and goes back to the black side, it is then logged correctly again.

4.5.4 Obscuration problems

The blue marker representing the gaze point increases in radius over time to represent gaze duration (i.e. fixation time) on the same object(s). Unfortunately the marker is opaque, so precise positions and jerky eye movements under the marker cannot be analyzed in detail. Since one fixation can have a duration of over 29 frames, this is a major disadvantage when analyzing detailed visual content from the AVI file.

Figure 32. A typical case when a fixation is targeting one object and the marker representing it increases over time to obscure the detail beneath. If the marker is turned off, gaze time information is lost.
Figure 33. This is an example of how visual interpretation problems occur due to the opaque fixation marker increasing over time. The transcription says [germansoldier_972_1103] in column “I” and in column “J”, while the transcription says “fixation grows”. In column “A” the Object Log [wdfence_1042_1123] has logged differently. The reason for this is the increasing size of marker, which is not transparent.

4.5.5 Collision detection errors

Graphical objects within a virtual game world have a bounding volume that is referred to by the physics system to detect collisions between that object and other objects having bounding volumes. When adding these objects to a graphical model, the volumes can have different shapes from the model shape depending on what kind of triggers or actions one wants to implement for these objects, and depending upon the required accuracy of collision detection (in relation to the shape of the graphical mesh). If there is no bounding volume at all around an object, then it does not exist as far as the physics system is concerned. Since the integrated HFE/eyetracking system uses the collision of a ray trace with a bounding volume to identify the object under a gaze point, an object having no bounding volume will not be logged by the system as an object within the world.

Figure 34. In this case the wall/fence does not have any collision detection and therefore, since the object under the gaze point is detected by the engine using a collision volume/surface, the hangar is logged instead.
In this case the soldier sinks into the building after being shot. However, this is a collision detection error that is not relevant to verification of the eyetracker integration per se.

4.5.6 Collision mesh accuracy
Figures 37 and 38 show how ambiguous it can be to estimate what the gaze is actually directed at when collision detection is carried out on transparent parts that may be looked through [row12,224 / frame2682-6]. It would be possible in principle to create collision meshes corresponding exactly with all graphical boundaries. However, other spatial inaccuracies in the system (e.g. the spatial accuracy of eyetracking) suggest a limit to the usefulness of this, although it is also dependent upon virtual distance to the object: from far away, the spatial accuracy of the eyetracker is the dominant limiting accuracy factor, while very close up the accuracy of the collision volume in relation to the graphical object is the dominant factor.

4.5.7 Spatial Ambiguity due to distance

The problem of ambiguity due to the small-scale spatial inaccuracy of eyetracking, together with obscuration by the gaze point graphic, also applies for objects that are close together and have a far virtual distance from the player character viewpoint. Small scale, natural jerky eye movements also become potentially significant in size for distant objects that are close together.

In figures 39, 40 and 41 it is hard to interpret what is being fixated when the fixation dot is centered over a cluster of distant objects, especially when the dot itself is increasing over time. The problem can be illustrated more clearly in close up images.
From left there is an angel_sculpture, in the middle a fence in profile that consists of three individual models in depth, and to the right there is a soldier. The scene is shown from a distance in the left of figure 42.

The fence has two more pieces behind it that are not viewable from this angle (figure 43).

On the right there is a soldier (the grey box) and the fence is also visible from the right side (figure 43).
In fig. 44 a transparent circle is added and placed over the three objects to illustrate the disadvantage of the opaque (blue) marker currently available within the Clearview software, which can be the same size as the circle, obscuring all of the relevant detail. The system could alternatively show just the circle boundary so the transcriber can see what is behind the circle. A third alternative would be a cross, centred upon the fixation and with a size depending upon fixation time.

4.5.8 Physics errors

Physics errors other than raytracing collision errors are not immediately relevant to the verification of the integrated eyetracker/HFE system, but may have a bearing upon interpretation of the data provided by the system in later studies. This is because any such errors that are visually observable by players may represent novel events that attract the attention of a player by virtue of their unusualness and unexpectedness. The resulting gaze behaviour will then not be related to visual attention patterns that characterise task-related decision processes.
4.5.9 Other engine errors

Specified “distance logging”, referring to the virtual distance from the player character observation point to the object that is the target of each fixation within the virtual world, is not carried out by the version of the HFE software used for the verification.

5 Analysis: Resulting Accuracy/Precision Model

5.1 Characterising Accuracy and Precision

The verification of the integrated eyetracker/HFE system reported here has been undertaken by comparison of the eye gaze data logged by the HiFi engine with captured video data of the history of gameplay with eye gaze data superimposed over the image in the form of dots representing fixation points and lines representing saccades. For analytical clarity and for transfer of the results of this study to the interpretation of data in experimental (rather than verification) studies based upon the use of the integrated eyetracker/HFE system, it is necessary to derive an abstracted and transferable model of the performance of the system from the detailed data used in and derived from the verification analysis. Such a model represents the conceptual result of the verification study.

It is useful and appropriate for this conceptual result to take the form of an empirically derived model of the accuracy and precision of the system. In general, it may be stated that:

“In the fields of science, engineering, industry and statistics, accuracy is the degree of conformity of a measured or calculated quantity to its actual (true) value. Accuracy is closely related to precision, also called reproducibility or repeatability, the degree to which further measurements or calculations will show the same or similar results.”

“The results of calculations or a measurement can be accurate but not precise; precise but not accurate; neither; or both. A result is called valid if it is both accurate and precise.”


In other words, accuracy is the degree of veracity (or closeness to the truth) of a measure, while precision is the reproducibility of the measure of accuracy.
In the current verification study, data can be extracted from the sorted and analysed object log and video transcript table to consider accuracy and precision in both the spatial and the temporal performance of the system. Considering accuracy and precision specifically relating to the integrated system (and not its two primary component subsystems, HFE and the eyetracking system) these terms can be applied in the following ways:

Spatial:

Accuracy:

Each object log entry reports one object under the gaze point. Accuracy can be assessed by comparison with a known actual number of objects within the region of accuracy of the eyetracking system. The result can be expressed in terms of: for a given record for an object within the object log, what is the probability that there were actually \( n \) objects under the gaze point, for an arbitrary +ve integer \( n \)?

Precision:

Given a characterisation of spatial accuracy as described above, spatial precision can then be assessed in terms of the repeatability of the probabilities obtained for different values of \( n \) objects under the gaze point.

Temporal:

Accuracy:

Each object log entry reports the time at which a specific object was under the gaze point during one simulation time cycle of the game engine. Accuracy can be assessed by comparison with a known actual time of that
specific object under the gaze point during one simulation time cycle of the
game engine. The result can be expressed in terms of: for a given record for an
object within the object log occurring at a time tlog, what is the probability
that the actual time at which the gaze occurred is tlog +/- toffset, for toffset
in some known real number range of seconds?

Precision:

Given a characterisation of temporal accuracy as described above, temporal
precision can then be assessed in terms of the repeatability of the
probabilities obtained for different values of toffset from the time at which
an arbitrary object is recorded as being under the gaze point.

Repeatability in both cases could be assessed by looking at probabilities derived: a) from
different subsets of data within a single verification test run, or b) from data across multiple
verification test runs.

In developing accuracy and precision models in detail below, other aspects of accuracy and
precision will be considered in addition to the above components specifically relating to
the integrated system. This results in the following categories of the components of spatial
and temporal accuracy and precision:

– the visual system of the player

– the HFE system

– the Tobii eyetracking system

These categories will be considered in turn below, first for spatial accuracy and precision
and then for temporal accuracy and precision.

5.2 Detailed Model of Spatial Accuracy and Precision

5.2.1 The Visual System of the Player

It is useful to briefly summarise the accuracy or precision of the human visual system,
especially in relation to the other systems. Foveal resolution refers to the visual angular
spatial resolution of the most sensitive, high resolution area of the retina within the eye.
The highest resolution of the human eye (i.e. highest perceivable spatial frequency cycle)
has been characterized to be around 30 arc seconds. At the optimal viewing distance of
the Tobii eyetracking system of 60 cm (and hence the distance of the game screen from
the player’s eye in this study), this represents a distance of 4.6e-7 m on the screen. The
Tobii screen measures 0.34 x 0.27 m, which at a pixel resolution of 1024 x 728 gives pixel dimensions of 3.3e-4 x 3.7e-4. Clearly, foveal resolution is insignificant compared to the pixel size on the screen.

Eye pointing accuracy can be considered in terms of muscular tremor, which has a variation of 20-40 arc seconds. Since this is around the same value as the 30 arc seconds of foveal resolution, it can similarly be disregarded as a factor compared with apparent pixel size.

A potentially significant factor in the accuracy of vision is the role of peripheral vision. This is dependent upon the task, the task environment and the learned visual competence of the observer within the environment and so should be studied within that context.

5.2.2 The HFE System

The HFE system and the game environment are sources of basic limitations in spatial accuracy. If the game engine is subjected to high processing demand during gameplay, it could slow down to a point of introducing noticeable visual delays that might interfere with the mapping of screen (x,y) coordinates onto virtual visual objects. However, this was neither observed nor specifically tested for in the study. Graphical discrimination due to factors of hue, saturation, lighting, materials and textures can make a difference to the perceivability of objects (e.g. whether a player perceives a boundary between objects or not). These factors are dependent upon the basic graphics and rendering functions of the system. Again, these factors were not specifically tested for in the study.

A factor that did have an observed impact upon the accuracy of the system was the correspondence of specified object bounding volumes within the game design with the graphical objects that they are associated with. This factor is subsumed within the human/machine error statistics described below.

Maximum screen resolution, with its influence on the position of edges, is a function both of screen capacity and rendering capacity. Since this was dependent in this study upon the Tobii eyetracker screen, it is considered in the next section.

Additional game engine errors having a bearing upon the study include rendering errors (e.g. an object is not visible when it should be), raytracing errors (concerning engine functions added for integration with the eyetracker, amounting to failure to intercept objects under the point of gaze even though the gaze coordinates have been successfully retrieved from the eyetracker) and errors in the position of objects. These factors are subsumed within the human/machine error statistics described below.
5.2.3 The Tobii Eyetracking System

As noted above, the recommended screen resolution of the Tobii eyetracker is 1024 x 768 pixels, resulting in pixel dimensions of 3.3e-4 x 3.7e-4 m. The Tobii eyetracker has a stated accuracy of 30 pixels, which can be interpreted as a diameter of about 1 cm at the recommended viewing distance of 60 cm. This is a significant factor in spatial accuracy.

Nyquist (or Shannon’s) sampling theorem states that a signal must be sampled at least at the same rate as its highest frequency component in order to avoid aliasing (i.e. reading higher frequencies as lower frequencies). 1 cm accuracy of the eyetracker (corresponding to 100 spatial wavelengths per metre) means that in principle, the highest spatial frequency that can accurately be reproduced is 50 cycles per metre, or 2 cm. Hence periodic spatial frequency variations of a higher frequency than that cannot be accurately discriminated as gaze objects.

Regarding screen resolution, sampling theorem allows spatial frequencies greater than or equal to 2 pixels in size to be reproduced without spatial aliasing, amounting to a visual size of 6.6e-4 m horizontally and 7.4e-4 m vertically. Clearly, the eyetracker spatial frequency limits are a far more significant factor than screen resolution in determining the spatial accuracy of the system. Hence screen resolution can be disregarded.

Graphical discrimination (as affected by hue, saturation, light, material and texture) is a function of the game design, rendering engine and the Tobii screen itself. This factor is not considered in the study.

Calibration of the eye point of gaze is a critical factor in determining the spatial accuracy of eyetracking. However, the quality of calibration data is assessed automatically (the calibration procedure can be repeated in the case of poor calibrations), with no quantified indication of the meaning of a good calibration. Hence this factor is not considered in the study, other than by using only what the ClearView system regards as ‘good’ calibrations.

For the verification procedure, the spatial accuracy of video capture and the representation of gaze position over the captured video for interpretation are significant factors. The recommended screen resolution of captured video is 640 x 480 pixels, much lower than that of the Tobii screen itself. The impact of this is significant, as manually assessed and reflected in the figures for machine versus visual interpretation error presented below.

Gaze data is represented in the video file as a coloured line for saccades and as an opaque dot for fixation points. The line is thin enough to be disregarded as a limiting spatial accuracy factor. The fixation point dot can be set with a radius of between 10 and 100 pixels, with a default value of 30. Since the eyetracker has a rated accuracy of 30 pixels, a setting of 15 pixels radius corresponds to the spatial accuracy of eyetracking. If the video is analysed with a dot of this size, ambiguity due to the dot obscuring object detail should be an accurate
representation of the basic spatial accuracy limits of the eyetracker. Using a larger dot size can obscure image detail and reduce the accuracy of visual interpretation of objects under
the represented point of gaze to a level less accurate than that due to the accuracy limits of
the eyetracker. For the study, a radius of 30 pixels was used, with the resulting ambiguities
being taken into account in the summary accuracy model (below). As discussed below, this is not regarded as a critical issue, since object ambiguity under the point of gaze is
a function of game level design and player tasks, making this aspect of spatial accuracy
difficult to generalise into a precision model across level designs and players.

5.3 Resulting Spatial Accuracy and Precision Model

5.3.1 Spatial Accuracy

As noted above, the spatial accuracy of the integrated HFE/Tobii system can be expressed
as the probability for a given record for a specific object within the object log that there
were actually n objects under the gaze point, for an arbitrary +ve integer n.

Examination of the captured video data showing the game screen with superimposed gaze
data revealed many cases where visual examination could not disambiguate the number
of objects under the gaze point, especially within the 30 pixel diameter accuracy limits
of the eyetracking system. The following table shows the frequency of the occurrence of
n ambiguous objects under the gaze point, together with unrelated errors in the final
column.

<table>
<thead>
<tr>
<th>n</th>
<th>0</th>
<th>1</th>
<th>2</th>
<th>3</th>
<th>4</th>
<th>5</th>
<th>6</th>
<th>7</th>
<th>error</th>
</tr>
</thead>
<tbody>
<tr>
<td>count</td>
<td>80</td>
<td>3331</td>
<td>235</td>
<td>112</td>
<td>12</td>
<td>3</td>
<td>1</td>
<td>0</td>
<td>3</td>
</tr>
</tbody>
</table>

Figure 47.

In Figure 47 these figures are depicted as probabilities of n Objects under the gaze point.
The meaning of this probability distribution is that if the integrated HFE/Tobii system
is used in a study, for each entry in the object log these is actually a probability of not 1
but n objects being under the gaze point, where the probability of there being n objects,
P(n), is as shown on Figure 48.
5.3.2 Spatial Precision

The spatial precision of the integrated HFE/Tobii system is a matter of how the accuracy derived above might generalise to other studies. No detailed spatial precision model has been derived. In general it is doubtful whether such a model can be derived, since a spatial accuracy model derived using the method above may vary a great deal depending upon stimulus design and the detailed task(s) given to a player under study. For example, a stimulus game level could be constructed with a small number of very simple objects (e.g. buildings) within the game level design. In this case, there may be few instances where the player experiences overlapping objects within a small visual field and hence a very low frequency of highly ambiguous perceptual situations. On the other hand, a stimulus could use a very large number of densely located objects (e.g. using many small plants and artefacts very close together). In this case there may be very many situations in which many small objects fall within the 30 pixel accuracy limits of the eyetracker. In both cases, the nature of the tasks given to a player may draw attention towards or away from spatially ambiguous areas of the game. For instance, a task to search for a small object may lead the player to look at areas where many plants are close together, providing good hiding places. Alternatively a large scale navigation task may lead the player to disregard clusters of small plants while looking more at large scale navigational features, such as buildings.

5.4 Detailed Model of Temporal Accuracy and Precision

5.4.1 The Visual System of the Player

The frequency of high frequency tremor of the eye is generally within a range of 30 to 90 Hz. However, since the spatial consequences of this are insignificant, this factor is disregarded. The general sampling rate of the eye for sharp edges in the visual field, due to high frequency tremor, is 90 Hz. Sampling rates due to saccadic motion can vary within a wide range,
depending upon the visual stimulus, the task, the person and the occasion. This is not considered directly here in terms of general saccadic movement, but it may be noted that the Tobii sampling rate of 50 Hz limits accurately sampled saccadic movement frequency components to those having frequencies of less than or equal to 25 Hz (see below).

Spontaneous EyeBlinking Rate (SEBR) generally occurs at 2 to 10 sec intervals, having a duration of 0.2 to 0.4 sec per blink. This is significant and results in corresponding loss of tracking by the Tobii system. However, this was not found to have any significant analytical consequences and so is disregarded in this study.

5.4.2 The HFE System

The HFE frame rate varies depending upon transient processing loads, with an average frame interval of about 30 ms; this is the same as the simulation update rate (more precisely, the game engine update rate appears to vary between 28.6Hz and 50Hz, with an average of 30 Hz). According to sampling theory, this update results in the unaliased synthesis of time frequency components less than or equal to approximately 15 Hz. The effects of this will be dependent upon the time frequencies built into a game level design, e.g. frequency components of animation sequences.

The more immediate and intrinsic effects of the variable HFE frame rate for the integrated HFE/Tobii system occur in terms of time offsets of object log entries compared to captured video frames. This variability subsumes lag, effectively encompassing the effects of both lag and frame rate variations in comparison with gaze data log data obtained from the Tobii system.

5.4.3 The Tobii Eyetracking System

The Tobii eyetracking system has a sample rate of 50 Hz, and hence a sample interval of 20ms. By sampling theory, this means that only eye movements having a temporal frequency of less than or equal to 25 Hz can be reproduced without aliasing. This is a fundamental limitation of the eyetracking system, so no higher frequency information about eye movement is available for this study. Since this frequency is nevertheless higher than that of the HiFi game engine and the frame grabber, the Tobii gaze movement data must be used as the basis from which accuracy and precision characterisations are obtained. Temporal accuracy therefore refers to the time accuracy of correlation between the eyetracker and the game engine object logging.

Video screen capture is conducted at a rate of 10 Hz, or 10 fps (hence with an inter-frame time of 100ms). Using a gaze history time of 100 ms results in all captured gaze data from one frame to the next being represented in the video output generated by the Clearview software. By sampling theory, this frame capture rate means accurate capture of frequency components having a frequency of less than or equal to 5 Hz. The relatively small scale
of changes between frames as a function of the game design and its dynamics makes this acceptable as a basis for interpreting the objects under the gaze point within each frame, especially given the 30 pixel resolution of the eyetracker (since faster changes tend to occur within smaller spatial regions). Variations due to the engine frame rate are analysed in relation to the relative difference in time between when a frame is grabbed and when the corresponding object log record is made by the HiFi engine.

5.5 Resulting Temporal Accuracy and Precision Model

5.5.1 Temporal Accuracy

As noted above, temporal accuracy of a given record for an object within the object log occurring at a time log can be expressed in terms of the probability that the actual time at which the gaze occurred is tlog +/- toffset, for toffset in some known real number range of seconds.

If the game engine had a fixed frame rate that is a multiple of 20 ms and the engine and the eyetracker were precisely synchronised, every engine frame would correspond with a subsample of every n-th gaze sample taken by the game engine.

However, the game engine frame rate and lag in the interface with the eyetracker are both variable. For the verification study, the only possible reference for ‘the actual time at which the gaze occurred’ is the video file with overlayed gaze data. The timing of video frame capture is consistent, but at a much slower frame rate of 10 Hz, compared to the eyetracker sample rate of 50 Hz. With the gaze history set to 100 ms, a video frame and its gaze data represent the best available representation of where the gaze point was at the time of frame capture and for 100 ms prior to and leading up to that time.

During 100 ms the video image changes; this change is small and was not quantified during the study, but is assumed to be small enough for the frame image to represent a reasonable basis for interpreting the objects under the gaze point through the whole 100 ms gaze history represented on the frame image (no more accurate data than this is available).

Variability in the frame rate and lag in the connection with the eyetracker result in object log entries having highly variable sample timing, having an average frame interval, and hence sample interval, of 28.6 ms (with a population standard deviation of 6 ms). The object log entry is stamped with a time stamp from the game computer. Hence the log entry time value includes the variable lag in the request and reception of a gaze data point from the TET eyetracking server.

Given these uncertainties, in order to obtain an approximate time accuracy characterisation, it is assumed to be a reasonable approximation to equate time variability with relative entry
position in the sorted log/transcript file. That is, from the time of a given captured video frame, the temporally closest corresponding object log file entry is found and the number of entries ahead or behind the video transcript in the sorted log file is taken to reflect its time accuracy, based upon the 28.6 ms average frame rate.

![Object Count of Offsets - 1 Gaze Object](image)

Figure 49. Frequency count of offsets of object log entries from object visibility in the video transcript.

The count of offset values is then given in the following tables. In the first the case, one unambiguous object under the gaze point is considered (i.e. the object log record corresponds with the only object under the gaze point in the video transcript). The resulting tables are as follows:

**Negative Offsets (-):**

```
<table>
<thead>
<tr>
<th>Offset</th>
<th>Count</th>
</tr>
</thead>
<tbody>
<tr>
<td>-15</td>
<td>1</td>
</tr>
<tr>
<td>-14</td>
<td>1</td>
</tr>
<tr>
<td>-13</td>
<td>0</td>
</tr>
<tr>
<td>-12</td>
<td>0</td>
</tr>
<tr>
<td>-11</td>
<td>0</td>
</tr>
<tr>
<td>-10</td>
<td>2</td>
</tr>
<tr>
<td>-9</td>
<td>4</td>
</tr>
<tr>
<td>-8</td>
<td>5</td>
</tr>
<tr>
<td>-7</td>
<td>3</td>
</tr>
<tr>
<td>-6</td>
<td>16</td>
</tr>
<tr>
<td>-5</td>
<td>32</td>
</tr>
<tr>
<td>-4</td>
<td>42</td>
</tr>
<tr>
<td>-3</td>
<td>80</td>
</tr>
<tr>
<td>-2</td>
<td>144</td>
</tr>
<tr>
<td>-1</td>
<td>216</td>
</tr>
</tbody>
</table>
```

**0 Offset:**

```
<table>
<thead>
<tr>
<th>Offset</th>
<th>Count</th>
</tr>
</thead>
<tbody>
<tr>
<td>0</td>
<td>2232</td>
</tr>
</tbody>
</table>
```

**Positive Offsets (+):**

```
<table>
<thead>
<tr>
<th>Offset</th>
<th>Count</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>95</td>
</tr>
<tr>
<td>2</td>
<td>46</td>
</tr>
<tr>
<td>3</td>
<td>19</td>
</tr>
<tr>
<td>4</td>
<td>6</td>
</tr>
<tr>
<td>5</td>
<td>4</td>
</tr>
<tr>
<td>6</td>
<td>1</td>
</tr>
<tr>
<td>7</td>
<td>1</td>
</tr>
<tr>
<td>8</td>
<td>0</td>
</tr>
<tr>
<td>9</td>
<td>0</td>
</tr>
<tr>
<td>10</td>
<td>0</td>
</tr>
<tr>
<td>11</td>
<td>1</td>
</tr>
<tr>
<td>12</td>
<td>0</td>
</tr>
<tr>
<td>13</td>
<td>0</td>
</tr>
<tr>
<td>14</td>
<td>0</td>
</tr>
<tr>
<td>15</td>
<td>1</td>
</tr>
</tbody>
</table>
```
These values are shown on the chart on Figure 49. Note that the ‘total match’ condition occurs when the video transcript entry time exactly matches the time of a corresponding object log entry (actually there are 0 cases of this), or when a the video transcript entry is flanked in the sorted time list by two corresponding object log entries. The chart on Figure 50 shows the same frequency distribution using a logarithmic scale.

![Chart](image)

Figure 50. Frequency count of offsets of object log entries from object visibility in the video transcript, using a logarithmic scale.

For the case of any number of objects under the gaze point (i.e. the object log record corresponds with at least one object out of any number that are under the gaze point in the video transcript), the resulting tables are as follows:

**Negative Offsets (-):**

<table>
<thead>
<tr>
<th>Offset</th>
<th>Count</th>
</tr>
</thead>
<tbody>
<tr>
<td>-15</td>
<td>2</td>
</tr>
<tr>
<td>-14</td>
<td>2</td>
</tr>
<tr>
<td>-13</td>
<td>1</td>
</tr>
<tr>
<td>-12</td>
<td>0</td>
</tr>
<tr>
<td>-11</td>
<td>0</td>
</tr>
<tr>
<td>-10</td>
<td>3</td>
</tr>
<tr>
<td>-9</td>
<td>5</td>
</tr>
<tr>
<td>-8</td>
<td>6</td>
</tr>
<tr>
<td>-7</td>
<td>4</td>
</tr>
<tr>
<td>-6</td>
<td>18</td>
</tr>
<tr>
<td>-5</td>
<td>36</td>
</tr>
<tr>
<td>-4</td>
<td>45</td>
</tr>
<tr>
<td>-3</td>
<td>103</td>
</tr>
<tr>
<td>-2</td>
<td>183</td>
</tr>
<tr>
<td>-1</td>
<td>247</td>
</tr>
</tbody>
</table>

**0 Offset:**

<table>
<thead>
<tr>
<th>Offset</th>
<th>Count</th>
</tr>
</thead>
<tbody>
<tr>
<td>0</td>
<td>1</td>
</tr>
</tbody>
</table>

**Positive Offsets (+):**

<table>
<thead>
<tr>
<th>Offset</th>
<th>Count</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>1</td>
</tr>
<tr>
<td>2</td>
<td>101</td>
</tr>
<tr>
<td>3</td>
<td>60</td>
</tr>
<tr>
<td>4</td>
<td>26</td>
</tr>
<tr>
<td>5</td>
<td>65</td>
</tr>
<tr>
<td>6</td>
<td>62</td>
</tr>
<tr>
<td>7</td>
<td>0</td>
</tr>
<tr>
<td>8</td>
<td>0</td>
</tr>
<tr>
<td>9</td>
<td>1</td>
</tr>
<tr>
<td>10</td>
<td>1</td>
</tr>
<tr>
<td>11</td>
<td>0</td>
</tr>
<tr>
<td>12</td>
<td>0</td>
</tr>
<tr>
<td>13</td>
<td>1</td>
</tr>
<tr>
<td>14</td>
<td>15</td>
</tr>
<tr>
<td>15</td>
<td>1</td>
</tr>
</tbody>
</table>
Figure 51 shows probabilities of offsets based upon these frequency counts, both for single objects and for multiple objects under the gaze point. Hence for a given object log entry, a reasonable heuristic is that the actual time of the entry has a probability of being the stated time +/- 28.6 P(offset), where offset has a range from -15 to 15 and P(offset) is as depicted on Figure 51.

Since a video frame covers 100 ms and the object log entries have an average inter-frame rate of 28.6 ms, it can be considered that a log entry falls within an acceptable range if it is within 100/28.6 = 3.50 entries of the video transcript within the sorted log file. There are 3780 transcribed video frames. There are 3072 objects log entries that fall within +/- 3 frames of the corresponding video frame time. Hence the overall temporal accuracy of object logging might be said to be 3072/3780 = ~81%. Note that this follows from the sampling rate represented by the captured video file of 100 msec per sample (frame), acting as an accurate sample rate only for sampled data periods greater than or equal to 200 msec. +/- 3 frames amounts to a 6 frame interval, equivalent to 200 msec but rounded down to the integer 6 multiple of average log entry durations, thereby representing a time span of 6 x 28.6 = 171.6 msec. In effect, interpretation of the video frame capture rate in terms of corresponding object log entries within +/- 3 log entries, functions as a temporal filter on the object log data, filtering out small scale variations in line with the temporal resolution of video frame grabbing.
5.5.2 Temporal Precision

Temporal precision has not been derived by considering a series of measurements over different trials. However, unlike the case of spatial precision, since the accuracy as analysed above is derived from system characteristics, it is a reasonable assumption that this represents a good characterisation of temporal precision across many stimuli. Further testing as described in this report will be necessary to verify this assumption.

5.6 Additional Sources of Error Qualify the Accuracy Models

The probabilistic accuracy models proposed above use the overall count of object log entries. This count includes a number of errors that are attributable neither to time variations nor spatial ambiguities. Instead, these are errors either within the HFE system or made by the human analyst in interpreting the video data. HFE system errors include rendering errors, collision detection errors and raytracing errors, as discussed earlier in this document. Human errors include wrong interpretations of fixated objects and saccades due to fatigue, pixelation, distance, or a dot representing a gaze point that is growing over time and obscuring what is behind it. The overall error count is broken down into percentages by type in Figure 52. Note that raytracing errors could most likely be fixed with further software development, and collision detection errors could be eliminated by careful construction of collision volumes to match as closely as possible their corresponding graphical meshes.

![Errors [999] by Machine and Human](image)

Fig. 52. The distribution of errors within the total of category 999. The total amount of errors includes 376 occasions of one gazed object and also if more than one object has been included due fixation size (ratio of 1 cm). In total, on 528 occasions
6 Conclusion

The verification study reported in this document has shown that the integrated HFE/Tobii eyetracking system performs well, with the accuracies described above. Such accuracy figures are well within what is regarded as constituting a useful system, especially for characterising gaze behaviour and variations in gaze behaviour for statistically significant numbers of test subjects. The availability of this system makes it feasible to consider experiments using large numbers of subjects and long test sessions, generating data that would involve impractical amounts of time to analyse if done purely manually. The system is now regarded as being a highly valuable tool suitable for ongoing use in such empirical studies.

7 Acknowledgements

The primary author would like to thank Alf Svensson for discussions and advice regarding helicopter flight operations, VFR and IFR, as well as many engaging conversations and much support over the years.
References


JAA Administrative & Guidance Material, Section Four: Operations, Part Three: Temporary Guidance Leaflet. LEAFLET NO. 34: NIGHT VISION IMAGING SYSTEM (NVIS) OPERATIONS. Attachment 2, Ground Training – Areas of Instruction. Section 4/Part 3 (JAR-OPS) 01.06.03


Evaluation of Real-time Eye Gaze Logging by a 3D Game Engine

Charlotte Sennersten, Sweden, charlotte.sennersten@bth.se Craig Lindley, craig.lindley@bth.se
Blekinge Institute of Technology, Karlshamn

Abstract
Human Computer Interaction studies of visual attention in dynamic 3D computer gameplay can be greatly facilitated by automated gaze object logging implemented by integration of eye gaze tracking systems with game engines. This verification study reports the spatial and temporal accuracy of such an integrated system.

Keywords: Eye-tracking, gameplay, methodology.

1 Introduction
Empirical research on computer gameplay can provide scientific and empirically validated knowledge about the affects of game design features on the player and their experience, providing deeper foundations for game designers in choosing from design options to achieve specific motivational, emotional and cognitive affects.

For Serious Games there are specific issues of learning outcomes and how to create effective situated learning environments, while for entertainment games it is important to achieve a deeper understanding of game engagement and immersion, and how design decisions can shape the player experience. Studying games from a cognitive science perspective holds the promise of providing some answers to these kinds of questions. At the same time, computer games may reveal more about cognitive processes than, for example, traditional simplified cognitive stimuli [1].

3D computer games are visually rich, interactive virtual environments in which visual cognition has a primary role. Understanding how visual cognition functions in gameplay involves basic questions, such as how a player distributes their visual attention during gameplay, how visual attention is involved in cognitive task performance, how visual behavior and task performance interact with game design features, and how these factors lead to variations of player experience and cognitive learning outcomes. Studying visual cognition in 3D computer gameplay can also help in the investigation of broader questions such as: what is digital- and visual literacy?

Eye gaze tracking is a method for determining where on a computer screen a computer game player is looking. Eyetracking is a method that has been used commonly over the last 40 years to give insights into how attention is distributed when carrying out different vision-based tasks. The technique involves tracking eye movements and recording the distribution of gaze over time. Among a variety of eyetracking techniques (reviewed by [2, 3, 4]), video-based eyetracking, using combined pupil and corneal reflection, is currently the most commonly used method due to considerations of ecological validity and ease of use [5]. In this method, following a calibration process, the eyetracking system uses a camera to track variations in reflections of diode array patterns in the eye of a subject, and these variations are algorithmically decoded to calculate where upon a computer screen a person is looking at the moment when a gaze data sample is taken.

Stimuli used for gaze behaviour analysis to date have been predominantly 2-dimensional, limited to screens of 2D data with an overall static structure containing, for example, text, still images and 2D animations in fixed positions. A typical example is [6], investigating visual distribution while viewing newspapers. Recently eye gaze tracking has been used extensively to study visual attention patterns of users of web pages, e.g. to determine layout schemes and the effectiveness of advertising placement. Analysis of gaze behavior over 2D stimuli can be based upon hotspot analysis, where gaze density is shown by color maps laid over an image, by direct plotting of gaze traces, or by gathering statistics for specified Areas of Interest (AOIs) in a static 2D representation. AOIs are specified as polygons on the 2D surface mapping out zones within the perceptual field of the flat screen. Gaze behavior can be analyzed using statistics describing how gaze (x,y) positions fall within AOIs, providing 2D spatial distribution statistics summarizing extended visual times in specific screen areas.

There have been a number of studies where these analysis techniques have been applied to the dynamic, 3D stimuli provided by interactive computer games, e.g. [7] analyse 3D gameplay using AOI’s. [8], [9] and [10] investigate correlations of AOI statistics with player immersion. [11] use eyetracking AOIs to investigate prediction of visual attention in complex 3D dynamic environments. [12] have developed a method for creating dynamic AOIs, based upon manual markup of AOIs on keyframes of the display of a captured gameplay session, with automated calculation of whether gaze data falls within those marked up AOIs. [13] describe a method for synchronising gaze data with gameplay screen content by 2-d regions (not by automated 3d-object intersection from a gaze point, as described in this paper), functionality that is now available in commercial eyetracking systems. Eyetracking has been used extensively as an interaction technology, based upon gaze points falling within the bounds of 2d screen objects; [14] provides an early study of this technique. More recently, gaze tracking has been investigated as an interaction technology for immersive stereoscopic virtual reality systems [15][3], and also as an input method for commercial 3d computer games, substituting gaze position for mouse input in controlling a game cursor [e.g. 16, 17]. However, this technique applied to game
interaction is not used for unmodified gameplay analysis per se, and has not gone further than providing an alternative interaction technology for existing game mechanics.

Eyetracking studies of dynamic visual stimuli have been generally regarded in the eyetracking community as being a very difficult problem since established analysis methods based upon statistics are only meaningful because they are collected from 2D areas within static screens for which the corresponding visual objects are also static and known. Since most of the features of a 3D game stimulus are not fixed within a 2D screen space, analysis of gaze behaviour during gameplay has required manual examination of a recorded video capture of a game session with gaze data superimposed on the game display and manual transcription of objects under gaze points for each frame. To analyse a First Person Shooter (FPS) game session manually in this way means to go through a number of video capture frames dependent upon the screen capture rate, and then manually transcribing the objects under the gaze points (e.g. fixations) [18]. For example, at a 10 fps (frames per second) capture rate, 600 frames for every minute of gameplay must be examined and transcribed. This is very time consuming, making statistically significant studies very difficult to conduct. For example, [18] reports that four hours of subject recordings took 3 months to analyse manually.

A solution to this analysis problem, described in this paper is the development of an integrated system, where data from an eyetracker is used by the game engine to automatically log objects of interest under the gaze point in real time. This paper describes this solution and the detailed methodology and summary of results of the verification study that was conducted, not only to verify correct operation of the system, but to obtain a characterization of its accuracy for use in future studies. The paper begins with a more detailed description of the problem. Specifications for an integrated system are then described, including details both of the game engine and the eyetracking system that are to be integrated. Details of the implementation of the system are then described. The paper then describes the detailed verification methodology and the resulting spatial and temporal accuracy model of the integrated system. The accuracy model is critical for being able to characterise the accuracy and hence validity of statistics obtained by using the system in ongoing studies; without such information, data obtained cannot be regarded as reliable - the accuracy model quantifies reliability. Implications of the verification study for the use of gaze-based virtual world interaction are also discussed.

2 Problem to solve
In the initial study of gaze behaviour in 3D computer gameplay conducted by [18], eyetracking was already well established as a technology and a method primarily for studying 2D stimuli, such as text, static images and web pages. Methods used in these former studies come into question when studying gameplay in 3D environments: a simple 2D solution cannot easily be applied to dynamic stimuli like 3D game worlds. AOIs are
useful when considering visual attention directed towards visually static\textsuperscript{31} 2D areas of a computer game display, such as Head Up Displays (HUD’s) and static Graphical User Interface (GUI) elements, but this approach cannot capture gaze behaviour in a meaningful way for perceptual objects that change in their 2D display position. This means that when it comes to 3D games, we haven’t been able to log gaze data automatically using AOIs because the 360 degree dynamic view around the player’s first-person point of view axis within the game world makes everything seen within the world dynamic. In 3D games there are also typically hundreds of objects the player has to move through, with the objects within the perceptual field constantly changing both in position and type, which means that automated methods of gathering gaze object statistics based upon static 2D AOIs are unusable. Analysis is possible, but only by extremely time-consuming frame-by-frame data collection. As reported by [18], a game stimulus (e.g. a commercial game level) can be played through by player participants while visual attention is recorded by the eyetracker and represented superimposed on top of frame capture data of the stimulus. The captured video file with the stimulus as background and the superimposed eye gaze targets on top of the stimulus then has to be analysed frame by frame in order to identify and record what the players attention is focussed upon while performing tasks to meet in-game challenges. This is a very time consuming procedure that becomes impractical for a statistically significant number of players playing for meaningful play session lengths of up to half an hour each: manual analysis of 0.5hrs of gameplay, corresponding to one player/play session, requires approximately 45hrs of transcription analysis in a controlled experiment.

To eliminate this impractical manual analysis time, the 2D $(x,y)$ coordinates of each gaze point can be automatically mapped into the in-game space of the 3D environment and the object(s) of gaze picked and automatically logged (see Fig.1).

![Fig.1. A ray trace from a gaze point to an object within a 3D virtual game space.](image)

The solution developed for this purpose involves connecting two independent systems, an eye gaze tracker and a suitable game engine.

\textsuperscript{31} Where ‘static’ refers to 2D position, since these elements may change within position, e.g. graphical changes like boundaries and colours indicating activated buttons, animations in place, etc.
3. Software Requirements for Integrating an Eyetracker and a Computer game Engine

3.1. Eye-Gaze Tracker Description

The eyetracker used for the work described in this paper is a Tobii Eye Tracker 1750. The 1750 runs on a PC running Windows. The system can be used either in single or double computer setups. In this study a game application is running at the same time as eye gaze tracking, which requires too much processing to do on one computer, hence we use the double computer and double screen setup.

The eye gaze tracking system detects and collects eye gaze data at a rate of 50 samples per second (50 Hz), with frame grabbing of the corresponding game display occurring at a lower but selectable frame rate. The eye gaze tracking system monitors the eyes of the subject/participant based upon reflected infrared light patterns and calculates gaze positions on the screen automatically, by hardware in combination with advanced software algorithms. The system has head motion compensation and low drift effects, with binocular tracking, which yields higher accuracy than a monocular method. Another advantage of this system is that if the participant happens to move, especially the head, tracking is not lost if one eye is occluded during the movement. (With other systems, the participants may have helmets, headrests or markers, with different respective freedoms and constraints of movement.) For ecological validity, the eye gaze tracking camera is implemented at the bottom part of a normal sized screen. The Tobii 1750 system can only track what is presented on the screen and not anything outside the screen area. The calibration time is 3 minutes if no problems occur due to eye disorders, blinking, etc.. The recommended eye distance to screen is approximately 60 cm. The accuracy of an eyetracker is measured in degrees. The accuracy or bias error of the Tobii 1750 system has been tested over a set of individuals and found to be about 0.5 degrees of visual angle when using standard accuracy measurement principles for eye trackers. One degree of accuracy corresponds to an average error range of about 1 cm between the measured and intended gaze point on the target screen at 60 cm distance from the user to the screen. The system tolerates head movements slower than about 10 cm sec\(^{-1}\).

For each eye (left and right) the following data is available derived from the eye tracker data processed by the Clearview™ software and TET Server:

- Time, a timestamp in microseconds,
- Screen X, horizontal position,
- Screen Y, vertical position,
- Cam X, horizontal location of the pupil in the camera image,
– Cam Y, vertical location of the pupil in the camera image,
– Distance, from camera to eye,
– Pupil, size in mm
– Code, (validity of gaze data).

Validity codes ranging from 0-4 are logged for each eye, with every gaze data point. 0 is the highest validity value and means that all relevant data for the particular eye should be correct.

![Fig.2. The Tobii 1750 eyetracker display with 2 infrared diode trackers and central camera at the bottom of the screen.](image)

### 3.2. Computer Game Engine Description

A game engine is a software system that generates an interactive visualisation of a 3D virtual world. This is one of several ‘layers’ that constitute a complete game. Other layers include game logic (game rules, artificial intelligence, or AI, and interaction rules), which implement the logic of action within the game world in order to facilitate gameplay. Game rules decide what the consequences of interaction(s) will be. Interaction with the system is via devices such as a mouse, keyboard and/or console, but could include other kinds of devices such as motion, sound and haptics. The third layer of a game is the game media layer, which includes such things as 3D models, 2D graphics, animation sequences, and audio components.

It is important to understand the difference between virtual worlds and games as such because virtual worlds (e.g. *Second Life*, secondlife.com, *ActiveWorlds*, www.activeworlds.com, *There*, www.there.com/, and *Metaverse*, metaverse.sourceforge.net) do not include instrinsic game rules by default. A computer game has certain tasks and challenges built/programmed/designed into the game and those particular tasks and challenges form the basis of game genres. Studies by Yarbus in the 1960s (summarized in [4]) demonstrate
significant differences in gaze paths when viewers look at a scene with different objectives. Similar differences are apparent when investigating computer gameplay where the different objectives are associated with different game tasks and challenges.

The game platform used in this verification experiment is the HiFi Engine (HFE), a game/simulation engine developed at Man-System-Interaction, Swedish Defence Research Agency (FOI) in Linköping, Sweden. The high level operation of the game engine is a simulation loop that follows the steps shown on Figure 3. The ‘Start/Initialise’ phase is a gloss over what could include many metagame functions, such as forming and choosing characters, loading levels, setting control parameters, etc. Software initialisation functions also occur here, such as instantiating classes for the GameWorld, GameObjects, ModelObjects, etc., depending upon the implementation.

![Diagram of game engine algorithm](image)

**Fig.3. Typical Game Engine Algorithm per frame and/or simulation tick.**

The simulation loop executes the game world simulation in which the player interacts with the world through keyboard, mouse or other devices. A simulation cycle might, for example, begin with getting user inputs. Then the game state is updated, based upon the game state in the last simulation cycle, together with the player inputs. Updates include updating simulations of physical events, actions of player characters, non-player characters and enemies (i.e. ‘mobs’). Once the game state is updated, the graphics can be updated, rendered and displayed.

### 1.3. Requirements for an Integrated System

![Diagram of eye tracker and game engine integration](image)

**Fig.4. Eyetracker and game engine integration requirements.**
The basic requirements for an integrated system are based upon the game engine being able to interrogate the eyetracking system to obtain the (x,y) screen coordinates of a gaze point, and then using that (x,y) coordinate to determine the underlying gaze object and generate a log entry for that object of gaze, as shown in Figure 4. On the game engine side, it is critical to be able to take the screen (x,y) coordinates from the eyetracker and trace back from these into the virtual 3D game world to identify the first object that is encountered, corresponding to the object of gaze, a process referred to as picking. This is very similar to the processing of a click at a mouse-controlled cursor position on the screen, e.g. for targeting weapons or other functions. However, the source in the current case is the eyetracker rather than a mouse, and processing of the interception point must address the generation of a gaze object log entry. The game engine must support the implementation of these functions in addition to any normal mouse (x,y) coordinate processing that occurs. (Note that for gaze-based interaction, the gaze point might replace the mouse function for implementing in-game actions, such as aiming a weapon.)

A further requirement for real time logging of objects within the game world is that each object must have understandable identity and position data for use in later analysis. Finally, it is conceptually and practically simple to implement object picking using the existing collision detection functions of an engine; then objects must have individual collision boxes so the inverted ray trace under an x, y screen coordinate can be logged in response to a specific type of collision event, in this case a collision with the ray traced from the gaze point. The veracity of object logging is then dependent upon the collision volumes around in-game objects; too large or too small a collision box for its associated graphical object can result in a misidentification of what is actually gazed upon within the game world.

4. Technical Implementation

The implementation of the integrated eyetracker/game engine system requires three additional steps in the basic game algorithm described above (see Figure 5). After the rendering step in the initial algorithm, current gaze point (x,y) information is retrieved from the eyetracker. Then the object collision box under the x,y coordinates is identified, and finally the engine logs the gaze position and object data.

Fig. 5. Modifications to the Game Engine Algorithm per frame and/or simulation tick.
In the implementation of the integrated system, an operator computer runs the eyetracking software ClearView™ (2.6.3) and the TET eyetracking server, while a gaming computer runs the game application. The two computers are connected via TCP/IP. The TET eyetracking server gets the data from the eyetracker and sends it to the ClearView software. For calibration of the eyes of the player, the operator switches the eyetracker display over to the second screen of the operating computer that runs the Clearview software, for display of a calibration pattern (hence a dual head graphics card is needed on the operating computer for the double display setup). The output from the Game Computer goes through a splitter where two video outputs are created, one going to the video capture card on the operator computer, where the captured video data is input to ClearView, while the other output goes to a switch that selects either the game display or the operator computer second monitor display (with the calibration display) to send to the eyetracker display. The routing of the video display is illustrated in Figure 6.

![Diagram](image)

Fig. 6. Data interconnection diagram showing for display routing in the dual-computer setup.

A typical play session using this system involves the game player playing the game in front of the eyetracker screen. After briefing the player with any instructions relevant to the study, the calibration process is run, calibration patterns being displayed on the eyetracker screen while Clearview™ uses the known positions of elements in the display to calibrate eyetracking for that specific player. Then the eyetracker screen is switched over to the game display via the routing switch. The operator starts game object logging within the game engine using a command line instruction, at which a synchronisation message is sent by the game engine via the TET server to Clearview™ to commence eyetracking. Synchronisation is necessary to simplify time correlation of object log entries on the game computer with video frames captured by the operating computer for the verification analysis. At an instruction from the operator, the player then commences playing while eyetracking and game engine object logging are running.

During gameplay, with each simulation tick or game display frame the HiFi (game) engine requests a gaze data sample from the Clearview™ software via the TET server. The game engine then traces a ray from the \( (x,y) \) coordinates of the gaze point to intercept the first collision mesh within the game world under that \( (x,y) \) point, corresponding with the
object being looked upon by the player. Within the game engine, a log entry is
generated that includes the name of the 3D object model appended with the (x,y)
coordinates of the model instance on the horizontal plane of the 3D game world,
ensuring that each object instance is uniquely identified by type and location,
together with a time stamp generated by a trigger signal from the eyetracker
collected with a high-resolution timer, based on Windows multimedia
timer. The log record is then appended by the game engine to the
object log file stored on the gaming computer. When the goals of the session
have been met, the operator asks the player to stop playing. A command line
instruction on the gaming computer is then used to turn off the object log and
eyetracking.

5. Verification Study of the Integrated Eyetracker and Computer
Game Engine

The verification of the integrated system aims to establish correct
functionality of the software and to characterise the accuracy of logged data
based upon an investigation of the forms and degrees of error in the system. Such
a study is critical for use of the system in experimental studies. Steps in the
Verification Study include:

i) Prepare Stimulus
ii) Select Player Participant(s)
iii) Log Play Data (generated by HiFi Engine) and Game Capture Screen Data
iv) Create Transcription (the old manual method)
v) Synchronise record times from the Game Engine Object Log and Video
   Transcription
vi) Run Time Ordering Macro
vii) Conduct detailed analysis of time ordered data — compare auto vs manually
     interpreted objects of gaze
viii) Derive Accuracy Model

The verification study uses a game level stimulus consisting of a 3D virtual model of
a World War II European town, with scattered enemy soldiers serving as
combative game challenges. The aim for the player is to reach a central church in the
town. Verification of automated gaze logging involves logging all gazed objects in
time during a play session, with parallel video capture of game play with
superimposed eye gaze positions. To determine if the engine logged the gaze
objects correctly, the object log data is manually compared with the video capture
data. The video capture data with the superimposed eye
gaze data has to be transcribed manually into text records, with a transcription entry
stating for each log entry: the object identity number, the kind of object (suburban
house, gun, fence, soldier, etc.) and also the duration of the eye gaze on the object.
When the recording and the transcription are done, the actual verification procedure starts. Verification of the real time logging by the game engine involves manual comparison of corresponding object identity numbers, object type and log record time from the engine log with the transcribed information obtained from the video capture data. Frame rates and timing are crucial to consider in analysing and verifying the data generated by this system. The game engine has its own tick-by-tick simulation update rate, the eye tracker has a different sampling rate and the real time video capture has yet another frame rate, as shown in Figure 7. The gaze point sampling rate of the 1750 eyetracker is a constant rate of 50Hz. The display frame update rate of the game engine varies, depending upon the complexity of the visual scene, within the range 25-35 ms. Video data capture was set at 10 fps (frames per second) for the present study. These different rates must be correlated during analysis in order to obtain the most accurate correspondences between gaze data, game frames and video frames, although, due to variable latencies in the system, the correlations are never perfect. Latency is defined as the time taken from the generation of a signal to its reception, in this case including time taken for eyetracking camera exposure, transfer to the TET server, calculation and delays in the server, transfer to Clearview®, request generation from the game engine, processing time in the TET server, and response propagation back to the game engine.

<table>
<thead>
<tr>
<th>Game Engine</th>
<th>Eyetracker</th>
<th>Video Capture</th>
</tr>
</thead>
<tbody>
<tr>
<td>Every ~30ms</td>
<td>Every 20ms</td>
<td>Every 100ms</td>
</tr>
<tr>
<td>2 samples</td>
<td>3 samples</td>
<td>1 sample</td>
</tr>
</tbody>
</table>

Fig. 7. Sampling time of logging, gaze- and video capture.

The start time of the Game Engine also has to be synchronised with real time video capture in order to correlate log entries and video frames. Hence the starting time of one of the files must be offset and all subsequent times within the file adjusted accordingly to gain the best overall synchronisation between log record times and frame times for apparently corresponding gazed objects. The object logs and time-adjusted transcription records are then placed in a single spreadsheet and a sorting macro is run in order to sort all records of both kinds into a single time order.

The analysis then consists of a manual row-by-row comparison of object log records with what should be the corresponding video transcript data, noting: when objects correspond, time offsets if there appears to be a lead or lag of a corresponding object, and also object ambiguity, where a single object in the log corresponds with more than one candidate object in the video transcription. In general it is possible to take a subset of samples within these files to reduce the amount of data that needs to be analysed for a verification study, but in this initial study all records were used. This amounted to the manual investigation
of 26,570 rows of log data and 7,560 captured video frames for the first run through the verification process. For the first attempt, the system could not be verified, and no plausible synchronisation between the object log and the video transcript could be found. It was then determined that the ray tracing did not work properly, actually invalidating the log data in a way that was not obvious by looking at the log file alone. The verification study had to be repeated once this fault was corrected, this time generating 13,285 game engine log entries and 3,780 transcribed video frames entries for a 5 minute play session.

Analysis of the log data in comparison with the video transcript results in an accuracy model for the system. Accuracy is the degree of veracity (or closeness to the truth) of a measure, while precision is the reproducibility of the measure of accuracy. In the verification study, data was extracted from the sorted and analysed object log and video transcript table to characterise accuracy in both the spatial and the temporal performance of the system. Spatial accuracy is assessed by comparison with a known actual number of objects within the region of accuracy of the eyetracking system. The result can be expressed in terms of: for a given record for an object within the object log, what is the probability that there were actually n objects under the gaze point, for an arbitrary +ve integer n? Temporal accuracy is assessed by comparison of log entry time with the best estimate of the actual time when a specific object fell under the gaze point during one simulation time cycle of the game engine. The result can be expressed in terms of: for a given record for an object within the object log occurring at a time tlog, what is the probability that the actual time at which the gaze occurred is tlog +/- toffset, for toffset in some known real number range of seconds? Both temporal and spatial precision were not analysed in detail, since this would require a comparison across several studies, requiring an impractical amount of investigation time.

6. Results
After analysing the data and categorized frequency count of time offsets, human and technical errors it was found that system performance in this study has an 81% time accuracy of total offsets of log entries, meaning that 81% of log entries fall within 1 standard deviation of the time noted in the log file (see Figure 8). The positive offset time error to the extreme right represents cases that are uninterpreted due to image occlusion by a gaze fixation indicator.

Fig.8. Frequency count of time offsets of object log entries from object visibility in the video transcript.
The verification study also resulted in a spatial accuracy characterization, expressed in terms of the probability of \( n \) objects falling within the spatial accuracy of the eyetracker, when a single object is logged under the gaze point by the game engine. As shown on Figure 9, one object logged corresponds with one object within the gaze accuracy area in 89% of cases.

7. Discussion and Conclusion

The verification study described in this report has shown that the integrated game engine and eyetracking systems can operate together with a quantifiable accuracy that is sufficient to conduct statistical studies of eye gaze behaviour in relation to dynamic, 3D stimuli. There are some changes that could improve the accuracy of the verification study, such as improving accuracy over time by higher frame/sampling rates in the game engine, synchronised start times for the game engine and video capture, and having a cursor on top of the video file that is transparent so verification objects are never obscured and can be more easily interpreted (e.g. as provided by Senso Motoric Instruments, SMI, http://www.smi.de/home/index.html, which uses an unfilled contour as a marker/cursor for representing gaze data superimposed on top of captured video). The accuracy of the integrated system could be increased by higher frame/sampling rates in the game engine and/or eyetracker, and by higher spatial resolution of the eyetracker. Spatial accuracy is highly dependent upon the design of the stimulus and can be increased by having collision volumes that closely bound graphical meshes, and minimising high spatial frequency areas of a level, especially in the form of reducing clustering of small objects or object parts and boundaries within the scale of accuracy of the eyetracker. There is also a fundamental question of how graphics are assimilated to the concept of an object of interest. E.g. for a human character, is an object of interest the whole character, a body part such as a head or a finger, or a pocket on clothing? The answer to this may depend upon the goals of a study, a particular study participant, their tasks, etc., and in this sense is a factor to be determined by a study rather than presumed. More flexibility in the system in this respect would be facilitated by using collision volumes organized as hierarchical object assemblies, e.g. representing a whole character and their perceivable subdivisions; this could, however,
greatly complicate stimulus development. In any case, it is clear that the method can be used for investigating 3d gameplay with statistically significant participation group sizes. The basic gaze object analysis time then decreases from 45 hours per 30 minutes of play time [18] to 0.

Accuracy considerations are relevant to the use of gaze tracking for game and virtual world interaction, as well as for analysis. Gaze interaction has been used to provide access to common computer functions, such as email, word processing and web browsing for people who are unable to use mouse and keyboard interaction (e.g. http://www.tobii.com/assistive_technology/products/product_overview.aspx). Eye gaze tracking systems are expected to decrease in cost, eventually to a level of accessibility as mass market products. This raises the possibility that gaze-based interaction could become a common feature of mass market software systems, including computer games. There have been a number of experiments using gaze-based interaction for control of computer gameplay. For example, [19] describes a gaze-driven chess game, [16] compares several mouse and gaze interaction modes for the games Sacrifice and Half-Life, [20] compare gaze and mouse steering in the Breakout game, and [21] compares gaze and mouse control for the simple Blob game. The commercial game industry is also reported to have expressed interest in a controller that can simplify interaction [22].

The use of gaze tracking as an interaction technology requires design practices that minimize the impact of error sources, such as the mismatch of collision volumes with graphics, upon the player experience. The observed timing inaccuracies are not likely to have a major impact in computer gameplay, since they are within a small number of frames at a high frame rate. Spatial ambiguities can have an impact, and this should be minimized by minimizing spatial clustering and reducing requirements for game actions that occur at large (virtual) distances within the game design.

8. Acknowledgements

This work has been conducted in the context of a joint project between Gotland University, Blekinge Institute of Technology and Swedish Defence Research Agency, Sweden, and also in the context of the EU FUGA (NEST- 28765) project.
References


PART IV
EMPIRICAL EXPERIMENTS & RESULTS

An Investigation of Visual Attention in FPS Computer Gameplay

Sennersten, C., Castor, M., Hedström, J., Lindahl, B., Lindley, C., Svensson, E.
Blekinge Institute of Technology, Karlsknam, Sweden
charlotte.sennersten@bth.se, craig.lindley@bth.se

Abstract
Cognitive science provides a useful approach to studying computer gameplay, especially from the perspective of determining the cognitive skills that players learn during play. Computer games are highly visual medium and game interaction involves continuous visual cognition. A system integrating an eyetracker with a computer game engine allows real time gaze object logging, providing a fast and convenient way of collecting data for analysis. Such a system has been used to test three hypotheses concerning visual attention in a first-person shooter (FPS) computer game. Firstly, the cuing effect of the passive gun graphic on visual attention was tested, but no evidence was found to support this. Data supported the second hypothesis, that a player attends to the target opponent while shooting at them, in most cases, while in a small percentage of cases this is achieved in peripheral vision. Finally, in most cases, a player targets the nearest opponent. These results provide a baseline for further investigations in which the stimulus game design may be modified to provide more detailed models of the visual cognitive processes involved in gameplay.

Keywords: Visual Attention, Eye-Tracking, Decision-making, Gameplay

Introduction
Recent trends in computer game research include the use of empirical methodologies for gaining a better understanding of gameplay behavior. While much of this empirical work has been based upon questionnaires for testing player self-evaluations over multiple gameplay sessions (e.g. Yee, 2006), psychophysiological and neuroscientific methods are emerging for studying emotional factors of gameplay (Ravaja et al, 2008), detailed activation of neurophysiological areas (Mathiak et al, 2006) and the orchestration of visual attention processes in cognitive task performance in gameplay (Lindley and Sennersten, 2007) at a level of fine detail (e.g. milliseconds) within play sessions. These more detailed methods are laying the foundations for a much deeper understanding of the motivational and reward factors that lead players to continue playing, as well as the cognitive skill learning outcomes of gameplay. This knowledge may be of benefit for entertainment game development in providing designers with a more scientific understanding of the foundations of ‘fun’ and how the emotional substructure of fun relates to details of design features. Scientific studies of gameplay also provide direct benefits for the development of serious games, where explicit cognitive skill training and emotional orchestration within decision processes are the primary drivers of game designs.
The development of detailed cognitive and psychophysiological understandings of gameplay and its affects is a large scale project that will require many different individual studies due to the high resolution with which games are investigated, and the very large space of possible combinations and detailed layouts of game design features that may influence player cognitive and emotional affects under study. This paper presents one such study, addressing firstly the question of the degree to which a passive graphical element, the gun position, in a first-person shooter computer game may have an influence in cuing visual attention during gameplay task performance, secondly the extent to which visual attention is focused upon the opponent before/while shooting, and thirdly, the effect of virtual distance.

**Hypotheses Deriving from the Task Context**

3D FPS (first-person shooter) game environments are virtual worlds in which the player moves their avatar, a player character from the visual perspective of which the game world is viewed, through the game world and overcomes barriers in the form of enemies by engaging and defeating them in combat. Primary player tasks are survival, movement and navigation through the world, and executing tactical combat operations, typically in the form of shooting at enemies until they are dead.

Visual attention is the main perceptual instrument for managing challenges in this kind of setting. Visual perception provides information about the nature and location of challenges/threats, providing input to cognitive decision processes addressing where and how to move, and how and when to trigger attack commands, triggered by the player using a mouse or command key and implemented by the simulation engine of the game as the activation of a directed weapon discharge. 3D FPS worlds have virtual depth, so task performance must also take virtual distance into account.

The detailed decision process of a player can be modeled in the form of a cognitive task model, representing a cognitive schema or algorithm for task performance (Lindley and Sennriesten, 2006, 2007). The development of gameplay schema models represents a key outcome for understanding the mechanisms of gameplay, representing the results of the cognitive learning process of learning to play and succeed within a particular game. The hypotheses investigated in the current study address detailed aspects of visual processing in gameplay task performance, as an aspect of developing a more complete schema model for FPS gameplay. That is, firstly, to what extent is visual attention cued by a graphical representation of the player weapon that has no function in aiming the weapon? Secondly, does a player always look at a target while shooting?

The weapon representation in an FPS game is the first person view of the virtual gun that the player avatar holds. The actual aiming point for the player firing a weapon in the virtual world is represented by a cross-hair graphic in the centre of the screen. Gibson (1977, 1979) used the term *affordances* to refer to action potentials within an environment. In an FPS
game, a representation of the aiming point of a virtual weapon, such as a gunsight, affords aiming the virtual weapon. In the case of a static gun graphic with a separate active aiming point (e.g. a central crosshair in the screen), the gun graphic actually has no affordance value for aiming. The graphic is usually visually active, in that it moves to represent the walking or running motion of an avatar, and typically has animations indicating firing states and reloading actions. It also provides a direct representation of which weapon is currently selected and active for the player. These indicators present informational inputs to decision processes, indicating the satisfaction or not of preconditions for actions, but they do not directly locate the objects of actions contributing to achieving in-game goals. The graphic is also static in the sense that it remains on the computer screen representing the character field of view as the character moves or rotates within the game world.

If the gun graphic provides a cue leading visual attention away from potential targets or the aiming cross during targeting during combat, then its design may be a distraction from optimal task performance. The hypothesis that this kind of distraction may occur derives from an earlier study (Sennersten, 2004) in which eye gaze tracking was used to investigate gameplay in a Counterstrike tutorial. In the earlier study, it was found that: 1. slightly more than 50% of gaze behavior fell within the left of the game display, where a priori an even distribution between left and right might be expected, and 2. approximately 50% of players visually fixate on the far end of the graphical representation of the barrel of the gun, that actually has no functionality in the performance of shooting tasks, and in the region in between the end of the gun and the actual aiming cross sight. These results motivate two of the hypotheses investigated in the current study:

H1: the diagonal shape and position of the gun graphic cues and directs visual attention within the screen, preceding and/or independently of the attentional demands of the player’s in-game task(s) (Figure 1). This hypothesis is of interest for many serious games contexts where graphical elements designed into game environments may or may not represent functional affordances. If a graphical element does not represent direct task-oriented affordances, how is the passive graphical design relevant, is it just decoration with no purpose, might it distract from task performance and what are the implications of this for the transfer of player performance competence to out-of-game contexts? Might the graphical design be changed to eliminate task distractions and instead cue visual attention to task relevant perceptual elements?

H2: gaze is firstly directed upon an opponent prior to shooting the opponent. This is a counter-hypothesis to finding 2 from (Sennersten, 2004) as noted above. H2 relates to the degree to which the most important visual perceptual target for the performance of a task is at the centre of vision while the task is performed, as opposed to being attended to in peripheral vision.
As noted in Lindley and Sennersten (2006), computer games provide a rich task environment for cognitive experiments, and as such it is possible to design experiments that use a single stimulus (game/level) to address multiple hypotheses. For the current study, the combat situation was also considered as a source of situation and task specific priorities in decision making. In particular, in addition to hypotheses H1 and H2, a further fundamental principle of combat is addressed by the hypothesis:

H3: for otherwise equivalent opponents, the closest opponent will be targeted first. Equivalence here relates to similar toughness (ease or difficulty to kill), potency (amount of harm they can inflict upon the player) and accessibility/visibility.

H3 may seem obvious, but the point of developing a scientific approach to gameplay is to provide quantitative evidence for what may otherwise be taken for granted or assumed. Data collected for H3 and H1 may also provide an indication of the relative importance of visual cuing and (virtual) proximity if a cuing effect is found. Moreover, this study functions as a baseline, providing data for comparison with results from ongoing experiments that may be designed to investigate the distinction between apparent size and apparent virtual distance (i.e. is it the closest opponent or the one that appears the largest that is chosen first?), and then exploring the relative importance of factors such as variations in apparent toughness and potency in relation to distance and apparent size. Underlying questions here from a cognitive skill perspective include the degree to which players may be attempting to optimize combat behavior by the allocation of importance among these factors, what influences these weightings (experience, PC role preferences, adapting the level of experienced challenge?), the interplay between emotional responses and rational decision-making, etc.
Apparatus

Eye gaze behavior can be captured by an eyetracking system. The eyetracker used for the work described in this paper is a Tobii Eye Tracker 1750 in which the eyetracking technology is integrated with a graphical screen upon which the stimulus is displayed. The 1750 runs on a PC running Windows (Windows 2000 service pack 3 or Windows XP service pack 1 or 2). The eyetracker delivers an \((x, y)\) coordinate representing the gaze position in 2D screen coordinates. The study used a double computer setup with the game application running on a separate computer from the eyetracking software but displaying the game via the eyetracker screen. The game computer executes the stimulus game, in this case using the HiFi game engine developed by the Swedish Defence Research Agency (Försvarets Forsknings Institut, FOI).

Traditionally, eye tracking studies of gaze behavior on computer screens have been limited to static 2D stimuli analysed in terms of the statistics of gaze falling within static 2D subareas (e.g. Holsanova, Rahm and Holmqvist, 2006), generally referred to as Areas Of Interest (AOIs). In a 3D game setting we are more interested in Volumes Of Interest (VOIs), or Objects of Interest (OOIs), as objects of gaze, which move in relation to their projection onto the 2D surface of the computer screen. While OOIs can be identified by examining a plot of gaze positions superimposed over the game display, this is extremely time-consuming. For our studies, in collaboration with FOI we developed a system in which the eyetracker is integrated with the HiFi game engine so that the engine received gaze coordinates from the eyetracker and performs a ray trace (see Duchowski, 2003/2007) within the game world from the gaze coordinate to the first intercepted object within the game world. The HiFi engine then records an object log entry for the gaze point, including the time, gaze coordinates, id and location of the intercepted object. The resulting object log then includes information on all objects under the gaze point for each participant and each experimental session. Details of the integrated system, its verification process and accuracy characterization are presented in Sennersten and Lindley (2006), and Sennersten et al. (2007).

A typical play session within a gameplay study using this system involves the player participant playing the game in front of the eyetracker. After briefing the player with any instructions relevant to the study, the calibration process is run. Then the eyetracker screen is switched over to the game display. The operator starts game object logging and at an instruction from the operator, the player commences playing while eyetracking and game engine object logging are running. Logging stops at the end of the session and the log data is saved for analysis.

Experiment Design

In order to test hypothesis H1 in a 3D FPS game, an experiment design is required in which it is possible to discern the effect of the position of the diagonal gun graphic in influencing the direction of visual attention during weapon firing actions. This can be
achieved by varying the gun graphic between pointing toward the upper left from the lower right, as shown on Figure 1, and pointing toward the upper right from the lower left, as shown on Figure 2.

![Figure 2. The alternate Left gun position.](image)

These variations must occur in situations where there is a choice of who to shoot first among combat opponents on each of the left and right sides of the screen. Assuming no other biasing factors (and with handedness being tested by questionnaires during the study), if there is no visual cuing effect, then the choice between shooting the left or the right opponent first should be random and hence equally probable, leading to them having comparable frequencies of selection during gameplay. H3 can be addressed by presenting the further variable of opponent distance. Hence, if two opponents are encountered, one may be nearer than the other. Again, if distance is not a decision factor, then the near and far opponents will be equally likely to be selected, and hence have comparable selection frequencies for first attack during play.

From these considerations it is sufficient to provide a decision situation for the study where a player attack decision is made upon encountering a pair of opponents, one on the left and one on the right, in combinations of near and far distance and with either a Left or Right gun position. The dependent variable for a decision/choice point is the decision about which opponent to attack first. Each decision point can be characterized by 3 independent variables having the binary values:

- **V1: Distance of L_opponent**
  - Value 1: Near
  - Value 2: Far

- **V2: Distance of R_opponent**
  - Value 1: Near
  - Value 2: Far

- **V3: GunGraphic**
  - Value 1: Left
  - Value 2: Right
H2 is also accommodated by these variables. To make these independent variables statistically valid there should be at least 10 samples of each combination of possible values for the three variables. This is an estimate is based upon assumptions of multiple regression (Pallant, 2005). Multiple regression analyses will not be reported in this paper, but is used as a criterion for the number of samples in order to support later data correlations. In this case there are 3 independent variables with binary values that make for 8 possible combinations. To obtain 10 samples of each of 8 possible combinations, a single experimental run must therefore present a participant with 80 combat encounters, each of which is a decision situation addressing which opponent to shoot first. Each encounter must have both a Left and a Right opponent, at standard distances to the left and right, and the Near and Far distances are also standardized to be the same for each encounter. Randomising the gun graphic position from one encounter to another would be disturbing and attract attention in a way that is atypical in gameplay, so the first half of the encounters for each participant use one gun position which then changes to the other position for the second half of the encounters.

Too few samples may also lead to a result that does not generalize (cannot be repeated) with other samples, and is hence of little scientific value (Pallant, 2005). Hence 20 participants were involved in the experiment, resulting in data for 1600 combat encounters.

**Stimulus Design**

The stimulus is a computer game level implementing the combat encounters to the specification described above. Only 40 encounters needed to be implemented, since the 80 encounters experienced by the player can repeat a sequence of 40 encounters but using the alternate gun position, with the starting gun position being randomized for each player. Each combat encounter was designed to take place within a single room, while the level as a whole is a series of interconnected rooms. Every room contains two NPC’s (non-player characters), with one each on the left and right of the room, and with their Near/Far distance positioned randomly so that players can’t predict their distance beforehand.

The enemy positions within the 40 rooms of the level used the same layout and distances for each room. The LeftNear opponent and the RightNear opponent have a 2 (virtual) meter distance from the player character (PC) soldier’s entry, and the lateral distance between the two opponent soldiers is also 2 meters. The LeftFar and RightFar opponent positions are 10 meters from the PC soldier’s entrypoint to a room, five meters in depth from one another. The PC enters each room through a narrow passage and with vision blocked by a blind until within the room in order to ensure some consistency in the configuration of opponents when visually encountered. However, the distances and visual angles vary a little for each player and each room depending upon whether the PC is running or walking, orientation and also according to some freedom in position when a room is entered.

Figure 3 shows the top-down map of the game level stimulus from above. The player participants start out from the bottom right corner of the map. All rooms are approximately
of the same size and the left and right handed curves of the map occur in equal numbers to eliminate visual biases due to path topology that may affect the result. The level has been created in the BattleCraft Editor developed for the Battlefield 1942 game by the game development company DICE.

Figure 3. TopDownView, from the editor, of the stimulus level with 40 rooms

Results
This section summarises results obtained for all rooms and players, covering the 1600 data samples (20 participants with 80 enemy encounters each) obtained in the study. Results are considered in relation to the three hypotheses of the study. In the data representations, ‘N=200’ means that all 20 player participants experienced 10 occurrences of a specific Left/Right, Near/Far combination, adding up to 200 samples of possible visual attentive choices for each combination.

H1
H1, that the diagonal shape and position of the gun graphic cues and directs visual attention within the screen, preceding and/or independently of the attentional demands of the player’s in-game task(s), is most easily tested using the subset of data in which both opponents are either near or far. In this case, evidence for H1 will show preferential attention upon either the left or the right opponent, correlated with a specific gun position.
Right Gun Position

Figure 4. Attention towards the LeftNear NPC occurs \( f = 101 \) times (50.5%) and attention towards the RightNear NPC occurs \( f = 97 \) times (48.5%) \((N=200)\). “No Data” means no registered gaze, which occurs twice.

Figure 4 shows the sum of choices for the first opponent looked at when the opponents were in the LeftNear and RightNear positions for all encounters. The attentional distribution differs by only 2 percent, indicating no significant cuing effect.

![Graph showing frequency of attention in different positions]

Figure 5. “LeftFar” has \( f = 99 \) (49.5%) and “RightFar” has \( f = 99 \) (49.5%), “No Data” has \( f = 2 \) (1%) \((N=200)\).

Figure 5 shows the sum of choices for the first first opponent looked at when the opponents were in the LeftFar and RightFar positions for all encounters. The attentional distribution differs by only 1 percent, indicating no significant cuing effect.

Left Gun Position

Figure 6 shows the sum of choices for the first first opponent looked at when the opponents were in the LeftNear and RightNear positions for all encounters. The attentional distribution differs by 19.5 percent, indicating a possible cuing effect in favour of the Left opponent. However, this is not in the side of the screen towards which the gun graphic is pointing and hence does not provide evidence for H1.
Figure 6. “LeftNear” has $f = 114$ (57%) and “RightNear” has $f = 77$ (38.5%), “No Data” has $f = 9$ (4.5%) (N=200).

Figure 7 shows the sum of choices for the first first opponent looked at when the opponents were in the LeftFar and RightFar positions for all encounters. The attentional distribution differs by only 4 percent, indicating no significant cuing effect.

Figure 7. “LeftFar” has $f = 102$ (51%) and “RightFar” has $f = 94$ (47%), “No Data” = 4 (2%) (N=200).

**Conclusion for H1**

These results for both the left and right gun graphic positions represent the random choice case in relation to H1, providing no evidence in favour of H1 and suggesting that the gun graphic has no effect in cuing visual attention towards the left or right of screen.

**H2**

There are two possible alternative attention behaviors in relation to, H2, that *gaze is firstly directed upon an opponent* prior to shooting the opponent, addressed in this study. Firstly, the player fires before looking at any opponent target at the centre of vision. This is labeled “Fire_Look” (for ‘fire then look’). Secondly, the player directs their gaze towards one opponent while firing at the other NPC. This case is labeled “Dual_Att” (for dual
or divided attention). The tables in this section present the relationship between looking and firing at an opponent. For each combination of the opponent positions (LeftNear, LeftFar, RightNear and RightFar), the frequency is given for which opponent was fired upon first. Frequencies are also given for the Fire_Look and DualAtt cases. Tables 1 to 4 show the results for the RightGunPosition for each of the four possible combinations of opponent position.

“Dual_Att” has an overall occurrence for the RightGunPosition of 7%, while “Fire_Look” has an overall occurrence of 4.5%. These figures represent 11.5% of cases where an opponent is not first looked upon and then directly shot at.

<table>
<thead>
<tr>
<th>Frequency</th>
<th>Valid Percent</th>
<th>Cumulative Percent</th>
</tr>
</thead>
<tbody>
<tr>
<td>Valid LeftNear</td>
<td>93</td>
<td>46.5</td>
</tr>
<tr>
<td>RightNear</td>
<td>84</td>
<td>42.0</td>
</tr>
<tr>
<td>Fire_Look</td>
<td>9</td>
<td>4.5</td>
</tr>
<tr>
<td>Dual_Att</td>
<td>14</td>
<td>7.0</td>
</tr>
<tr>
<td>Total</td>
<td>200</td>
<td>100.0</td>
</tr>
</tbody>
</table>

Table 1. “Fire_Look” = 4.5% of cases and “Dual_Att” = 7% (N=200).

<table>
<thead>
<tr>
<th>Frequency</th>
<th>Valid Percent</th>
<th>Cumulative Percent</th>
</tr>
</thead>
<tbody>
<tr>
<td>Valid LeftNear</td>
<td>158</td>
<td>79.0</td>
</tr>
<tr>
<td>RightFar</td>
<td>28</td>
<td>14.0</td>
</tr>
<tr>
<td>Fire_Look</td>
<td>8</td>
<td>4.0</td>
</tr>
<tr>
<td>Dual_Att</td>
<td>6</td>
<td>3.0</td>
</tr>
<tr>
<td>Total</td>
<td>200</td>
<td>100.0</td>
</tr>
</tbody>
</table>

Table 2. “Fire_Look” = 4% and “Dual_Att” = 3% (N=200).

<table>
<thead>
<tr>
<th>Frequency</th>
<th>Valid Percent</th>
<th>Cumulative Percent</th>
</tr>
</thead>
<tbody>
<tr>
<td>Valid LeftFar</td>
<td>23</td>
<td>11.5</td>
</tr>
<tr>
<td>RightNear</td>
<td>144</td>
<td>72.0</td>
</tr>
<tr>
<td>Fire_Look</td>
<td>14</td>
<td>7.0</td>
</tr>
<tr>
<td>Dual_Att</td>
<td>19</td>
<td>9.5</td>
</tr>
<tr>
<td>Total</td>
<td>200</td>
<td>100.0</td>
</tr>
</tbody>
</table>

Table 3. “Fire_Look” = 7% and “Dual_Att” = 9.5% (N=200).

189
Tables 5 to 8 show the results for the LeftGunPosition for each of the four possible combinations of opponent position. “Dual_Att” has an overall occurrence for the LeftGunPosition of 7.8%, while “Fire_Look” has an overall occurrence of 5.5%. These figures represent 13.3% of cases where an opponent is not first looked upon and then directly shot at.
Figure 7 summarises the frequencies of the “Fire_Look” for all opponent combinations and for both gun positions. Here the combination LeftFar/RightNear for the LeftGunPosition creates the most cases of firing on an opponent before looking directly at any opponent.

<table>
<thead>
<tr>
<th></th>
<th>Frequency</th>
<th>Valid Percent</th>
<th>Cumulative Percent</th>
</tr>
</thead>
<tbody>
<tr>
<td>Valid LeftFar</td>
<td>26</td>
<td>13.0</td>
<td>13.0</td>
</tr>
<tr>
<td>RightNear</td>
<td>137</td>
<td>68.5</td>
<td>81.5</td>
</tr>
<tr>
<td>Fire_Look</td>
<td>19</td>
<td>9.5</td>
<td>91.0</td>
</tr>
<tr>
<td>Dual_Att</td>
<td>18</td>
<td>9.0</td>
<td>100.0</td>
</tr>
<tr>
<td>Total</td>
<td>200</td>
<td>100.0</td>
<td></td>
</tr>
</tbody>
</table>

Table 7. “Fire_Look” = 9.5% and “Dual_Att” = 9% (N=200).

Figure 8 presents a summary for the “Dual_Att” category. Here the combination LeftFar/RightNear for the RightGunPosition creates the most cases of firing on one opponent while looking directly at the other.

<table>
<thead>
<tr>
<th></th>
<th>Frequency</th>
<th>Valid Percent</th>
<th>Cumulative Percent</th>
</tr>
</thead>
<tbody>
<tr>
<td>Valid LeftFar</td>
<td>94</td>
<td>47.0</td>
<td>47.0</td>
</tr>
<tr>
<td>RightFar</td>
<td>83</td>
<td>41.5</td>
<td>88.5</td>
</tr>
<tr>
<td>Fire_Look</td>
<td>7</td>
<td>3.5</td>
<td>92.0</td>
</tr>
<tr>
<td>Dual_Att</td>
<td>16</td>
<td>8.0</td>
<td>100.0</td>
</tr>
<tr>
<td>Total</td>
<td>200</td>
<td>100.0</td>
<td></td>
</tr>
</tbody>
</table>

Table 8. “Fire_Look” = 3.5% and “Dual_Att” = 8% (N=200).
Conclusion for H2
The overall conclusion is that hypothesis H2, that *gaze is firstly directed upon an opponent prior to shooting the opponent*, appears to be correct in about 88% of initial encounters. The remaining 12% represents cases where aiming at the moment of firing is achieved by peripheral vision.

H3
H1, that *for otherwise equivalent opponents, the closest opponent will be targeted first*, is most easily tested using the subset of data in which one opponent is near while the other opponent is far. In this case, evidence for H3 will show preferential attention upon the closest opponent, irrespectively of its left or the right opponent or the gun position.

Right Gun Position
Figure 9 shows the sum of choices for the first opponent looked at when the opponents were in the LeftNear and RightFar positions for all encounters. The attentional distribution differs by 62 percent in favour of the nearest opponent, suggesting a significant preference for the near opponent.

![Figure 9](image)

Figure 9. “LeftNear” has f = 162 (81%) and “RightFar” has f = 38 (19%). “No Data” = 0 (N=200).

Figure 10 shows the sum of choices for the first opponent looked at when the opponents were in the LeftFar and RightNear positions for all encounters. The attentional distribution differs by 56.5 percent in favour of the nearest opponent, suggesting a significant preference for the near opponent.

![Figure 10](image)

Figure 10. “LeftFar” has f = 43 (21.5%) and “RightNear” has f = 156 (78%). “No Data” = 1 (0.5%) (N=200).
**Left Gun Position**

Figure 11 shows the sum of choices for the first opponent looked at when the opponents were in the LeftNear and RightFar positions for all encounters. The attentional distribution differs by 56 percent in favour of the nearest opponent, suggesting a significant preference for the near opponent.

![Graph showing the distribution of choices for LeftNear, RightFar, and No Data](image)

Figure 11. “LeftNear” has $f=155$ (77.5%) and “RightFar” has $f=43$ (21.5%). “No Data” = 2 (1%) (N=200).

Figure 12 shows the sum of choices for the first opponent looked at when the opponents were in the LeftFar and RightNear positions for all encounters. The attentional distribution differs by 57.5 percent in favour of the nearest opponent, suggesting a significant preference for the near opponent.

![Graph showing the distribution of choices for LeftFar, RightNear, and No Data](image)

Figure 12. “LeftFar” has $f=48$ (24%) and “RightNear” has $f=143$ (71.5%). “No Data” = 9 (4.5%) (N=200).

**Conclusion for H3**

These results show strong evidence in favour of hypothesis H3, that in an average of 77% of cases the closest opponent will be looked at first, irrespectively of the left or right gun graphic position.
Discussion and Conclusion

This experiment has resulted in no evidence in support of hypothesis H1, hence the graphical representation of the weapon does not appear to be cuing visual attention. Hypothesis H2 is supported most of the time, with only 12% of firing events being carried out using peripheral vision. Finally, H3, that the closest opponent will be the first one to be shot at, is also strongly supported (77% of firing events).

In general the graphical representation of the weapon, which does not have any affordances for firing the weapon (although it does carry relevant state information) does not have a strong role in cuing visual attention. The position of the weapon graphic also has no strong discriminating power for determining whether peripheral vision will be used or not, or for determining circumstances under which either the nearer or further opponent will be targeted first.

These results do raise the question of what factors account for the 12% of cases that do not conform to H2 and the 23% of cases that do not conform to H3. These cases may arise due to specific details of the geometry of the encounter, or may arise as stochastic variations in the decision processes of players. It also remains to be determined whether there is any correlation in these cases with other features of players, such as their gameplay experience, play preferences, handedness, etc.. Answering these questions may be facilitated by further investigation of the data collected by this study, or may require the design of ongoing experimental studies.
References


Ongoing research directions

There are many potential ongoing research directions that can be pursued based upon the work reported in this thesis.

The empirical studies that have been conducted so far have generated a great deal of data. Analysis of this data so far, as reported in Sennersten et al (2007, 2008) and Sennersten and Lindley (2008) merely scratches the surface of what may eventually be obtained from it. Additional analysis can address the pace of learning over multiple play sessions, different styles and forms of gameplay, correlating different styles of gameplay with player features determined by questionnaires, and correlating variations of eye gaze behavior with player features determined by questionnaires.

Modelling different play styles includes the surface description of gameplay patterns, or gameplay gestalts (Lindley, 2002). These descriptions can form the basis for defining more systematic gameplay task or schema models. A schema modeling methodology for computer games can be defined and schema descriptions developed, for basic analysis, for comparative studies of different players, and to represent cognitive skill learning outcomes of gameplay. Schema models can also potentially be integrated within computer games as a basis for player performance monitoring and adaptation of game mechanics (Lindley and Sennersten, 2007).

FPS games were chosen for this study since FPS gameplay involves strict goal-directed activity that requires high visual attention due to the needs of survival. FPS games have less noise to complicate cognitive activity than some other game forms, while presenting demanding challenges with wide scope for behavioural variations in meeting those challenges. After starting out with game levels with low workload and minimum complexity, ongoing studies can increase the complexity of the game stimuli and/or the difficulty of game challenges to see how gameplay behaviours and underlying schemas change accordingly. Further complexities can be explored by considering multiplayer FPS games.

Similar studies can be conducted with role playing games, where there is more diverse gameplay, social interaction with both other players and computational non-player characters (NPCs), and lower pace in overall gameplay. Cognitive tasks may change to deal with other issues than those that are typical for FPS games, including a wider variety of tasks and challenges, such as navigation, puzzle solving, trading, inventory management and character configuration. These differences will have corresponding differences in gaze distribution indicating different underlying schemas than those in a typical FPS.

A further direction for ongoing research is to consider psychophysiological measurements as inputs to a game engine, and to investigate how game mechanics can take advantage of these inputs to create new and interesting forms of game experience. Other kinds of...
interfaces can also be explored, including those supporting gameplay in a physical space and where audiovisual content can be synchronized and investigated with sensors in relation to theater, dance, music creation and gaze. Studies of the effects of gameplay in these new environments can be conducted using similar apparatus and techniques to those used to study conventional computer games.

These directions in potential ongoing research represent a very broad scope for ongoing work, requiring many researchers and groups, and it is hoped that this thesis will contribute to providing some kind of basis and inspiration for others undertaking this ongoing program of research.
Epilog

Being part of the big picture of game research and society in the early 21st century is both fun and troublesome. The painful part is that game research and also the game subject have an independent history no longer than 7-10 years back in time, in the academic field. Of course there have been activities going on before this, but only as minor parts of other programs or disciplines.

Young and old players and gamers are not a minority group any longer, they are everywhere and may play games at any time of the day. Beyond entertainment, this project has relevance to many different domains and areas and there is an increasing interest not just in studying and tracking gameplay behaviors but also in using biometrics as input data for affecting the game state or other kinds of behavior of the game engine. Applying game design beyond entertainment, e.g. for training and learning, requires answers to questions of the transferability of skills learned by gameplay. It is hoped that the method develop here may contribute to answering some of these questions.

The government has a great responsibility to take this research field seriously because game media are actually changing the knowledge domains of players to something that has not been investigated thoroughly nor been taken seriously outside a limited number of fields, such as military application areas. What is this knowledge about? We are more or less interacting with visual content and this visual content provides the basis for knowledge construction which is fundamentally different to text based content. Knowledge gained via non-textual dynamic visual semiotics may dominate the word, leading to knowledge built through the word to decrease in importance and value.

Games may provide a way of gaining knowledge about aspects of the world situation by situated learning. Recorded videos or 2D presentations of news do not give experience-based knowledge about important events in the world; those medias are informative and descriptive but not self-evaluated in action. Games can give a useful practice of “being in the world” for understanding other kinds of world contexts in a way that cannot be read about or watched. Political Science, Economics, Serious Games and Entertainment have different takes on gameplay and game design, and have reward systems for both the same and different objectives. They are also forms of games in themselves, which games as media can provide direct experience of.
References


ABSTRACT

Research into gameplay can contribute to more self-conscious approaches to design, allowing designers to create effective gameplay with less testing or to target specific cognitive and emotional affects of gameplay for serious games applications. Self-conscious design includes theoretically-motivated design of game systems to facilitate gameplay motivated by cognitive, scientific and/or rhetorical theories of game affect and functionality. Deepening the understanding of gameplay requires a consideration of basic epistemological questions about the nature of understanding. Understanding gameplay is a matter of generating mappings to explanatory frameworks in alternative interpretation paradigms. All games are cognitive skill learning environments, and an especially useful approach that may aid in the creation of more self-conscious game design practices is to conduct research into gameplay using theories and methods of cognitive science and cognitive psychology. On this basis, a framework is proposed based upon the integration of schema theory with attention theory. Cognitive task analysis provides a foundation for developing schema descriptions, which can then be elaborated according to more detailed models of cognitive and attentional processes. This approach provides a rich explanatory framework for the cognitive processes underlying gameplay. Playing a commercial PC or console game is a highly visual activity regardless of whether the purpose is entertainment or situated learning. Information about the visual attention of the player is an important foundation for detailed schema modeling. A range of different eye-tracking equipment has been used in many studies of visual cognition. However, very few studies describe dynamic stimuli involving the visual interaction of a user/player with a moving 3D scene displayed on a computer screen. In order to address this, a software interface has been developed linking a Tobii™ eyetracking system with the HiFi game engine for use in automated logging of dynamic 3D objects of gaze attention. The system has been verified in a detailed study, confirming correct operation of the system as well as providing a characterisation of its spatial and temporal accuracy. The integrated Tobii/HiFi system has been validated in a study to test three hypotheses concerning visual attention in a first-person shooter (FPS) computer game. Firstly, the cuing effect of the passive gun graphic on visual attention was tested, with no evidence being found to support this hypothesis. A second hypothesis, that a player directs their gaze at a target opponent while shooting at them, was found to be supported in most cases, while in a small percentage of cases targeting is achieved in peripheral vision. Finally, in most cases, a player targets the nearest opponent. These results provide a baseline for further investigations in which the stimulus game design may be modified to provide more detailed models of the visual cognitive processes involved in gameplay and how they are involved in player decision-making.

GAMEPLAY

(3D GAME ENGINE + RAY TRACING = VISUAL ATTENTION THROUGH EYETRACKING)

Charlotte Sennersten

Blekinge Institute of Technology
Licentiate Dissertation Series No. 2008:10
School of Technoculture, Humanities and Planning