EXPLORING PERSONALITY TYPE FOR COMPOSITION-BASED GROUP USER MODELLING IN AN EMBODIED INTERACTION SYSTEM

by

Ying Jiang
B.Sc., Beijing Information Science and Technology University 1997

PROJECT SUBMITTED IN PARTIAL FULFILLMENT OF
THE REQUIREMENTS FOR THE DEGREE OF

MASTER OF APPLIED SCIENCE IN INTERACTIVE ARTS

In the
School of Interactive Arts and Technology

SIMON FRASER UNIVERSITY

Summer 2008

APPROVAL

Name: **Ying Jiang**

Degree: **M.A.Sc. in Interactive Arts**

Title of Thesis: **Exploring Personality Types for Composition-Based Group User Modelling in an Embodied Interaction System**

Examining Committee:

Chair:

Dr. Tom Calvert
Professor of School of Interactive Arts and Technology

Prof. Ron Wakkary
Senior Supervisor
Associate Professor of School of Interactive Arts and Technology

Dr. Marek Hatala
Supervisor
Associate Professor of School of Interactive Arts and Technology

Dr. Dragan Gasevic
External Examiner
Assistant Professor, School of Computing and Information Systems, Athabasca University

Date Defended/Approved: July 18, 2008

STATEMENT OF ETHICS APPROVAL

The author, whose name appears on the title page of this work, has obtained, for the research described in this work, either:

(a) Human research ethics approval from the Simon Fraser University Office of Research Ethics,

or

(b) Advance approval of the animal care protocol from the University Animal Care Committee of Simon Fraser University;

or has conducted the research

(c) as a co-investigator, in a research project approved in advance,

or

(d) as a member of a course approved in advance for minimal risk human research, by the Office of Research Ethics.

A copy of the approval letter has been filed at the Theses Office of the University Library at the time of submission of this thesis or project.

The original application for approval and letter of approval are filed with the relevant offices. Inquiries may be directed to those authorities.

Simon Fraser University Library
Burnaby, BC, Canada

ABSTRACT

Embodied interaction system design faces increasing challenges to design for groups within social contexts; therefore, group user modeling within an embodied interaction environment becomes a critical area warranting exploration. Most existing group models are derived using quantitative-based approaches, which pose limitations when being applied to groups and embodied interaction. An alternative approach based on compositions and personality classifiers may address these issues, however it remains underexplored. This study discusses the rationale for choosing Keirsey Temperament as a composition classifier, presents design and implementation of our research in which we incorporated both qualitative and quantitative methods with an ecological approach. Results suggest excluding Keirsey Temperament as an effective basis for constructing composition-based group models in an embodied interaction system yet shed light on a future research focus of investigating psychological types for a wider range of dimensions as a possible basis for constructing composition-based group models in embodied interaction system design.

Keywords: group model; embodied interaction; personality type

Subject Terms: HCI; user model; interaction design; artificial intelligence

For Wei, Aaron and Robin-to-come

ACKNOWLEDGEMENTS

The project would not have been completed without the support of many people. I would like to give special thanks to my graduate program committee. Their tremendous guidance and support made this study possible. I especially appreciate the assistance from my senior supervisor Prof. Ron Wakkary, who has provided me with perceptive insights to develop research ideas in this study and great encouragement to overcome all the obstacles. I also want to express my gratitude to my supervisor Prof. Marek Hatala, who has offered enormous supervision and valuable advice during my process of accomplishing this study.

I am also thankful to Dr. Dragan Gasevic for being my external examiner. Thanks to Prof. Jim Budd for his guidance in the early years of my graduate study and Ms. Penny Simpson for her kind assistance on thesis formatting. Many thanks to Wei Wang and Mingming Zhou for their help on statistics.

My greatest appreciation to my family: I am especially thankful to my husband and soul mate Wei - thank you for always being there for me; your unconditional love and support has made this possible. I want to thank my beloved son Aaron, and Robin (who is still on its way to this world) – you are my inspiration. I love you guys!

I want to thank my parents for their care and support over the years – thank you Mom and Dad for all you have done for me.

Many thanks to my friend Cheryl Qian for her encouragement and helpful advice.

TABLE OF CONTENTS

LIST OF FIGURES

LIST OF TABLES

GLOSSARY

KTS Keirsey Temperament Sorter

CSCW Computer Supported Collaborative Work

CSCL Computer Supported Collaborative Learning

HCI Human Computer Interaction

MBTI Myers-Briggs Type Indicator

1: INTRODUCTION

1.1 Value of Research

Embodied interaction is an emerging approach to interacting with computers. Embodied interaction uses the real world as a *medium* instead of a *metaphor*, and it is by nature tangible and social. Embodied interaction systems "are manifest in our environment and are incorporated in our everyday activities" (Dourish, 2001). Embodied interactions can be characterized as full-body, haptic, and physical; and users' experience in embodied interaction is based on both physical and social reality (Hornecker & Buur, 2006).

The evolution of tangible and embodied interaction has shifted the emphasis from creating virtual systems to creating social and embodied artefacts to address the imperative needs in design for both groups and social context. Tangible and embodied interaction systems have advantages over virtual systems in that they tend to integrate better with social and physical contexts. Tangible and embodied interaction systems resemble real world experience by adopting real world characteristics to the extent of becoming part of our everyday social and physical environment. By utilizing our experiences with everyday physical objects and environment, these systems invite fluid interaction models that are less discrete and more dynamic. Tangible and embodied interaction systems enable an immediate level of social interaction among the users and between the users and their environment (Jiang et al., 2008).

Researchers of tangible user interfaces and embodied interaction systems have progressed greatly toward achieving novel interactions that leverage real world familiarities and explore embodiment. While recent studies have shifted attention in tangible computing to collaborative group environments (Binder et al., 2004; Grudin, 2002), less progress has been made in the areas of social interplay and responsiveness within an embodied context.

Our motivation in this research is to explore an adaptive group modelling approach to address the requirements of group and embodied interaction. The value of our research lies in its methodological assessment and empirical findings in exploring such an approach in order to tackle the following two major issues in the area of embodied interaction that are still in need of ways to address: 1) groups and 2) social context:

Design for groups

In life and work, for a large part of time, people form dynamic and static groups from time to time to communicate, share, and so on: socializing is one of the fundamental human needs. In Maslow's hierarchy (1943), being part of groups is the third basic type of human need, following physiological needs and safety. Groups have been inseparable from every aspect of people's lives. Most direct inter-personal interactions happen in groups.

Interaction designers nowadays are facing more challenges to design for group of users as the advance of networking technology are connecting people to each other more than ever. Personal computers have become interpersonal computers (Craiger, 1997). When technology is used to facilitate such

communication, it is essential for designers to attend the needs of the group as well as individuals. Furthermore, the complexity of modern tasks often indicates the need for teamwork instead of individual effort. Today, an increasing number of tasks require use of teams and collaboration between team members. Most tasks need collaboration between people at the different stages of the workflow and many tasks often require people to work closely together at the same time. In many situations, to optimize group performance, it is crucial for the system to be able to support all the team members as a whole.

A scenario planned for ordinary people living in 2010 with an interactive system by European Communities (Ducatel et al., 2001) envisions people surrounded by intuitive interfaces embedded in various everyday objects and provides a vision of future interactive systems in which support for human interactions, especially toward community enhancement, is a critical characteristic for the success of such systems.

Therefore, as groups are becoming prevalent in many design situations including tangible and embodied interaction, design for only individual users is often inadequate; designs that only focus on individual users' needs and neglect the group may yield unsatisfactory, sometimes even unusable outcomes. It is imperative to take group factors into consideration in order to better facilitate group interactions.

Design for social context

Designing for groups is not simply accommodating the capacity of multiple users; it requires a deep-level understanding on how a group of users interact

within certain context – at both individual level and group level – and the perspective to view the group as a whole.

Zack and McKenny's study revealed that different groups operating within different social contexts adopted the communication technology differently but in a fashion that accommodated and reinforced their existing social structure despite that they were functionally structured in the same way, performed the same task and used same communication technologies (Zach & McKenny, 1995). In their study, *social context* includes "the culture, distribution of power, and the social norms, habits, practices, expectations and preferences held by a group regarding its present and past interaction" (Zack & McKenny 1995) which has great impact on the group interaction patterns.

This finding emphasized that when designing a system that facilitates group interactions, it is crucially important to understand group differences not only structurally but also culturally, psychologically and so on.

1.2 Limitation of Previous Research

To tackle the above questions, group user modelling within an embodied interaction environment becomes a critical area in need of exploration. Understanding group interaction patterns and identifying their contributing factors are becoming increasingly important.

Most existing approaches to group modelling are quantitative-based, relying on user preferences as the basis for constructing group models. Using

such approaches, group models are represented by quantitative data computed from individual user models using a certain algorithm.

Although quantitative-based approaches allow accurate representation for the group model, they increase complexity to the degree it can become unmanageable and unrepeatable in a group situation. In addition, as an individual user's rating scheme often varies, the reliance on user preferences raises the issue of inconsistency, which can affect the accuracy of the group model. Opportunities to collect user preference data may not be available in an embodied interaction situation in which actions are often multi-modal (Jiang et al., 2008).

When constructing group model using user preferences, the choice of strategy has a significant impact on the group model constructed. Which strategies to use remains an external choice by the researcher. A bad choice of strategies may jeopardize the effectiveness of the group model. Additionally, strategies can be multiple and emergent in a given context, especially within embodied interaction. Furthermore, a quantitative group model cannot reflect individual identity as they are compromised in the process of group model computation.

There also remains some challenges for researchers in terms of predictability in group modelling. The quantitative-based approaches allow us to construct models for existing groups. However, they are limited when we need to configure groups to achieve preselected group models to take advantage of the group models. An alternative approach to group modelling based on group

compositions and behavioural classifiers may address the above limitations. However, unlike quantitative-based approaches, this alternative approach to group modelling remains underexplored.

1.3 Purpose of this Research

The aim of this research is to explore the use of personality classifiers as a basis for composition-based group modelling in an embodied interaction system.

Our interests in personality classifiers originated from the use of Bartle types (Bartle, 1996) in online gaming environments. We intended to extend the use of Bartle's concept to a physical group gameplay environment, especially our own prototype socio-ec(h)o. However, research literature supporting the validity of Bartle types is limited yet research into the roles of personality types in respect to group collaboration and problem solving is extensive and promising in our context (Jiang et al., 2008).

Personality classifiers provide a good heuristic for defining group types while maintaining identities of individuals within groups. Hence, we believed they would provide effective and efficient ways to identify group types. Most classifier methods characterize a limited set of types that makes finite possible group compositions, an important consideration for managing complexity of group interaction. Classifier systems have reliable and easy-to-administer instruments for assessing and sorting of personality types that we believed was promising in relation to developing a quick and less intrusive technique for defining types prior

to interaction thus supporting the ready-use and social interplay qualities of embodied and tangible interaction.

We reviewed commonly used personality classifiers and chose Keirsey Temperament Sorter (Keirsey, 1998) over the others due to its advantages over the other classifiers. We also chose an ecological or natural experiment approach to investigating user model strategies in a real context.

The main objective of this study is to investigate the use of Keirsey Temperament as a basis for a compositional approach to group modelling in an as ecologically valid way as possible within our embodied interaction system, named socio-ec(h)o, to provide a validity check in a real context. It aimed to provide research evidence of the effectiveness of using Keirsey Temperament for constructing composition-based group models in the context of physical gameplay in an embodied interaction system.

1.4 Structure of Thesis

This thesis consists of six chapters. Chapter 1 contains an introduction of the value of this research, limitations of previous research, and purpose of this research.

Chapter 2 reviews user modelling in respect to tangible and embodied interaction systems, group user modelling approaches, strategies and limitations. Also included is a literature review on Bartle player types, Keirsey Temperament Sorter, Myers-Briggs Type Indicator and the Big Five.

Chapter 3 introduces our research design and describes the prototype of socio-ec(h)o, an embodied interaction system.

Chapter 4 describes the research methodologies used in this study. This chapter also discusses the qualitative and quantitative nature of this study and presents the data gathering procedure.

Chapter 5 describes the three stages of data analysis procedure, presents the results from the data analysis in each stage, and discusses the results.

Chapter 6 interprets the results of the data in this study, discusses the implications and limitations of the present study, and provides suggestions for future research.

Chapter 7 summarizes the present study and proposes future work that needs to be pursued.

2: LITERATURE REVIEW

2.1 User Modelling In Embodied Interaction Systems

User modeling refers to the construction of models of users' psychological activities and behaviours to more accurately predict users' knowledge, preferences, plans and so on, which can be exploited to improve the interaction to make the system more user-centred (Kobsa & Wahlster, 1989).

User modeling was first studied about three decades ago in the field of natural-language dialog systems (Hayes & Rosner, 1976; Cohen & Perrault, 1979; Rich, 1979) and spread into other domains that are concerned with design of computer systems to be used by different users, such as information filtering and content recommendation (Hirashima et al., 1997; Linton & Schaefer, 2000) information retrieval (Newell, 1997) computer games (Albrecht et al., 1998), computer-aided education (Corbett et al., 2000; Kay, 2001).

In recent years, user modelling has attracted increasing attention from the field of human-computer interaction (HCI) because of its potential benefit to enhance the adaptivity of the system (Benyon & Murray, 1993) and to improve the collaborative nature of human-computer systems (Fischer, 2001).

In tangible and embodied interaction system design, social compatibility is a significant challenge that needs to be addressed. Being capable of dynamically adjusting to the parameters of its participants is a pivotal requirement for the system (Bohn et al., 2004). In recent research, user modelling in embodied

interaction systems is being increasingly explored. Hatala and Wakkary (2004) developed ec(h)o, which provides an "augmented reality interface" for a natural history and science museum, using knowledge-based user modelling approach. PEACH (Stock et al., 2007), a "multifaceted system that accompanies the visitor and augments her overall museum experience" also used knowledge-based approach along with stereotype-based approach. Vildjiounaite and Kallio (2007) used a combination of explicit, stereotype-based and knowledge-based user modelling approaches in Smart Home – an embodied interaction environment that recognized contexts of its inhabitants and responded accordingly.

2.2 Group Modelling

2.2.1 Definition

Group modeling combines user models of the individuals to synthesize a group model. Despite the strong interest in user models, group user models, especially in respect to tangible and embodied interaction systems, have been little explored and as a consequence, literature on group modelling is limited.

The concept of group user modelling emerged from research where numerically represented individual user preferences were combined to represent the preferences of a group of users as a whole (Kaplan et al., 1993; Mathé & Chen, 1996).

2.2.2 Research Domains

Domains in which group user modelling has been investigated include but are not limited to: adaptive information retrieval (Kaplan et al., 1993; Mathé &

Chen, 1996), group decision making (Molin, 1997; Jameson et al., 2003; Kudenko et al., 2003), content recommendation for groups (O'Conner et al., 2001; Masthoff, 2002; Masthoff, 2004), group judgment analysis (Jacobsen, 2004), CSCW (Computer Supported Collaborative Work) (McCarthy & Anagnost, 1998; Jameson et al., 2003) and CSCL (Computer Supported Collaborative Learning) (Suebnukarn & Haddawy, 2006).

2.2.3 Strategies

Most existing studies on group modelling focus on the question of how to satisfy the entire group based on the preferences of each individual in the group and the strategies they used closely relate to the issue of social choice. Many strategies (McCarthy & Anagnost, 1998; Hogg & Jennings, 2000; O'Conner et al., 2001; Ardissono et al., 2002; Masthoff, 2002) that were adopted were derived from and devised based on the Social Choice Theory to serve the purpose of deciding the sequence of items for a group. Commonly used strategies include:

- *Utilitarian Strategy – Additive:* This strategy derives a sequence based upon the sum of all the user ratings; the larger the sum the earlier the alternative appears in the sequence. User rating refers to users' assessment of statements or items, which is usually represented by a response to a Likert scale question or a number within a certain range. Hogg and Jennings (1999) used this strategy in their research on a multi-agent system to facilitate group decision-making. Sorting the average of individual ratings in a descending order can derive the same sequence.

Therefore, it was called "Average Strategy" in Masthoff's (2002) study on adaptive television on user groups.

- *Least Misery Strategy:* This strategy decides the sequence of a list of items based on the minimum of the individual ratings on each item, the higher the minimum of the ratings, the earlier the item appears in the sequence. This strategy was adopted in PolyLens (O'Conner et al., 2001), a movie recommender system for groups of users based on the assumption that the group is happy to the degree of its least happy member.
- *Average without Misery Strategy:* This strategy is the combination of the above two strategies – a sequence is derived based upon the average of the individual ratings in a descending order, but ratings that are below a certain threshold are excluded from the sequence. MusicFX (McCarthy & Anagnost, 1998) used this strategy to select music to suit the interests of a group of people in a fitness center.
- *Most Respected Person Strategy:* This strategy decides a sequence for the group based only on the ratings of the most respected member in the group. This strategy was used in INTRIGUE (Ardissono et al., 2002), a system that provides information about tourist attractions and services to a group of people travelling together.

2.2.4 Limitations

Most existing group models are quantitative, represented by data computed from individual user models using a certain algorithm. Many of these strategies pose limitations in respect to applying them to groups and embodied interaction, such as:

- *Unlimited variations:* Quantitative-based approaches employ user preferences as the basis for constructing group models. Numerically represented individual user preferences can have virtually unlimited variations, and therefore so does the group model. Although this allows accurate representation for the group model, it on the other hand increases complexity to the point of being unmanageable and unrepeatable in a group situation.
- *Consistency:* The reliance on user preferences raises the issue of inconsistency. Different users often have different rating schemes: some are cautious about being extreme and prefer to use the middle of the scale; while others avoid being neutral and tend to use only the ends of the scale. As a result, a 7 out 10 can mean "excellent" to one user but "mediocre" to another. Depending on the individual rating tendency variation in the group, this issue of inconsistency can impact on the accuracy of the group model.
- *Assessibility:* Opportunities to gather user preferences ratings may not be available in an embodied interaction situation in which actions are multi-

modal involving physical movements, gestures, and non-verbal communication with other users and the system.

- *Choice of strategies:* The choice of strategy has significant impact on the group model constructed. However, which strategies to use is an external choice by the researcher. In some cases, the best choice of strategy is obvious, whereas very often, it can be hard to decide the optimum strategies and therefore there is the potential risk of ineffective group model caused by bad choice of strategies. Additionally, strategies can be multiple and emergent in a given context, especially within embodied interaction.
- *Loss of individual identities*: Using quantitative data based group modelling approaches, individual identities cannot be reflected in a group model derived, as they are compromised in the process of group model computation.
- *Predictability:* A quantitative-based approach is straightforward and easy to use in constructing models for existing groups. But what if we want to do it the other way around? That is, how can we configure groups to achieve preselected group models in order to take advantage of the group models?

2.3 Common Practices of Defining and Measuring Personality

An alternative approach based on group compositions and behavioural classifiers may address the above issues. In this study, we chose to investigate

personality classifiers as the basis for a composition-based group modelling approach.

We started our interests in personality classifiers with Bartle types (Bartle, 1996) which has been widely used in online gaming environments. We intended to extend the use of Bartle's concept to a physical group gameplay environment, especially our own prototype socio-ec(h)o; however, little supporting literature is available on its validity. We then reviewed some other common personality classifiers, namely Myers-Briggs Type Indicator (MBTI), Keirsey Temperament Sorter (KTS) and the Big Five.

We narrowed down our candidates for the personality classifier to MBTI and KTS based on our general guideline in choosing a classifier:

• An ideal classifier should be measured using a standard procedure to define its value for each individual instead of being subjectively determined by the individuals themselves to avoid the issue of inconsistency;

• An ideal classifier should yield results that are different enough to identify individuality yet have limited variations to make experiment practically operationalizable;

• Individuals of each type as assessed with this classifier should be abundantly available, thus we can easily design and conduct experiments;

• A valid instrument for identifying types should be available and potentially extensible for implementation in an embodied interaction situation.

We later chose KTS as the classifier in our study and the rationale is discussed in Chapter 4. The following is the review of these four personality classifiers.

2.3.1 Bartle Types

Bartle conducted research on player personality types in collaborative play in massively-multiplayer online games, specifically Multi-User Dungeons (MUDs), and developed a model known as Bartle's Player Types (Bartle, 1996). The underlying concept of this model is two axes representing gaming preferences in two dimensions: the x-axis going from an emphasis on players to an emphasis on the environment and the y-axis going from interacting with other players to acting alone. This model creates taxonomy of players based on these two axes, with each type of player in one of the quadrants. The four types of players are: *Achievers,* who are interested in acting in the environment; *Explorers,* who are interested in interacting with the environment; *Socializers,* who are interested in interacting with other players in the environment; and *Killers,* who are interested in acting on other players.

Bartle's concept has been adapted into an online test (GuildCafe, 2008) generally referred to as the Bartle Test, which helps online game players to identify their game-playing preferences and classifies them into these four categories. However, several limitations may cause some issues when applying the Bartle's model.

First, Bartle's model is purely theoretical and itself provides no empirical ways to assess player's types. Second, the Bartle Test has been challenged for its dichotomous and forced-choice nature (Yee, 2004) and was deemed to "merely perpetuates the assumptions of Bartle's Types rather than validating them" (Yee, 2005). More importantly, there is very limited literature supporting the validity of Bartle's model.

Yee (2005) argued that, due to the lack of evidence from empirical data that support its validity, it would be essentially hard to adopt Bartle's model pragmatically. He conducted empirical research that involved 3200 participants for the purposes of testing the validity of Bartle's types and correcting for inherent problems with a purely theoretical model. The data indicated that the components that hypothetically distribute bi-modally in Bartle's model were in fact normally distributed. The results from Yee's study challenged Bartle's concept which inherently puts people in one of the four boxes, and indicated that players are multi-faceted and should be defined by a cluster of preferences instead of a single preference.

2.3.2 Myers-Briggs Type Indicator

Katharine Cook Briggs and Isabel Briggs Myers developed Myers-Briggs Type Indicator (MBTI) (Myers, 1962) based on the framework of Jung's (Jung, 1921/1971) theory of psychological types. The purpose of MBTI is to help individuals to identify their prominent personal traits.

MBTI contains four dichotomies to represent four aspects of people's psychological properties: (1) sensing*[S]/intuition[N]* to represent the way people prefer to use for perceiving information; (2) *thinking[T]/feeling[F]* to represent the way people prefer to use for judgment or decision making; (3) *extraversion[E]/introversion[I]* to represent people's preferred orientation between inner world and outer world; (4) *judging[J]/perceiving[P]* to represent people's orientation in the outer world. The current North American English version of MBTI Step I questionnaire includes 93 questions and each question has two possible answers, one of which has to be chosen. The two possible answers to each question are not necessarily opposite in meaning but different enough to reflect participants' two possible preferences.

Much research has been conducted to examine the validity of MBTI and there is sufficient evidence for its reliability (Carlyn, 1977; Carbon, 1985; Myers & McCaulley, 1985) and validity (Carlyn, 1977; Carlson, 1985). MBTI has been widely used in various fields, including career choosing, team building, learning, relationship, personal growth, etc. Today, more than two million people worldwide take the Indicator each year (MyersBriggs.org, 2007).

2.3.3 Keirsey Temperament Sorter (KTS)

Keirsey Temperament Sorter (KTS) is a personality self-assessment questionnaire designed by Keirsey (1978) and KTS II is its current version that is available in both paper and online format (Keirsey, 1998). Keirsey states that there are four temperaments, namely Artisan, Guardian, Idealist, and Rational and KTS is designed to help people indentify their dominant temperament.

KTS is closely related to MBTI in that it combined the existing Myers-Briggs system groupings SP, SJ, NF and NT and mapped them respectively to four temperaments Artisan, Guardian Idealist, and Rational. However, KTS is essentially different from MBTI in two aspects: First, KTS assesses personality type using a hierarchical system that contains four temperaments at the top level and four subtypes within each temperament; whereas, MBTI assesses personality type in a linear way: it uses four sets of dichotomies and yields results of sixteen possible psychological types represented by the combination of four letters, one from each dichotomies. Second, Keirsey argues one's character trait should be assessed based on the observations of his or her behaviour (Keirsey, 1998) and as a result, KTS is behaviourally focused; while MBTI concentrate on cognitive and emotional qualities of types and is thus psychologically focused.

Much research has been conducted in order to find out how closely KTS and MBTI are related to each other, and many researchers have found correlations between these two measures of personality type. Tucker and Gillespie (Tucker & Gillespie, 1993) found correlations between MBTI and KTS in both paper and online format (values ranging from .68 to .86). Kelly and Jugovic (Kelly & Jugovic, 2001) also found strong correlations between KTS II and MBTI. Therefore, the above research has provided evidence to support that KTS and MBTI were measuring the same constructs, specifically the four sets of psychological types, namely E/I, S/N, T/F and J/P.

2.3.4 The Big Five

Based on the decades of pioneering work on taxonomy of personality traits conducted by previous researchers (Allport & Odbert, 1936; Cattell 1943, 1945a, 1945b; Norman 1963, 1967), Goldberg (1981) initially labelled five factors as the major dimensions of personality traits, which eventually became known as the "Big Five". The Big Five are five dimensions of personality traits. These five dimensions are Extraversion, Agreeableness, Conscientiousness, Neuroticism, Openness to Experience. Each of these dimensions consists of a number of more specific traits. For example, Conscientiousness includes such related qualities as perseverance, dependability, and controllability.

Although various studies (Goldberg, 1990; Saucier & Goldberg, 1996b; Saucier, 1997) demonstrated consistency and replicability of the Big Five, several limitations question its generalizability and validity. The Big Five has been generally investigated using the lexical approach that is based on English language (John et al., 1988; Saucier & Goldberg, 1996a), and as a consequence, this may raise the issue of generalizability across languages and cultures. Furthermore, the Big Five was discovered and investigated based on purely empirical studies and it not grounded by any theoretical foundation.

There are various measuring instruments for the Big Five developed by different researchers. John et al. (1991) developed Big Five Inventory that is comprised of 44 short phrases. Costa and McCrea developed a questionnaire known as NEO Personality Inventory (Costa & McCrea, 1985) to measure three of the five dimensions of the Big Five, namely Neuroticism, Extraversion, and

Openness to Experience, and later they further extended the NEO Personality Inventory to cover all five dimensions (Costa & McCrea, 1992). The revised questionnaire initially contains 240 items and was abbreviated to 60 items. Goldberg compiled a list that includes 100 trait descriptive adjectives (TDA) (Goldberg, 1992) to be measured using scales.

John and Srivastava (1999) conducted research to investigate the validity of the above three instruments and found that, all these instruments showed satisfactory coefficient alphas which indicated they may be reliable. However, it is also found that, across instruments, Agreeableness and Openness tended to be less reliable than the other three dimensions.

3: SOCIO-EC(H)O

3.1 Research Design

An indirect approach (Mastoff, 2004) may allow easier investigation of several strategies or classifiers in one study. However, such an approach reduces elements in isolation with stricter control, and as a result, this poses limitations when the subjects need to be studied together within their context.

Suchman (1987) stated that activities are situated in their environment, in other words, they take place "in particular, concrete circumstances." We echo her point of view on the importance of context. In our study, we proposed an ecological approach to explore composition-based group user models. By "ecological", we referred to a complete system that is developed in order to test the whole experience including the relationships between subjects and their environment. Using an ecological approach, the entire continuously ongoing experience is being studied while it naturally emerges and unfolds.

We believe that a natural experiment environment is critical in respect to researching tangible and embodied interaction systems (Jiang et al., 2008). Our approach is influenced by the philosophical notions of a natural inquiry (Lincoln & Guba, 1985) in which "realities are multiple, constructed, and holistic", "all entities are in a state of mutual simultaneous shaping", and "only time- and context-bound working hypotheses are possible" in terms of generalization.

We considered the following factors important and designed our research accordingly:

1. The group models should be studied in a group collaboration situation where the groups are to perform real instead of fictional activities. Thus, we chose to study groups in a collaborative group gameplay to solve physical puzzles.
2. The group models should be studied in an environment that is as natural as possible. Bearing this in mind, we configured the sound and display in an ecological way in our experiment environment.
3. We aimed at providing a particular set of heuristics that help identify optimum group modelling approaches in systems like ours. Therefore, we chose to thoroughly study one particular personality classifier and investigate the group models in depth from several perspectives, namely behaviour, perception and performance with this particular personality classifier.

Our study protocol was based on playing the socio-ec(h)o game that builds on simple, physical, and structured embodied interaction models. We constructed a dynamic responsive ambient intelligent space as the game environment. We chose Keirsey Temperament as the classifier to construct and investigate group models. The rationale for choosing Keirsey Temperament is discussed in detail in Chapter 4.

3.2 socio-ec(h)o Game Structure and Scenario

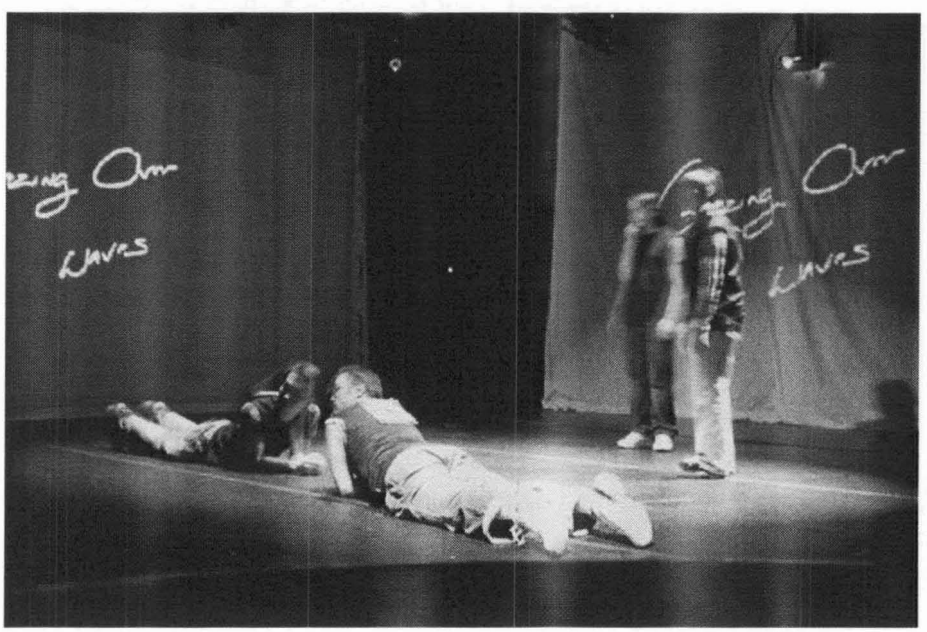
One single session of socio-ec(h)o game involves one team of four players. The four players are challenged to collaboratively solve physical puzzles and accomplish six levels.

The short scenario of the socio-ec(h)o environment below is excerpted from a previous paper about this study (Wakkary et al., 2005):

> *Madison, Corey, Elias and Trevor have just completed the first level of socio-ec(h)o. They discovered that each of them had to be low to the ground, still, practically on all fours. Once they had done that, the space became bathed in warm yellow light and filled with a wellspring sound of resonating cymbals. Minutes earlier, the space was very dim – almost pitch black until their eyes adjusted. A quiet soundscape of "electronic crickets" enveloped them. They discussed and tried out many possibilities to solving the word puzzle: "Opposites: Lo and behold." At Corey's urging, the four grouped together on the edge of the space and systematically sent a player at a time to the opposite side in order to gauge any change in the environment. Nothing changed. Madison, without communicating to anyone realized the obvious clue of "Lo" or "low". She lowered herself to a crouching position. The space immediately glowed red and became brighter. The audio changed into a rising chorus of cymbals – not loud but progressively more pronounced. Corey and Trevor stopped talking and looked around at the changing space. Madison, after a pause began to say "Get down! Get down!" Elias stooped down*

immediately and the space became even brighter. Corey and Trevor dropped down in unison and the space soon became bathed in a warm yellow light like daylight. The audio reverberated in the space. A loud cheer of recognition came from the group, "Aaaaahhh! We got it!"

Figure 3-1 Participants playing level 5 of socio-ec(h)o



The group needs to finish a six-level game by achieving specific body states and goals. The body states are the body movements and positions that players are required to achieve to accomplish a certain level. The goals the players are aiming to achieve are the changes responded by the environment in accordance with the body states. The goals are implicit and the player must discover them in each level during the game. After a beginning quality of light and audio started in each level, the players progress toward achieving the correct body states and correspondingly, the environment gradually changes toward the

goal state (see Figure 3-1, above). For instance, in the above scenario, the environment shifted toward the goal of creating day as Madison changed her body to a lower position, and then responded to the movements of the group as the other three players followed Madison. To complete the levels, the players need to gradually acquire skills to identify the goal state in order to manipulate the environment. The levels were designed to shift from low to high with progressive difficulty to enable the players to acquire more complex skills.

3.3 System Setup

3.3.1 Physical Setup

The experiment took place in a closed space approximately 30'x30' and 15' high. The space had black floor and walls, minimal décor, and was surrounded by thick black curtains and black walls, which help to make the space more immersive (See Figure 3-2, below).

The following are some of the main features of the experiment space:

- An eight-channel surrounding sound system and a lighting system were installed to play ambient sound and light. They were controlled by custom software programmed in Max/MSP.
- On three sidewalls of the space, three pieces of approximately 12'x12' white cloth were hung flat - each on one side of the space - as projection screens for ambient visual display.
- Two stationary video cameras were set up in two corners to record the entire process of the experiment for further analysis: Camera A

in the furthest upper corner in the space to capture the full scene and all three projection screens and Camera B in a near corner that is close to the area where the participants move around to capture activity details and two screens. A hand-held camcorder Camera C was used occasionally for close-ups and shots from complementary angles.

- Twelve infrared motion cameras were installed at the side and on the ceiling pointing the central area of the space. The motion cameras were spread out in a way that allows maximum coverage in the space (See Figure 3-3, below).

Figure 3-2 Physical setup of the experiment space

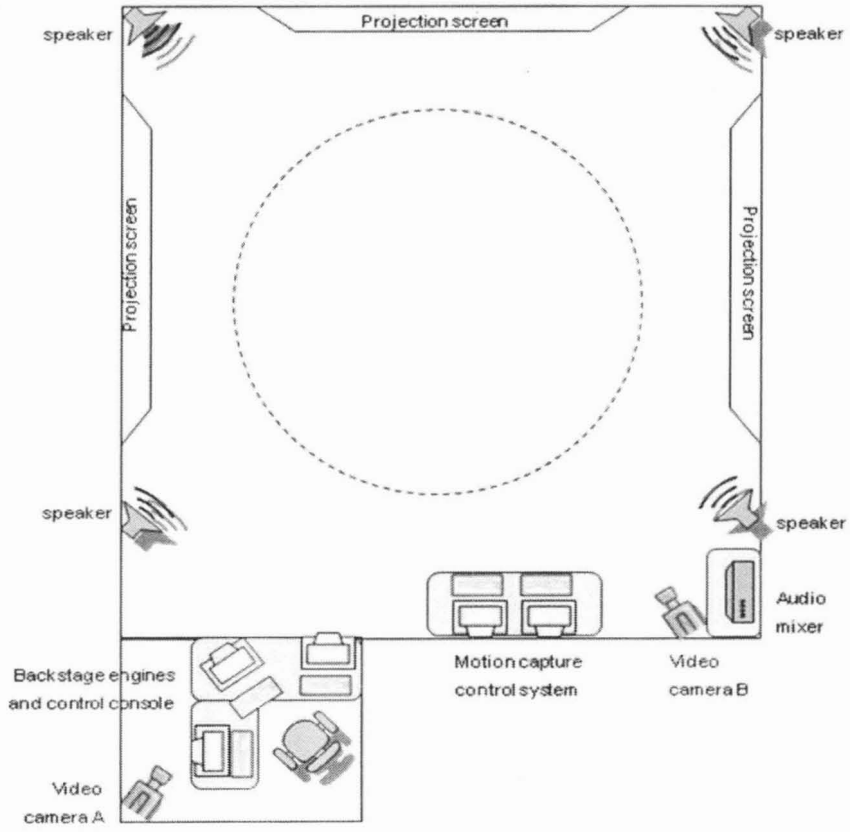
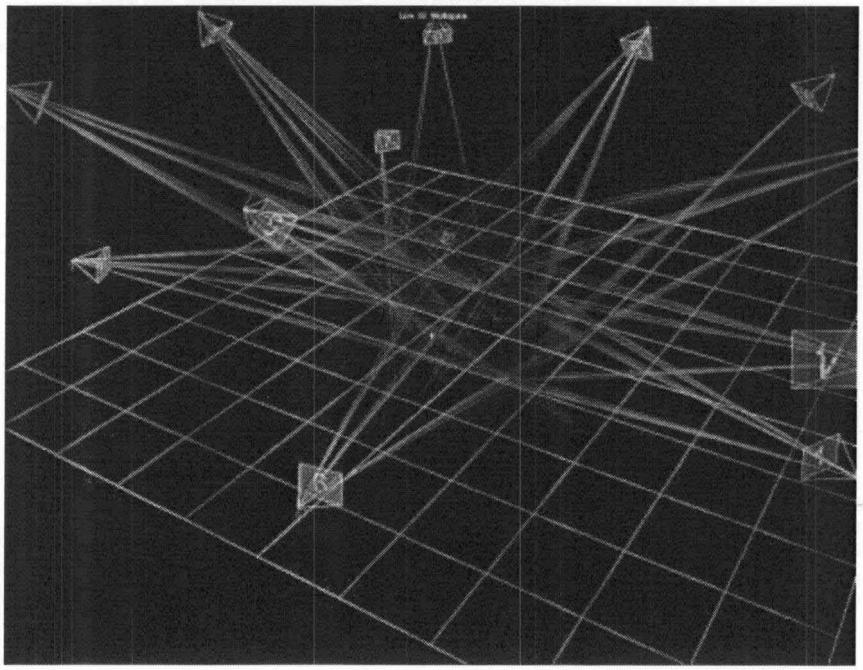


Figure 3-3 Motion Capture Camera Setup



Wakkary et al., 2005

A sound mixer and the motion capture control system were pushed off along the side where there was no projection. The backend systems and control console were setup at a far corner of the space. Computer displays were either turned off or placed backward from the participants to minimize distractions.

3.3.2 Technical Prototype

The technical system for socio-ec(h)o is comprised of three key components: a sensing system, reasoning engine and display engine, which are integrated by lightweight communication protocol that is transferred over the User Datagram Protocol (UDP) communication channel. The three key components are briefly described below. Readers who would like further technical details can refer to a previous paper related to this project (Wakkary et al., 2005).

Sensing System

A Vicon MX motion capture system that comprises twelve infrared cameras is the core component of the sensing system and it is controlled by custom software developed in Max/MSP. Each participant is distinguished by unique configuration of reflective markers worn on the back. The system detects each player's velocity, three-dimensional position, orientation, proximity and movement and transmits data to the reasoning engine for further analysis.

Reasoning Engine

The reasoning engine infers high-level group behaviour based on sensing data it received from the sensing system, control the narrative flow of the socio-ec(h)o experience, and feeds its output to the display engine.

Display Engine

The display engine includes audio and lighting components. The audio component provides a sound ecology for each game level as well as a gradient response to the participants. The lighting components create theatrical lighting effects and feedback to participants.

4: RESEARCH METHODOLOGY

4.1 Keirsey Temperament as Classifier

Based on the previous review in Chapter 2, both MBTI and KTS may be reliable ways to assign an individual a personality type and are good candidates for the classifier in our research. We chose KTS over MBTI for the following reasons:

- *Possibilities of less effort and more accurate answers from participants:* Myers' focus is on the cognitive and emotional qualities of types whereas Keirsey's focus is weighted toward behaviour. Keirsey argues that the best way to determine one's character trait is to carefully observe what he or she actually does (Keirsey, 1998). As a result, the questions in KTS seems more concrete and simpler when compared to those in MBTI. We believe this difference makes the KTS questionnaire more explicit and therefore it is more likely that the KTS questionnaire instrument has the potential to be modified to easily bootstrap the identification of individual and group types.
- *Easier operationalization by researchers:* MBTI assesses personality type using four sets of dichotomies, thus the results yield sixteen possible psychological types represented by the combination of four letters, one from each dichotomies, in a linear way. KTS assesses personality type using a hierarchical system that contains four temperaments at the top

level and four subtypes within each temperament. It is easier to form groups of certain personality combinations using four temperaments rather than sixteen psychological types. The hierarchical system of KTS can greatly simplify research design and implementation.

- *Easier and less expensive to implement:* MBTI is a commercial product and can require trained personnel for assessment. The current fee for the service of MBTI assessment is approximately USD $150 per participant. KTS II are now available both in paper format and online. The online KTS II assessment can be taken at no cost.

4.2 Measures

Team Performance

Team performance was measured by two factors: 1) level of completion, i.e., the highest level of puzzle accomplished and 2) completion times of each level.

Team Behaviours

We measured team behaviours based on two factors: *cohesion* and *goal focus* (Wakkary et al., 2008). *Cohesion* is a measure of team dynamics and can be described as the extent to which all members appear to be coordinating together and acting as a team – whether that is working on game solutions, playing, thinking or talking to each other. Seashore (1954) developed a four-item cohesiveness index and suggested that cohesion varies in work groups. In recent group studies, cohesion was a common construct to examine the affinity an

individual feels toward a group (Cammann et al., 1983). Mudrack (1989) addressed it as a high commitment to the solidarity of the group.

Goal focus can be described as the extent to which players appear to be or are attempting to "play the game" the way they understand it. Malone and Lepper (1987) identified four major factors of a learning game and proposed that games are intrinsic motivators enabling learners to focus on the mental images of things and situations that are not actually presented. Huizinga (1964) stated that play is a place for games in which players stepped out of the real life and intensely focus on the activities that go beyond the confines of the everyday world.

Perceptions

Subjects' perceptions on the system and their gameplay experiences were measured by a questionnaire, which solicits answers in five-point Likert scale, multiple choices and descriptive text format. The Likert scale answers range from very negative = 1 point to very positive = 5 points. For example, for a question like "How well did the system respond when you were having problems solving a puzzle (circle one)", the answers range from 1= "not well" to 5 = "very well". The answers to the descriptive questions are very open-ended. An example question can be "please describe in what respect you feel positively or negatively toward the system". Due to the time-consuming nature of content analysis, only Likert scale answers were analyzed in this stage of the research.

4.3 Data Collecting Methods

4.3.1 socio-ec(h)o Questionnaire

We developed a questionnaire of twenty-five questions including Likert scale and semi-structured questions as instrument to measure players' perceptions (Wakkary, 2006). The questionnaire included questions on the following four perspectives based on Activity Checklist (Kaptelinin et al., 1999):

1. *Means and ends:* how technology facilitated or constrained goals and how it caused or reconciled conflict goals.

   Sample question: How effective was the audio feedback in helping you solve the puzzles?

2. *Social and physical aspects of the environment:* how requirements, tools, resources, and social rules of the environment integrate with technology.

   Sample question: How well does the virtual system (lights, audio, video) integrate with the physical and architectural space?

3. *Learning and cognition:* the function of internal and external components of activity and their mutual transformation.

   Sample question: How well did the system provide useful cues in helping you evaluate how close you were to solving a puzzle?

4. *Development:* how the preceding activity transformed holistically.

   Sample question: Did your attitude toward the system become more or less positive as the evaluation progressed?

4.3.2 Video Recording

Video recording provided direct evidence of what the participants are exactly doing in the research experiment. Data could be captured by the video recording devices and those data could be transferred into standard statistical packages. Comparing to other data collection methods like note making in the field, video recording has crucial advantages because the audio visual inscriptions of events could be reviewed by multiple researchers after the event. The raw event data could be rated or coded to provide valuable findings for the research. In addition, the inscriptions are able to be stored and event records may be reanalyzed, examined for inter-coder reliability and reviewed by peer researchers later on.

4.3.3 System Log

Fine-grained data capturing body state of each participant characterized by six parameters inferred from the motion tracking systems were logged with the frequency of 200ms. The system response was logged with the same frequency, which allowed us to study the effect of group activity with a high accuracy and correlate it with video and audio data collection. Based on the system log, we could additionally reconstruct the audio display (Jiang et al., 2008).

4.4 Data Analysis Methods

4.4.1 Log and Questionnaire Analysis

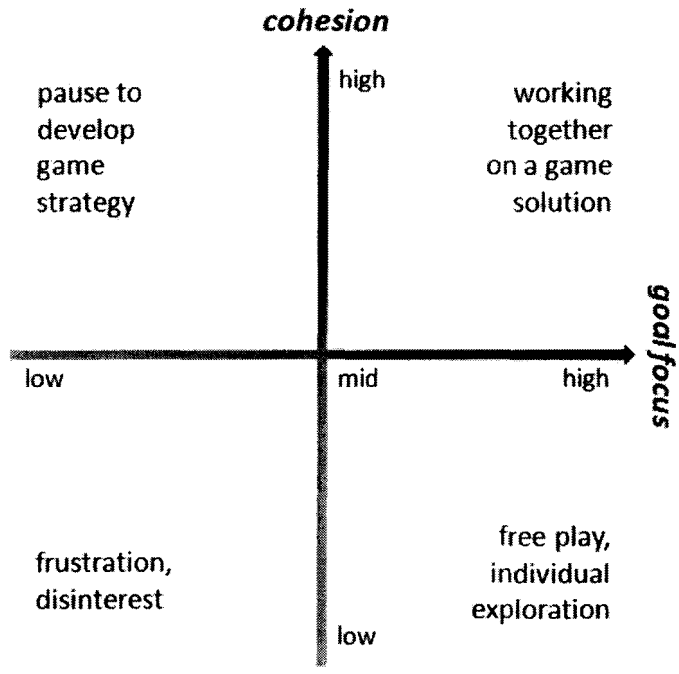
Log and questionnaire data was exported and sorted into spreadsheets. The sorted data was later transferred into a standard statistical package (SPSS) for analysis.

4.4.2 Video Coding

The video recording was analyzed using our coding scheme based on the two measures of team behaviour: *cohesion* and *goal focus* (Wakkary et al., 2008). These two measures were each analyzed for their different degrees, times of occurrences and their durations and sequences. Two researchers coded the videos together in consensus.

Each of the two measures was assessed for three degrees and their combination in a two-dimensional matrix show the degree of descriptive capacity (see figure 2, below).

Figure 4-1 Matrix showing the descriptive capacity of the two factors cohesion and goal focus



Wakkary et al., 2008

Cohesion and *goal focus* were independent of each other. Cohesion only indicates whether the players are acting in unison as a team, it does not necessarily reflect their focus on the game. Likewise, *goal focus* is not related to whether players are working as a team or not; it only refers to the degree of how active the players are playing the game. When playing the game, the player can be either moving or staying still, depending on their understanding of the requirements to solve the physical puzzle.

The three degrees of *cohesion* and *goal focus* were defined as following (Wakkary et al., 2008):

• *Low cohesion*, players are not together as a group or they are temporarily fragmented. They are not communicating or are individually exploring.

• *Medium cohesion*, players are in the process of becoming a group or are regrouping. Players are negotiating roles and establishing leadership or consensus.

• *High cohesion*, players constitute an established team. They make several agreements and are coordinated in their movements or are communicating with each other about strategy and solving the puzzles.

• *Low goal focus*, players are not involved in playing the game. They are resting, or are distracted, or engaged in activities not related to the game.

• *Medium goal focus*, players are in the process playing the game. They are experimenting with different actions, and communicating with each other about or reflecting on the effects of their actions.

• *High goal focus*, players are actively and consciously playing the game and attempting to solve the puzzle at hand. This is reflected in concerted efforts and good communication related to their performance in the game. Many ideas are shared on actions for solving the puzzle.

4.4.3 Triangulation

Triangulation of qualitative and quantitative, also known as mixed use of multiple methods, is a powerful data analysis strategy often used in the social and behavioral research. In this study, both qualitative and quantitative analysis

methods were adopted to align with the research questions and the nature of the data.

Qualitative approaches can be utilized to gain more understanding of specific situations and to provide a more contextualized perspective to empirical studies. For example, qualitative approaches helped us answer questions such as what was exactly happening in a team and to the environment at a particular moment in a game session, how long did the situation last, what were the causes and consequences of the situation, and so on. Qualitative approaches are powerful in identifying factors, defining their attributes and degrees; however, they are limited to make claims to generality about subtle differences in observed data captured in the video recordings.

In light of this, quantitative approaches are often used to discover what are likely to be the patterns or tendency of a population as a whole and which occurred merely by chance in well-designed experiment that carefully controls and manipulates relevant variables (Chi, 1997). In general, quantitative approaches afford opportunities to answer the how much and how often questions. For example, quantitative approaches helped us answer questions such as, how much of a certain level in a game session, the team acted in high cohesion and how often did they shift their goal focus to high level from medium or low level? These answers can be subjected to powerful statistical examinations by utilizing quantitative analysis methods.

In the present study, qualitative and quantitative approaches were each applied to analyze different dimensions in different phases. Quantitative data

analyses were conducted to assess team performance and investigate the relationship between team configuration and performance and the relationship between player temperament type and perceptions on the game. Qualitative data analyses were used to identify events and their durations and degrees of team behaviour for each team at each level; quantitative data analyses were then conducted to examine a team's behaviour pattern across a level and the whole session and investigate the relationship between team configuration and behaviour.

4.5 Data Gathering Procedure

These data were gathered by the socio-ec(h)o research team, led by Prof. Ron Wakkary, who granted permission to analyze and present it for the purpose of this thesis.

4.5.1 Participants

Participants were recruited (84 were recruited in total and 56 of them – 36 males and 20 females – were used for the actual study) from our university community. The majority of the participants were undergraduate students, but there were a number of graduate students and staff as well. The administrators noted participants' races only to the extent that they seemed to be random and varied.

Figure 4-2 Distribution of age ranges of subjects

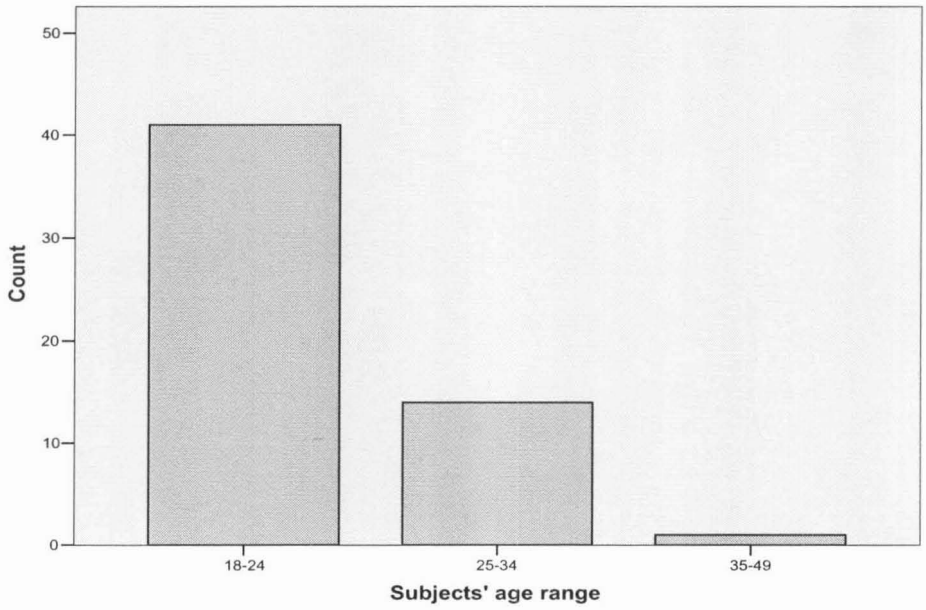
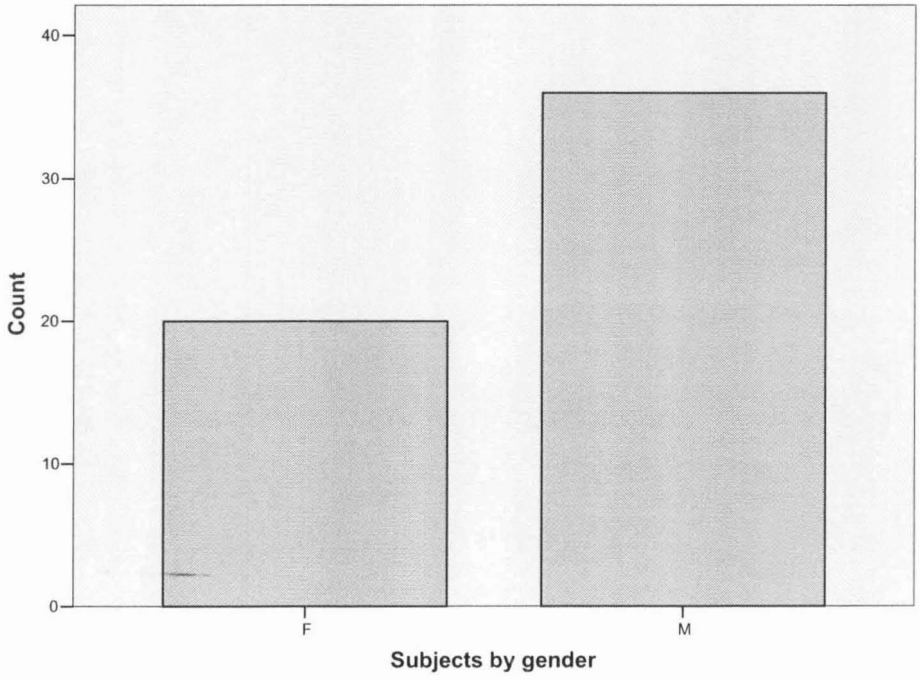


Figure 4-3 Subjects by gender



4.5.2 Procedure

4.5.2.1 Recruiting Subjects

In order to motivate as many people as possible to participate in the research, a number of incentives were used, including monetary (chances to win an iPod Nano upon registering and taking personality test and a $10 iTunes or Famous Players gift certificate for each individual to participate in the gameplay), social (opportunities to know other fellow students in game play and learn more about the state-of-the-art research project in campus), and psychological (challenges in solving the puzzles). Subjects were recruited using various ways, including class visits, posters and flyers, and campus news mailing list and advertising emails.

Class visits

Handouts and signup sheets were designed for recruiting subjects. On top of the handout, it noticeably indicated "Play a game (...plus answer a few questions) and receive iTunes gift certificate or Famous Players gift certificate + Win an iPod Nano!" The handout indicated that whoever wants to participate needs to sign up to take a 15 minute survey, and that eligible participants will be chosen to take part in the evaluation of socio-ec(h)o (and receive an iTunes or Famous Players gift certificate of approximately 10$ value. To encourage as many people as possible to sign up for the survey, it was emphasized in the handout that everyone who completes the survey will be eligible to win the iPod Nano in a draw. The handout also briefly outlined the processes to participate and listed the drop-in times and locations to sign up and take the survey.

Eight classes were chosen, because these classes were game/design/media related and students enrolled in these classes were more likely to be interested in participating the evaluation sessions. At the beginning of each class, a faculty member and a research assistant from the research team visited the class for 5 to 10 minutes, which was pre-arranged with the corresponding instructor. During a visit, we introduced our research project known as socio-ec(h)o and the research team to the class, distributed and walked students through the handouts, went over the instructions, drop-in times, and signup sheets, and answered any questions. See Appendix 2 for a copy of the handout and signup sheet.

Posters and flyers

The poster and flyer have exactly the same content and layout, except that they are in different sizes. They cover all the contents included in the handout made for the class visits, and additionally, introduced the socio-ec(h)o research project and the research teams. Posters were posted on bulletin boards around SFU Surrey campus; flyers were put in the front desk for the undergraduate students to pick up and in the mailboxes area to attract the attention from staff, faculty and graduate students. See Appendix 2 for a copy of the poster/flyer.

Campus news mailing list and advertising emails

The same information outlined in the posters/flyers was also distributed by SFU Surrey campus news mailing list known as "SFU Surrey Connect", which targets mainly the undergraduate students in SFU Surrey campus and also staff and faculty. To attract more graduate students, an advertising e-mail was also

sent to the SFU School of Interactive Arts and Technology graduate students mailing list.

Word of mouth

The information was also spread out to fellow graduate students, researchers and faculty in a personal way. Many showed interest and some of them registered to participate.

4.5.2.2 Administering KTS II Tests

Potential participants were invited to attend a Keirsey Temperament Sorter II test to assess their personality types. Six sessions were administered in after-class hours in previously booked classrooms on campus. Each session lasted from two to two and half hours. During each session, the interested people can drop in at anytime, fill in a form of contact information and take a paper-based KTS II test (see Appendix 3). The length of time each individual took to finish the test varied, but a typical time length is about 15 minutes. Everyone who took the test was asked to select one or more evaluation workshop sessions that were convenient for him or her, and this information would later be used for scheduling.

4.5.2.3 Selecting and Grouping Subjects

Each participant took a KTS II test and was assessed for his or her temperament based on the answers to the questionnaire. We labelled each test participant's temperament using the abbreviation of temperament (A for Artisan, G for Guardian, I for Idealist and R for Rational) or a combination of abbreviations

of the temperaments if there is a tie (i.e., the scores of two or more temperaments are equal), such as A/G or A/I/R.

We used the following guidelines in selecting and grouping the participants for the study:

- To make the results more reliable, mixed temperaments should be eliminated as many as possible, especially three temperaments, such as A/G/R or G/I/R;
- Based on the protocol and design of the gameplay we required four-person teams;
- To maximize the sample size of teams of a same personality type composition, we decided to use the least possible number of combinations of personality types (i.e., two);
- As the purpose of the evaluation workshops is to study the differences between teams of different personality type compositions, ideally the numbers of teams of each personality type composition should be equal.

Based on the above guidelines, we found that grouping participants into GGAA and GIIR teams allows us to get the most number of people out of the pool to participate in the our studies than any other ways of grouping, and it yields close to equal numbers of either type of groups.

4.5.2.4 Conducting Evaluation Workshops

Scheduling

First, we negotiated with another research group who also needed to use the same experiment space and found out the days and hours that the space were available to the evaluation workshops for this research. Based on this information we worked out 16 tentative timeslots spread out among 8 days.

We asked all the KTS test participants to provide us with their availability and preferences to all 16 timeslots and then compiled them along with their temperament types in a spreadsheet in a way that information is visualized. Based on such information, we configured either a GGAA or a GIIR group for each timeslot.

We then sent e-mails to those tentatively scheduled for the upcoming evaluation workshop session to confirm their participation. It turned out that scheduling was an iterative and laborious process, as inevitably people's schedule and availability often change. We often had to reschedule either people or the workshop time or recruit more participants to cope with the changes, and eventually we managed to run 14 evaluation workshop sessions, with four GGAA groups and five GIIR groups in one protocol and five GIIR groups in another protocol. This work only reports on one of the protocols and groups of thirty-six (36) participants. We also have chosen to focus our analysis on data from two critical levels of the game (i.e., level 3 and 4) rather than all the levels particularly for in-depth investigation on team behaviour.

On-site running

A single session involved one team. After introducing the research team to everyone, participants were given a warm-up session to get a grip of the idea of using metaphors to present puzzles and solving a puzzle through physical actions. The warm-up was a modification of the child's game of "hot-cold", which served as a good introduction to the type of responses the system would provide. We played a range of sonic cues and rewards in order to familiarize the participants to important aspects of our sound display.

The participants were also provided with an overview of the game and the system. Each participant was then equipped with a microphone for user testing and a vest that has reflective markers at the back for data acquisition by the motion capture system. After all the devices were tested and properly configured, the participants were then asked to participate the gameplay session that was conducted in two parts – in the first part, participants complete the first four levels within 60 minutes; and in the second part, participants attempt to complete the remaining two levels within 15 minutes. See Appendix 4 for socio-ec(h)o evaluation protocol.

Questionnaire and debriefing

After each evaluation workshop session was finished, participants were asked to fill out a questionnaire on site. See Appendix 5 for socio-echo questionnaire.

Rewarding subjects

Participants were thanked for participating. Those who chose a movie certificate as their preference was each given a Cineplex movie voucher with a value of $10 CAD on site right after each session. Participants who chose iTune certificates as their preference each received an e-mail afterwards that contains a code that can be used to redeem $10 CAD toward online purchases with iTune.

4.5.3 Data Collected

The following types of data were collected in this study:

• Age range, gender, Keirsey temperament type and psychological types (measured by the paper-based Keirsey Temperament Sorter II questionnaire);

• Observation: Researchers were present and observing each session;

• Video data: All sessions were recorded in their entirety by three stationary video cameras and selectively with one hand-held video camera;

• Verbal and audio data: Each participant was individually recorded using wireless microphones in addition to general audio recorded by the video cameras;

• System log: Body state of each participant characterized by six parameters inferred from the motion tracking systems were captured and logged, together with the system response with the frequency of 200ms;

• Participant perceptions: Immediately after each session, participants were asked to complete a questionnaire that contains eleven Likert questions, two multiple-choice questions and twelve semi-structured questions to reflect their perceptions on their game experience.

5: DATA ANALYSIS AND RESULTS

In our study, both performance and behaviour were examined at a group level. These two dimensions were naturally able to be measured at group level. However, perception is by nature individually differentiated. In this early phase of our research, perceptions were investigated at individual level to first explore if a connection can be established between individuals' temperament types and their perceptions in order to provide empirical understanding of their relationship from a new vantage point and establish a foundation for our future research on group perception model. Due to the time-consuming nature of content analysis, in this initial phase of our study, I did not scrutinize all the descriptive data collected by the questionnaire and only analyzed the answers to the Likert scale questions.

5.1 Stage 1

In this stage of analysis, I examined whether groups composed differently based on Keirsey Temperament have identifiably different characteristics in performance and behaviour. I also investigated whether players of different Keirsey Temperaments have different perceptions on their game experiences.

The following are the relationships I inspected and their respective results:

1) Team temperament composition and team performance:

Table 5-1 below shows the descriptive statistics of all variables examined in the correlations. Team temperament composition was correlated to both *level of*

completion and *duration* (see Table 5-2, below); no significant correlation was found.

2) Team temperament composition and team behaviour:

Table 5-3 below shows the descriptive statistics of all variables examined in the correlations. Team temperament composition was correlated to each level of *cohesion* (see Table 5-4, below) and *goal focus* (see Table 5-5, below) in level 3 and 4; no significant correlation was found.

3) Individual temperament type and perceptions on the game experience:

Players' Keirsey Temperament types were correlated to their answers to the Likert scale questions (see Table 5-7, 5-8 and 5-9, below). The results showed significant correlations between individual temperament type and answers to Question 2.1, 2.2 and 2.3, and significant correlations between team temperament composition and answers to Question 2.3 (see Table 5-8, below). Table 5-6 below shows the descriptive statistics of the answers to the Likert scale questions from the questionnaire.

Table 5-1 Descriptive statistics of highest level of completion and durations

	Mean	Std. Deviation	Number of Teams
Highest level accomplished	5.333	0.707	9
	Mean (minute)	**Std. Deviation (minute)**	**Number of Teams**
Level 1 duration	0.761	0.926	9
Level 2 duration	1.800	1.318	9
Level 3 duration	6.222	3.772	9
Level 4 duration	22.800	23.161	9
Level 5 duration	11.644	6.726	8
Level 6 duration	2.125	1.601	4

Table 5-2 Correlations between team composition and highest level of completion and durations

	1	2	3	4	5	6	7
1. Team temperament composition type	-	-	-	-	-	-	-
2. Highest level accomplished	0.224	-	-	-	-	-	-
3. Level 1 duration	-0.447	-0.016	-	-	-	-	-
4. Level 2 duration	0.432	0.194	-0.233	-	-	-	-
5. Level 3 duration	0.054	0.508	-0.012	0.318	-	-	-
6. Level 4 duration	-0.412	0.092	0.612	-0.254	0.357	-	-
7. Level 5 duration	-0.154	-0.621	0.150	-0.151	0.122	0.206	-
8. Level 6 duration	0.451	###	-0.080	0.710	0.772	-0.425	0.388

Cannot be computed because at least one of the variables is constant.

Table 5-3 Descriptive statistics of different levels of cohesion and goal focus in level 3 and 4

	Mean (%)	Std. Deviation (%)
High degree of cohesion in Level 3	74.333	36.246
Medium degree of cohesion in Level 3	15.444	11.069
Low degree of cohesion in Level 3	10.056	6.167
High degree of goal focus in Level 3	55.444	14.258
Medium degree of goal focus in Level 3	17.111	10.925
Low degree of goal focus in Level 3	27.333	9.513
High degree of cohesion in Level 4	79.556	21.284
Medium degree of cohesion in Level 4	11.222	13.599
Low degree of cohesion in Level 4	9.778	11.7130
High degree of goal focus in Level 4	67.000	16.897
Medium degree of goal focus in Level 4	13.222	11.065
Low degree of goal focus in Level 4	20.333	9.138

Table 5-4 Correlations between team composition and different levels of cohesion in level 3 and 4

	1	2	3	4	5	6
1. Team temperament composition type	-	-	-	-	-	-
2. High degree of cohesion in Level 3	-0.524	-	-	-	-	-
3. Medium degree of cohesion in Level 3	-0.509	-0.843**	-	-	-	-
4. Low degree of cohesion in Level 3	0.338	-0.398	-0.828**	-	-	-
5. High degree of cohesion in Level 4	0.265	-0.220	-0.419	0.455	-	-
6. Medium degree of cohesion in Level 4	-0.539	-0.631	0.758*	-0.611	-0.837**	-
7. Low degree of cohesion in Level 4	0.038	-0.198	-0.012	-0.172	-0.785*	0.358

**. $p < .01$, *. $p < .05$

Table 5-5 Correlations between team composition and different levels of goal focus in level 3 and 4

	1	2	3	4	5	6
1. Team temperament composition type	-	-	-	-	-	-
2. High degree of goal focus in Level 3	0.403	-	-	-	-	-
3. Medium degree of goal focus in Level 3	-0.444	-0.743*	-	-	-	-
4. Low degree of goal focus in Level 3	-0.108	-0.651	-0.023	-	-	-
5. High degree of goal focus in Level 4	0.281	0.199	-0.687*	0.484	-	-
6. Medium degree of goal focus in Level 4	-0.576	-0.511	0.846**	-0.190	-0.882**	-
7. Low degree of goal focus in Level 4	0.095	0.154	0.425	-0.726	-0.816**	0.486

**. $p < .01$, *. $p < .05$

Table 5-6 Descriptive statistics of Likert scale questions from the socio-ec(h)o questionnaire

	Mean	Std. Deviation	Number of Players
Individual temperament type	2.219	1.128	32
Team temperament composition	1.444	.504	36
Section1.1 How confident do you feel about using the system after this evaluation?	3.653	.970	36
Section1.2 Did the system require a large effort to learn?	3.086	1.173	35
Section1.3 Was the system more or less helpful in solving the puzzles as the evaluation progressed?	3.500	1.056	36
Section1.4 Did your attitude toward the system become more or less positive as the evaluation progressed?	3.471	1.187	34
Section2.1 How well did the system respond to all your actions from the start of solving the puzzle to solving it?	3.417	.967	36
Section2.2 How well did the system provide useful cues in helping you evaluate how close you were to solving a puzzle?	3.611	.934	36
Section2.3 How well did the system respond when you were having problems solving a puzzle?	3.056	1.351	36
Section2.4 How well did the system support coordinating information and actions within the team?	3.545	.971	33
Section3.1 Do the lights and audio work together to help you in the game?	3.900	.999	35
Section3.2 How effective was the audio feedback in helping you solve the puzzles?	4.097	.909	36
Section3.5 How well does the virtual system (lights, audio, video) integrate with the physical and architectural space?	3.788	.857	33
Section3.6 Is the environment created by the system conducive for a game?	3.686	.963	35

**. $p < .01$, *. $p < .05$

Table 5-7 Correlations between temperament and answers to Likert scale questions in section 1 of the socio-ec(h)o questionnaire

	1	2	3	4	5
1. Individual temperament type	-	-	-	-	-
2. Team temperament composition type	-.564**	-	-	-	-
3. Section1.1	-.180	.120	-	-	-
4. Section1.2	-.256	.180	.277	-	-
5. Section1.3	-.338	.107	.244	.081	-
6. Section1.4	-.358	.098	.220	.268	.582**

**. $p < .01$, *. $p < .05$

Table 5-8 Correlations between temperament and answers to Likert scale questions in section 2 of the socio-ec(h)o questionnaire

	1	2	3	4	5
1. Individual temperament type	-	-	-	-	-
2. Team temperament composition type	-0.56**	-	-	-	-
3. Section2.1	-0.382*	0.313	-	-	-
4. Section2.2	-0.429*	0.317	0.501**	-	-
5. Section2.3	-0.542**	0.382*	0.310	0.402*	-
6. Section2.4	0.164	-0.265	0.403*	0.132	0.101

**. $p < .01$, *. $p < .05$

Table 5-9 Correlations between temperament and answers to Likert scale questions in section 3 of the socio-ec(h)o questionnaire

	1	2	3	4	5
1. Individual temperament type	-	-	-	-	-
2. Team temperament composition type	-0.564**	-	-	-	-
3. Section3.1	-0.262	.000	-	-	-
4. Section3.2	-0.111	-0.128	0.530**	-	-
5. Section3.5	0.045	-0.075	0.356*	0.429*	-
6. Section3.6	0.169	0.043	0.033	0.178	0.625**

**. $p < .01$, *. $p < .05$

5.2 Stage 2

In our preliminary data analysis in stage 1, no significant correlation was found. This led me to the second stage analysis in which I examined our implementation process in order to narrow down the possible contributing issues.

First, I conducted a reliability analysis on the data collected from the Keirsey Temperament Sorter questionnaire. 67 out of 69 cases were valid for analysis (see Table 5-10, below). We examined coefficient alpha, which is often used as a measure of reliability of a psychometric instrument. The coefficient alpha is .764 (see Table 5-11, below), which indicated that the data collected from the KTS II questionnaire is internally consistent and therefore KTS II is reliable in this study.

Next, I examined the treatment of subjects that have mixed temperaments. In our study, subjects with three mixed temperaments were excluded and we tried to minimize those with two mixed temperaments. Three teams each contained one team member with two mixed temperaments. In these cases, one

of the two mixed temperaments was used to identify the individual to suit the preselected group models. These three teams were analyzed in stage 1 as part of the data set. In this stage, I excluded these three teams from the data set and re-conducted the same group level statistical analyses as in stage 1. Likewise, no significant correlations were found between team temperament composition and team performance or factors of *cohesion* and *goal focus*.

In stage 1, there was no significant correlation between players' temperament type and their perceptions. In this stage, an internal-consistency estimates of reliability (coefficient alphas) test was conducted to ensure that the items comprising each subscale are reliable. 29 out of 56 cases were valid for analysis (see Table 5-12, below). The coefficient alpha is .815 (see Table 5-13, below), which indicated that the data collected from the socio-ec(h)o questionnaire is internally consistent and therefore it is reliable in measuring the perceptions.

Table 5-10 Case processing summary

		Number	%
Cases	Valid	67	97.1
	Excluded[a]	2	2.9
	Total	69	100.0

a. Listwise deletion based on all variables in the procedure.

Table 5-11 Reliability statistics

Cronbach's Alpha	Number of Items
.764	70

Table 5-12 Case Processing Summary

		Number	%
Cases	Valid	29	51.8
	Excluded[a]	27	48.2
	Total	56	100.0

a. Listwise deletion based on all variables in the procedure.

Table 5-13 Reliability statistics

Cronbach's Alpha	Number of Items
.815	12

5.3 Stage 3

In this stage, I analyzed the data in finer granularity. A Keirsey temperament is comprised of four aspects of psychological preferences, also known as psychological types, each being represented as a dichotomy. These four dichotomies are: *extrovert* (E) and *introvert* (I), *sensing* (S) and *intuitive* (N), *thinking* (T) and *feeling* (F), and *judging* (J) and *perceiving* (P), and the degree of each psychological type is indicated by the corresponding score in the KTS questionnaire. A temperament is decided by the combination of the dominating preference in each of the four aspects. Using this method, information on the degree of each preference is discarded. For example, an individual that scores J:

19 / P: 1 and an individual that scores J: 11 / P: 9 are deemed equally as *judging* types.

To reflect how different the scores of each extreme of a preference are, and therefore achieve more accurate results of the impact of each psychological preference at the team level, for each psychological preference, the scores of the two ends of the dichotomy from each team member were added up respectively and the sum was used to represent the team's psychological preference.

For example, team A's four members are as following:

Player 1 – E: 6, I: 4, S: 16, N: 4, T: 9, F: 11, J: 12, P: 8,

Player 2 – E: 9, I: 1, S: 8, N: 12, T: 6, F: 14, J: 7, P: 13

Player 3 – E: 4, I: 6, S: 7, N: 13, T: 8, F: 12, J: 1, P: 19

Player 4 – E: 5, I: 5, S: 8, N: 12, T: 18, F: 2, J: 12, P: 8

In our previous analysis, team A members were identified as:

Player 1: ESFJ -> Guardian

Player 2: ENFP -> Idealist

Player 3: INFP -> Idealist

Player 4: (E/I)NTJ -> Rational

and the group model for team A is GIIR.

Here in this stage, the group model for team A is:

E: 24, I: 16, S: 39, N: 41, T: 41, F: 39, J: 32, P: 48

I correlated team performance to each of the four aspects of psychological preferences (see Table 5-14, 5-15, 5-16 and 5-17, below), and found significant correlations between team performance and group E/I preference in level 1 (see Table 5-14, below) and between team performance and group T/F preferences in level 3 (see Table 5-16, below).

Table 5-14 Correlations between team value of E/I (Extrovert/Introvert) and highest level of completion and durations

	1	2	3	4	5	6	7	8
1. Team value of E (Extrovert)	-	-	-	-	-	-	-	-
2. Team value of I (Introvert)	-1.000**	-	-	-	-	-	-	-
3. Highest level accomplished	.000	.000	-	-	-	-	-	-
4. Level 1 duration	.726*	-.726*	-.016	-	-	-	-	-
5. Level 2 duration	.198	-.198	.194	-.233	-	-	-	-
6. Level 3 duration	.276	-.276	.508	-.012	.318	-	-	-
7. Level 4 duration	.362	-.362	.092	.612	-.254	.357	-	-
8. Level 5 duration	.404	-.404	-.621	.150	-.151	.122	.206	-
9. Level 6 duration	.823	-.823	.[a]	-.080	.710	.772	-.425	.388

**. $p < .01$, *. $p < .05$

a. Cannot be computed because at least one of the variables is constant.

Table 5-15 Correlations between team value of S/N (Sensing/Intuitive) and highest level of completion and durations

	1	2	3	4	5	6	7	8
1. Team value of S (Sensing)	-	-	-	-	-	-	-	-
2. Team value of N (Intuitive)	-.999**	-	-	-	-	-	-	-
3. Highest level accomplished	.357	-.326	-	-	-	-	-	-
4. Level 1 duration	-.352	.365	-.016	-	-	-	-	-
5. Level 2 duration	.242	-.219	.194	-.233	-	-	-	-
6. Level 3 duration	.155	-.126	.508	-.012	.318	-	-	-
7. Level 4 duration	-.403	.409	.092	.612	-.254	.357	-	-
8. Level 5 duration	-.083	.084	-.621	.150	-.151	.122	.206	-
9. Level 6 duration	.403	-.367	.[a]	-.080	.710	.772	-.425	.388

**. $p < .01$, *. $p < .05$

a. Cannot be computed because at least one of the variables is constant.

Table 5-16 Correlations between team value of T/F (Thinking/Feeling) and highest level of completion and durations

	1	2	3	4	5	6	7	8
1. Team value of T (Thinking)	-	-	-	-	-	-	-	-
2. Team value of F (Feeling)	-.903**	-	-	-	-	-	-	-
3. Highest level accomplished	-.253	.482	-	-	-	-	-	-
4. Level 1 duration	.272	-.430	-.016	-	-	-	-	-
5. Level 2 duration	-.289	.445	.194	-.233	-	-	-	-
6. Level 3 duration	-.794*	.752*	.508	-.012	.318	-	-	-
7. Level 4 duration	.054	-.248	.092	.612	-.254	.357	-	-
8. Level 5 duration	-.212	-.090	-.621	.150	-.151	.122	.206	-
9. Level 6 duration	-.779	.916	.[a]	-.080	.710	.772	-.425	.388

**. $p < .01$, *. $p < .05$

a. Cannot be computed because at least one of the variables is constant.

Table 5-17 Correlations between team value of J/P (Judging/Perceiving) and highest level of completion and durations

	1	2	3	4	5	6	7	8
1. Team value of J (Judging)	-	-	-	-	-	-	-	-
2. Team value of P (Perceiving)	-.847**	-	-	-	-	-	-	-
3. Highest level accomplished	-.273	.396	-	-	-	-	-	-
4. Level 1 duration	-.552	.641	-.016	-	-	-	-	-
5. Level 2 duration	-.087	-.114	.194	-.233	-	-	-	-
6. Level 3 duration	.075	-.184	.508	-.012	.318	-	-	-
7. Level 4 duration	-.361	.142	.092	.612	-.254	.357	-	-
8. Level 5 duration	-.008	-.169	-.621	.150	-.151	.122	.206	-
9. Level 6 duration	.306	-.306	.[a]	-.080	.710	.772	-.425	.388

**. $p < .01$, *. $p < .05$

a. Cannot be computed because at least one of the variables is constant.

6: DISCUSSION

We compared two team types defined by the composition of Keirsey Temperaments and did not find statistically significant differences in respect to team performance and behaviours at group level and perceptions at individual level. We examined potential issues that might have occurred during the implementation of our study and excluded the causality between these issues and the non-existence of correlations. These findings suggest that we may rule out Keirsey Temperament as an effective basis for constructing composition-based group models to differentiate groups in the case of systems similar to our collaborative, physical, and embodied interaction system, specifically in the dimensions we investigated, namely performance, behaviour and perceptions. However, finer-granularity analysis in the third stage yielded some interesting results that revealed significant correlations between certain team psychological preference value and performance in certain levels of the game.

6.1 Empirical Findings and Implications of the Current Study

The correlations found in the first stage analysis can be interpreted as following:

- Individuals from different team composition have different perceptions on the system in terms of the effectiveness of its responses when they were having problems solving a puzzle.

- Individuals of different Keirsey temperament have different perceptions on the responsiveness of the system and the effectiveness of the system responses.

As we discussed in the beginning of the previous chapter, perceptions in the context of our present study is by nature individual and we investigated individual perceptions to verify the connection between individuals' temperament types and their perceptions in order to establish a foundation for our future research on group perception model. Although the correlation between team composition and individual perceptions did not support any claim on the effectiveness of using Keirsey Temperament as the basis for constructing composition-based group models, it did reveal that *the group settings that an individual belongs to have an impact on the individual's perceptions of the environment*. Hence, this finding echoed the emphasis on the importance of social context as we argued at the beginning of this thesis. The correlations between Keirsey temperament type and perceptions on the other hand validated the utility of Keirsey temperament in differentiating individual perceptions.

The correlations found in the third stage analysis can be interpreted as following:

- The higher the team value of *extrovert*, the longer the level 1 duration, thus the lower the team performance. Vice versa, the higher the team value of *introvert*, the shorter the level 1 duration, thus the higher the team performance.

- The higher the team value of *thinking*, the shorter the level 3 duration, thus the higher the team performance. Vice versa, the higher the team value of *feeling*, the longer the level 3 duration, thus the lower the team performance.

Although our findings are inconclusive and may merit further investigations of the causality between Keirsey sub-types (i.e., psychological types) and select levels in respect to team performances, it does suggest potential viability for psychological types to serve as an effective basis for constructing composition-based group models to differentiate team performance.

6.2 Limitations of the Current Study

There are a number of limitations of the current study that need to be considered:

1. First, the data analyzed in this study was not exhaustive. In our research, we collected abundant data in various formats, such as system logs, questionnaire data, and video and audio recordings. Currently, in the preliminary phase in evaluating our system, we have only analyzed part of the data we collected. We conducted quantitative analysis on the system log data and the answers to the Likert scale questions in the questionnaire and mixed analysis on the video recording specifically based on cohesion and goal focus. Some of the data we analyzed can be subject to analyses using alternative methods or finer-granular and more in-depth analyses. For example, the video

recording provides us with rich and contextualized data that can be analyzed in numerous dimensions and granularities. A substantial amount of data is still awaiting analysis, for example, the data collected using the semi-structured questions in the questionnaire and the transcript of conversations among the players.

2. Due to small number of teams of each composition, the results and conclusion were restricted in generalizability. In the present study, five GIIR teams and four GGAA teams were analyzed. We used grouping strategies that allow us to get the most possible number of people out of the participant pool to form close to equal numbers of teams of two compositions. However, we felt that a larger number of teams would be more helpful in confirming and generalizing patterns.

3. In this study, we investigated composition-based group models in terms of performance, behaviour and perceptions. We acknowledge that there are other important factors that need exploration, but a full investigation of all these factors was beyond the scope of this present study. In our future research, we plan to explore more dimensions as effective factors in differentiating groups.

4. Determined by the in-depth and ecological nature of our approach, our current study is limited to thoroughly exploring one classifier, namely the Keirsey Temperament. With reducible elements that could be studied in isolation with stricter controls or using an indirect approach, we might be able to investigate several classifiers in one study.

However, we argue that, given the fluid nature of the factors involved in the experience and the current exploratory nature of the research, tangible and embodied interactions are best examined in such a natural and holistic way, and therefore we chose to validate or invalidate any claims for Keirsey Temperament as an effective basis for constructing composition-based group models (Jiang et al., 2008).

6.3 Suggestions for Future Research

Here are some suggestions for future research to extend our present study:

1. We collected abundant data in the present study. As noted in the limitations of the current study, we are currently in the preliminary phase of our research, and the data collected have not yet been exhaustively analyzed. We believe that more intense qualitative analysis on the video recordings, questionnaire data and conversation transcripts should be a priority of our future research, as this will allow us to better interpret the present findings and gain empirical understanding on some of the questions that currently remain unanswered, and help us identify and investigate more factors.
2. Our present study investigated composition-based group models in three dimensions, namely performance, behaviour and perceptions. Although our findings suggest ruling out Keirsey Temperament as an effective classifier for constructing composition-based group models in

respect to these three dimensions with the specific measures we examined, there are numerous other measures within these dimensions and a wider range of other dimensions, and it is possible that KTS could be effective in differentiating other measures and dimensions. Given the restraint of our present study that only a limited number of measures and dimensions were explored, it is only appropriate to investigate other possible measures and dimensions before we can draw generalized conclusion on whether it is effective to use KTS in constructing composition-based group models in embodied interactions.

3. As we reviewed in Chapter 2, the concept of *Psychological Type* (Jung, 1921/1971) is the common fundamental concept of Keirsey Temperament Sorter and Myers-Briggs Type Indicator. We also reviewed the correlations between KTS and MBTI as evidences supporting that they were measuring the same constructs, specifically the four sets of psychological types, namely E/I, S/N, T/F and J/P (see Section 2.3). The findings in our finer-granular analysis provided empirical evidence of the effectiveness of using psychological types in constructing composition-based group models, which suggested that MBTI can be validated for its sub-types and deserves more investigation as an effective classifier. Although, as we discussed in Chapter 4, one of the limitations of using MBTI is the issue of operationalization. Unlike the hierarchical system used in KTS, MBTI

yields sixteen types, using the combination of four psychological types, one from each of the four dichotomies (see Section 4.1). Given the preceding findings and discussion, we think the utility of MBTI as the classifier and could become a subject of future research; however, how to simplify the utility of MBTI without affecting its validity is an issue that needs to be addressed first.

7: CONCLUSION AND FUTURE WORK

Our motivation for this research was to expand the avenue of group modelling in embodied interaction system design and explore the effectiveness of using personality classifiers.

In this work, we started with a discussion on the need to design for groups and social contexts in the area of embodied interaction and our intention to explore an adaptive group modelling approach in order to address such needs.

We reviewed user modelling in embodied interaction systems and commonly used quantitative-based group modelling approach, argued that quantitative-based group modelling approach is limited when considered for tangible and embodied interaction systems, and viewed composition-based group modelling approach using personality classifiers as a potentially viable alternative that addresses the problematic issues of quantitative-based approaches to group modelling.

This work discussed our rationale for choosing Keirsey Temperament as a composition classifier, presented the design and implementation of our research in which we incorporated both qualitative and quantitative methods within a natural experiment approach. We believe our description may serve as a methodological example for studying embodied interaction systems.

At the current stage of our research, we studied groups in three dimensions, namely performance, behaviour and perceptions. We conducted

three-stage analyses. In the first stage, we examined the identifiability of group models in the above dimensions with groups configured using compositions based on Keirsey Temperament and did not find any significant correlation between group composition and the factors we investigated. In the second stage, we conducted reliability analyses on the data collected from the Keirsey Temperament Sorter questionnaire and our socio-ec(h)o questionnaire. The coefficient alphas (.764 and .815) indicated that the data collected from both questionnaires were internally consistent and therefore were reliable in this study, and hence excluded the causality between them and the non-existence of correlations. In the third stage, we conducted analyses on revised group models represented using finer-granular component of Keirsey Temperament, namely the psychological types. The results from this stage revealed relationships between group psychological preferences and performance in select levels.

In summary, results from our three-stage analysis has suggested that Keirsey temperament was a reliable classifier in identifying individual perceptions for our participants however did not serve as a good basis for constructing composition-based group models in our specific context of physical gameplay in an embodied interaction system. Nonetheless, our analysis results have shed light on a future research focus of investigating the effectiveness of using psychological types for a wider range of measures and dimensions as a possible basis for constructing composition-based group models in embodied interaction system design.

The present study was an exploratory one and we were not aware of any other published research similar to ours. Our research is still in its early stage and the current study has its limitations; before these limitations are addressed and the research is replicated and extended, the results and conclusion from the present study should be considered tentative. The limitations of the present study include: First, the data was not exhaustively analyzed as we are in the preliminary phase in our system evaluation. Second, nine teams were studied in the present study, although we used optimal strategies for the most efficient grouping, the sample size is still somewhat small. Third, our conclusion is limited to the specific dimensions and measures we examined in this study. Fourth, due to the ecological nature of our research, our findings and conclusions on personality classifiers in the present study are limited to only the Keirsey Temperament.

We proposed the following future work to pursue to expand our present study and address its limitations: First, more thorough and intense data analyses should be conducted to help identify and investigate more factors. Second, more measures and dimensions should be examined to help provide a definitive conclusion on the generalizability of the effectiveness of using Keirsey Temperament as the basis for constructing composition-based group modelling. Third, the present study on using personality classifiers in group modelling can be expanded to other classifiers. The Myers-Briggs Type Indicator especially could become a subject of future research, as the present study provided positive evidence of the validity of its sub types.

APPENDICES

Appendix 1 – socio-ec(h)o Handout

Play a game (... plus answer a few questions) and receive:

+ iTunes gift certificate or
+ Famous Players gift certificate

+

Win an iPod Nano!

if...

... you participate in the socio-ec(h)o evaluation challenge, an ambient-intelligent team game that challenges you to collaborate, and physically solve puzzles.

we need you...

... to sign up to take a 15 minute survey. If you are one of the first 60 eligible participants, you'll be chosen to take part in the evaluation of socio-ec(h)o, and receive an iTunes or Famous Players gift certificate! (Approx. 10$ value) **Everyone who completes the survey will be eligible to win the iPod Nano in a draw!**

instructions:

1. Sign up by giving us your name and email.
2. Tell us when you can take the survey (see drop-in times and places below)
3. Tell us when you'll be available to do the evaluation. If you're chosen, we will notify you via email based on the availability you suggested to us on the sign up sheet.
4. Let us know which certificate you would like; iTunes or Famous Players.

Sign up sheet is attached on the back. Please leave your sign up sheet with your instructor after class.

Here are the drop in times. (Survey takes 15 minutes total)

Friday, February 24,
drop-in hours: 3:00 - 6:20pm,
SUR 800 ☐

Monday, February 27,
drop-in hours: 4:00 - 6:20pm,
SUR 616 ☐

Tuesday, February 28,
drop-in hours: 4:00 - 6:20pm,
SUR 616 ☐

Wednesday, March 1st
drop-in 4:00 - 6:20pm,
SUR 160 ☐

Thursday, March 2nd,
drop-in hours: 3:30 - 6:20pm,
SUR 630 ☐

Check off which time you signed up for, and use this sheet as your reference.

The socio-ec(h)o evaluation challenge

sign-up sheet

The socio-ec(h)o evaluation challenge

1 Name and email

Name: _____

Email: _____

2 I'll take the 15 min survey on...

(please check off one of the drop-in times)

- ☐ **Fri, February 24,** Drop-in hours: 3:00 - 5:20pm, SUR 800
- ☐ **Mon, February 27,** Drop-in hours: 4:00 - 6:20pm, SUR 515
- ☐ **Tue, February 28,** Drop-in hours: 4:00 - 6:20pm, SUR 515
- ☐ **Wed, March 1st,** Drop-in hours: 4:00 - 6:20pm, SUR 150
- ☐ **Thur, March 2nd,** Drop-in hours: 3:30 - 5:20pm, SUR 630

3 I can do the evaluation on...

(please check as many of the time slots as possible so we can place you in one of the sessions you indicate below.)

Mon, March 6th,
- ☐ 9:00 - 12:00pm
- ☐ 1:00 - 4:00pm

Tues, March 7th,
- ☐ 9:00 - 12:00pm
- ☐ 1:00 - 4:00pm

Wed, March 8th,
- ☐ 9:00 - 12:00pm
- ☐ 1:00 - 4:00pm

Fri, March 10th,
- ☐ 9:00 - 12:00pm
- ☐ 1:00 - 4:00pm

Mon, March 13th,
- ☐ 9:00 - 12:00pm
- ☐ 1:00 - 4:00pm

Tues, March 14th,
- ☐ 9:00 - 12:00pm
- ☐ 1:00 - 4:00pm

Wed, March 15th,
- ☐ 9:00 - 12:00pm
- ☐ 1:00 - 4:00pm

Fri, March 17th,
- ☐ 9:00 - 12:00pm
- ☐ 1:00 - 4:00pm

4 I want the... *(please check one)*

- ☐ iTunes 10$ Gift certificate
- ☐ Famous Players 10$ Gift certificate

Appendix 2 – socio-ec(h)o Flyer

Play a game (... plus answer a few questions) and receive:

+ iTunes gift certificate

or

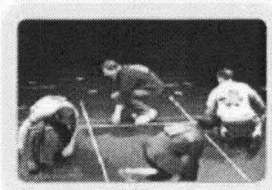
+ Famous Players gift certificate

+

Win an iPod Nano!

if...

... you participate in the socio-ec(h)o evaluation challenge, an ambient-intelligent team game that challenges you to collaborate, and physically solve puzzles.



what is socio-ec(h)o & who are we ?

Socio-ec(h)o is a research project researching group collaboration, communication and the role of group user modeling in an ambient intelligent environment. We have created an ambient intelligent game in an interactive space that is responsive to people's behaviour based on reasoning and the use of artificial intelligence.

The research team includes Ron Wakkary, Marek Hatala, Jim Bizzocchi and several SIAT graduate students.

we need you...

... to sign up to take a 15 minute personality survey. If you are one of the first 60 eligible participants, you'll be chosen to take part in the evaluation of socio-ec(h)o (maximum 3 hours), and receive an iTunes or Famous Players gift certificate! (Approx. 10$ value). **Everyone who completes the survey will be eligible to win the iPod Nano in a draw!**

interested ?

simply start with a 15 minute survey at any of the following drop-in times:

- **Friday, Feb 24**	3:00-5:20 pm	SUR800
- **Monday, Feb 27**	4:00-6:20 pm	SUR515
- **Tuesday, Feb 28**	4:00-6:20 pm	SUR515
- **Wednesday, Mar 1**	4:00-6:20 pm	SUR150
- **Thursday, Mar 2**	3:30-5:20 pm	SUR630

...and we will enter you into the draw and guide you through the rest of the processes :)

The socio-ec(h)o
evaluation challenge

Appendix 3 – The Keirsey Temperament Sorter II

Instructions:

1. Decide on answer **a** or **b** and circle it.
2. Your responses should reflect "the real you", not the way you want to be, think you should be, or are asked to be by someone else.
3. There are no right or wrong answers since about half of the population agrees with whatever answers you choose.
4. There are no right or wrong personality types.

1. When the phone rings do you:
   a. hurry to get it first
   b. hope someone else will answer

2. Are you more:
   a. observant than introspective
   b. introspective than observant

3. Is it worse to:
   a. have your head in the clouds
   b. be in a rut

4. With people are you usually more:
   a. firm than gentle
   b. gentle than firm

5. Are you more comfortable in making:
   a. critical judgments
   b. value judgments

6. Is clutter in the work place something you:
   a. take time to straighten up
   b. tolerate pretty well

7. Is it your way to:
   a. make up your mind quickly
   b. pick and choose at some length

8. Waiting in line, do you often:
   a. chat with others
   b. stick to business

9. Are you more:
   a. sensible than ideational
   b. ideational than sensible

10. Are you more interested in:
    a. what is actual
    b. what is possible

11. In making up your mind are you more likely to go by:
    a. data
    b. desires

12. In sizing up others, do you tend to be:
    a. objective and impersonal
    b. friendly and personal

13. Do you prefer contracts to be:
    a. signed, sealed, and delivered
    b. settled on a handshake

14. Are you more satisfied having:
    a. a finished product
    b. work in progress

15. At a party, do you:
    a. interact with many, even strangers
    b. interact with a few friends

16. Do you tend to be more:
    a. factual than speculative
    b. speculative than factual

17. Do you like writers who:
    a. say what they mean
    b. use metaphors and symbolism

18. Which appeals to you more:
    a. consistency of thought
    b. harmonious relationships

19. If you must disappoint someone are you usually:
    a. frank and straightforward
    b. warm and considerate

20. On the job, do you want your activities:
    a. scheduled
    b. unscheduled

21. Do you more often prefer:
    a. final, unalterable statements
    b. tentative, preliminary statements

22. Does interacting with strangers:
    a. energize you
    b. tax your reserves

23. Facts:
    a. speak for themselves
    b. illustrate principles

24. Do you find visionaries and theorists:
    a. somewhat annoying
    b. rather fascinating

25. In a heated discussion do you:
    a. stick to your guns
    b. look for common ground

26. Is it better to be:
    a. just
    b. merciful

27. At work, is it more natural for you to:
    a. point out mistakes
    b. try to please others

28. Are you more comfortable:
    a. after a decision
    b. before a decision

29. Do you tend to:
    a. say right out what's on your mind
    b. keep your ears open

30. Common sense is:
    a. usually reliable
    b. frequently questionable

31. Children often do not:
    a. make themselves useful enough
    b. exercise their fantasy enough

32. When in charge of others do you tend to be:
    a. firm and unbending
    b. forgiving and lenient

33. Are you more often:
    a. a cool-headed person
    b. a warm-hearted person

34. Are you prone to:
    a. nailing things down
    b. exploring the possibilities

35. In most situations are you more:
    a. deliberate than spontaneous
    b. spontaneous than deliberate

36. Do you think of yourself as:
    a. an outgoing person
    b. a private person

37. Are you more frequently:
    a. a practical sort of person
    b. a fanciful sort of person

38. Do you speak more in:
    a. particulars than generalities
    b. generalities than particulars

39. Which is more of a compliment:
    a. "There's a logical person"
    b. "There's a sentimental person"

40. Which rules you more:
    a. your thoughts
    b. your feelings

41. When finishing a job, do you like to:
    a. tie up all loose ends
    b. move on to something else

42. Do you prefer to work:
    a. to deadlines
    b. just whenever

43. Are you the kind of person who:
    a. is rather talkative
    b. doesn't miss much

44. Are you inclined to take what is said:
    a. more literally
    b. more figuratively

45. Do you more often see:
    a. what's right in front of you
    b. what can only be imagined

46. It is worse to be:
    a. a softy
    b. hard-nosed

47. In hard circumstances, are you sometimes:
    a. too unsympathetic
    b. too sympathetic

48. Do you tend to choose:
    a. rather carefully
    b. somewhat impulsively

49. Are you inclined to be more:
    a. hurried than leisurely
    b. leisurely than hurried

50. At work do you tend to:
    a. be sociable with your colleagues
    b. keep more to yourself

51. Are you more likely to trust:
    a. your experiences
    b. your conceptions

52. Are you more inclined to feel:
    a. down to earth
    b. somewhat removed

53. Do you think of yourself as a:
    a. tough-minded person
    b. tender-hearted person

54. Do you value in yourself more that you are:
    a. reasonable
    b. devoted

55. Do you usually want things:
    a. settled and decided
    b. just pencilled in

56. Would you say you are more:
    a. serious and determined
    b. easy going

57. Do you consider yourself:
    a. a good conversationalist
    b. a good listener

58. Do you prize in yourself:
    a. a strong hold on reality
    b. a vivid imagination

59. Are you drawn more to:
    a. fundamentals
    b. overtones

60. Which seems the greater fault:
    a. to be too compassionate
    b. to be too dispassionate

61. Are you swayed more by:
    a. convincing evidence
    b. a touching appeal

62. Do you feel better about:
    a. coming to closure
    b. keeping your options open

63. Is it preferable mostly to:
    a. make sure things are arranged
    b. just let things happen naturally

64. Are you inclined to be:
    a. easy to approach
    b. somewhat reserved

65. In stories, do you prefer:
    a. action and adventure
    b. fantasy and heroism

66. Is it easier for you to:
    a. put others to good use
    b. identify with others

67. Which do you wish more for yourself:
    a. strength of will
    b. strength of emotion

68. Do you see yourself as basically:
    a. thick-skinned
    b. thin-skinned

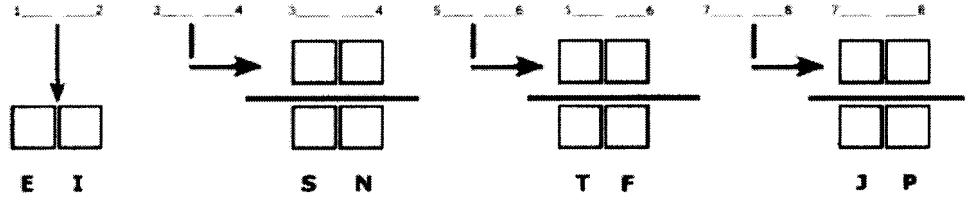
69. Do you tend to notice:
    a. disorderliness
    b. opportunities for change

70. Are you more:
    a. routinized than whimsical
    b. whimsical than routinized

Keirsey Temperament Sorter Answer Sheet

Next to the number for each question, make a check for your answer in column **a** or **b**.

a	b		a	b		a	b		a	b		a	b		a	b		a	b	
1			2			3			4			5			6			7		
8			9			10			11			12			13			14		
15			16			17			18			19			20			21		
22			23			24			25			26			27			28		
29			30			31			32			33			34			35		
36			37			38			39			40			41			42		
43			44			45			46			47			48			49		
50			51			52			53			54			55			56		
57			58			59			60			61			62			63		
64			65			66			67			68			69			70		



Directions for Scoring

1. Counting downward, add the total number of check marks in the first **a** column and write the sum on the line under the column. Do the same for all the **a** columns and all the **b** columns, so that all fourteen lines have a number written on them.
2. Next, copy the number on line 1 into the box below, marked E and the number on line 2 in the box marked I. Then, copy the numbers that are on lines 3 and 4 into the boxes indicated by the arrow. Skip to the next lines, marked 5 and 6, and do the same; then 7 and 8.
3. Now add each number in each box to the number on the line above it. (You'll see that you are adding together all your **a** check-marks and all you **b** check-marks.) Put each sum in the box beneath it, the ones labelled **E, I, S, N, T, F, J** and **P.**
4. Finally, in each pair of lettered boxes, circle the higher number, then circle the letter under it. If you have *the same* numbers in any set of boxes, cross out the letters below and write an **X**. The set of four circled letters represents your type. If you have an **X** in your type it means you are evenly balanced on that scale.

The 16 Combinations

SPs [Artisans]:	ESTP [Promoter], ISTP [Crafter], ESFP [Performer], ISFP [Composer]
SJs [Guardians]:	ESTJ [Supervisor], ISTJ [Inspector], ESFJ [Provider], ISFJ [Protector]
NFs [Idealists]:	ENFJ [Teacher], INFJ [Counselor], ENFP [Champion], INFP [Healer]
NTs [Rationals]:	ENTJ [Fieldmarshal], INTJ [Mastermind], ENTP [Inventor], INTP [Architect]

Appendix 4 – socio-ec(h)o Evaluation Protocol

1. Preparation

Documentation: None

1. Have 4 copies of the ethics form (Be sure to number the forms with the participant numbers & name)
2. Have 4 copies of the questionnaire & pens (Be sure to number the forms with the participant numbers ONLY)
3. Ready the vests and markers
4. Ready the wireless microphones
5. Ready both cameras with tapes inserted. LABEL the tapes properly!
6. Ready the mini-disc reorder with disc inserted. LABEL the disc properly!
7. Ready snacks and water (chips, bottled water, cookies)
8. Turn on the projectors

2. Meet participants: (5 minutes)

Documentation: None

1. Introduce everyone involved in the evaluation
2. Provide an overview of the evaluation session. STATE THAT THEY CAN STOP AT ANYTIME.
3. Provide ethics consent form for signature (Be sure to witness the forms – anyone of us can do that).

3. Introduction and warm-up session (15 minutes)

Documentation: Occasional Handheld camera for presentation purposes

The warm-up should be without the system. A series of simple and quick puzzles are given to the group who try to "solve" it through coordinated body movements (the goal is to connect puzzle solutions with body-movements and to see how the group solves the puzzle, i.e. by talking).

The facilitator introduces the concept of the game by comparing it to the game of "hot and cold." The facilitator leads the group through each of the puzzles by providing the word puzzle and telling the group whether they are close or not by using the words "cold", "cool", "warm", "hot". The facilitator also lets them know when they have completed the puzzle (In the process the facilitator may have to explain that gestures or absolute locations are not significant):

Puzzle 1: "high-low" Solution: look for movement that has the whole group low.

Puzzle 2: "sticks in the mud" Solution: look for the half the group to be standing up straight and still and the other half lying down still. Suggest that the group try to hold the solution for 3 or more seconds so we can be sure that they did not "accidentally" solve it.

Puzzle 3: "All Found around thinly – Fall into a swarm" Solution: a two part in sequence, first all spread out and then clustered together on the ground." Explain that the solution is a two part sequence.

4. Introduction to the system (10 minutes)

Documentation: Occasional Handheld camera for presentation purposes

1. Provide an overview of the game, describing the levels, puzzles, goals at each level, role of the system (mention the circle, and that this is a research prototype that has occasional glitches and can crash; that they need to be attentive to the system and that it has a slight delay in its response)
2. Play audio rewards to the players
3. Prepare each player with the vests and wireless microphones. (ATTENTION: Make sure that the numbers on the microphones and vests match the number the participant number (for example, participant 'C1" is "player 1" and should use the microphone and vest labelled "1"); Make sure that the microphone is securely attached to the back or in the pocket of the players; and the microphones are on.)

5. Playing the game (levels 1-4) (maximum 65 minutes)

Documentation: START to record with the two video cameras and mini-disc recorder; occasional Handheld camera for presentation purposes

1. Perform an audio test to be sure all the audio is being captured and be sure to record it. Ask player 1 to say player 1 so we can later recognize the voice. You will need to adjust the audio levels based on the level meter in the camera. Repeat this process for each player.
2. Players play levels 1-4 until all levels are completed or the maximum time is reached. Explain to them that they have a maximum of 60 minutes to play. If they do not complete the levels, explain the solution to the last level they are on.

6. Break (5 minutes)

Documentation: None PAUSE THE CAMERAS & MINIDISC!

Offer participants, juice, water and chips.

7. Playing the game (levels 5-6) (15 minutes)

Documentation: START to record with the two video cameras and mini-disc recorder; occasional Handheld camera for presentation purposes

Players play levels 5-6 until all levels are completed or the maximum time (15 minutes) is reached.

8. Questionnaire (15 minutes)

Documentation: None

Questionnaires are distributed and players complete the questionnaires. Players are thanked and certificates are given. Ask participants to tell others about participating in the evaluation.

Appendix 5 – socio-ec(h)o Questionnaire

Participant # ____

In the questions below, we ask you to rate an answer along a scale of 1-5 or circle "No opinion," if you cannot rate an answer. Some questions we ask you to provide written answers. If the space provided is not sufficient, please use the back of the sheet.

Please ask if you have any questions about completing this questionnaire.

Section 1

1.1. How confident do you feel about using the system after this evaluation (circle one)?

(not very)	no opinion	1	2	3	4	5	(very)

1.2. Did the system require a large effort to learn (circle one)?

(very large)	no opinion	1	2	3	4	5	(very little)

1.3. Was the system more or less helpful in solving the puzzles as the evaluation progressed (circle one)?

(less)	no opinion	1	2	3	4	5	(more)

1.4. Did your attitude toward the system become more or less positive as the evaluation progressed (circle one)?

(less)	no opinion	1	2	3	4	5	(more)

1.4a Please describe in what respect you feel positively or negatively toward the system.

Section 2

2.1 How well did the system respond to all your actions from the start of solving the puzzle to solving it (circle one)?

(not well)	no opinion	1	2	3	4	5	(very well)

2.2 How well did the system provide useful cues in helping you evaluate how close you were to solving a puzzle (circle one)?

(not well)	no opinion	1	2	3	4	5	(very well)

2.2a Can you describe how the system provided or did not provide useful cues in evaluating how close you were to solving a puzzle?

2.3 How well did the system respond when you were having problems solving a puzzle (circle one)?

(not well)	no opinion	1	2	3	4	5	(very well)

2.3a Can you describe how the system responded when you were having problems solving a puzzle?

2.4 How well did the system support coordinating information and actions within the team (circle one)?

(not well)	no opinion	1	2	3	4	5	(very well)

2.4a Can you describe how your team coordinated information and actions?

2.4b Can you describe how your team was organized and supported leadership?

Section 3

3.1 Do the lights and audio work together to help you in the game (circle one)?

(not well)	no opinion	1	2	3	4	5	(very well)

3.1a Can you describe how the lights and audio worked together to help you in the game?

3.2 How effective was the audio feedback in helping you solve the puzzles (circle one)?

(not effective)	no opinion	1	2	3	4	5	(very effective)

3.3 The audio response in which levels was most effective in helping you evaluate how close you were to solving a puzzle (circle one or more)?

Levels (in order)	no opinion	L1	L2	L3	L4	L5	L6

3.4 Which sounds worked better in the game (circle one or more)?

Sound category	L1(Music) L5(Water)	L2(Music) L6 (Forest)	L3(Rocks)	L4(Fire)

3.5 How well does the virtual system (lights, audio, video) integrate with the physical and architectural space (circle one)?

(not well)	no opinion	1	2	3	4	5	(very well)

3.6 Is the environment created by the system conducive for a game (circle one)?

(not well)	no opinion	1	2	3	4	5	(very well)

3.6a Can you describe how the environment created by the system was conducive or not conducive for a game?

Section 4

4.1 Were parts of the system not helpful or ignored?

4.2 Did the system not respond to some important actions in solving the puzzle?

4.3 How did the system respond when you were unsure or confused?

Section 5

5.1 Any additional comments?

Please note: If you are an **IAT332** student, please remember to write the 1-2 page reflection (as part of your assignment) right after the session when your memory is still fresh.

Thank you for completing this questionnaire and participating in this research!

REFERENCE LIST

Albrecht, D. W., Zukerman, I., & Nicholson, A. E. (1998). Bayesian models for keyhole plan recognition in an adventure game. *User Modeling and User-Adapted Interaction, 8,* 5-47.

Allport, G. W. (1937). *Personality: A psychological interpretation.* New York: Holt.

Allport, G. W., & Odbert, H. S. (1936). Trait-names: A psycho-lexical study. *Psychological Monographs, 47,* No. 211.

Ardissono, L., Goy, A., Petrone, G., Segnan, M., & Torasso P. (2002). Tailoring the recommendation of tourist information to heterogeneous user groups. In S. Reich, M. Tzagarakis, & P. De Bra (Eds.), *Hypermedia: Openness, structural awareness, and adaptivity,* International Workshops OHS-7, SC-3, and AH-3, 2001. Lecture Notes in Computer Science 2266, Berlin, pp. 280-295.

Bartle, R. (1996). Hearts, clubs, diamonds, spades: Players who suit MUDs, http://www.mud.co.uk/richard/hcds.

Bartle Test of Gamer Psychology, GuildCafe, Retrieved May 29, 2008, from http://www.guildcafe.com/bartle.php

Benyon, D., & Murray, D. (1993). Applying user modeling to human-computer interaction design. *Artificial Intelligence Review, 7,* 199-225.

Binder, T., DeMichelis, G., Gervautz, M. Iacucci, G., Matkovic, K., Psik, T., & Wagner, I. (2004). Supporting configurability in a tangibly augmented environment for design students, *Personal and Ubiquitous Computing,* 8(5), 310-325.

Bohn, J., Coroama, V., Langheinrich, M., Mattern, F., & Rohs, M. (2004). Social, economic, and ethical implications of ambient intelligence and ubiquitous computing. *Ambient Intelligence, pp. 5-29.*

Cammann, C., Fichman, M., Jenkins, G. D., & Klesh, J. R. (1983). Assessing the attitudes and perceptions of organizational members In S. E. Seashore, E. E. Lawler P. H. Mirvis & C. Cammann (Eds.), *Assessing Organizational Change*, John Wiley, New York, pp. 71-138.

Carlson, J. G. (1985). Recent assessments of the Myers-Briggs Type Indicator. Journal *of Personality Assessment, 49,* 356-365.

Carlyn, M. (1977). An assessment of the Myers-Briggs Type Indicator. *Journal of Personality Assessment, 41,* 461-473.

Cattell, R. B. (1943). The description of personality: Basic traits resolved into clusters. *Journal of Abnormal and Social Psychology, 38,* 476-506.

Cattell, R. B. (1945a). The description of personality: Principles and findings in a factor analysis. *American Journal of Psychology, 58,* 69-90.

Cattell, R. B. (1945b). The principle trait clusters for describing personality. *Psychological Bulletin, 42,* 129-161.

Chi, M. T. H. (1997). Quantifying qualitative analyses of verbal data: A practical guide. *Journal of the Learning Sciences, 6,* 271–315.

Cohen, P. R., & Perrault, C. R. (1979). Elements of a plan-based theory of speech acts. *Cognitive Science 3,* 177-212.

Corbett, A., McLaughlin, M., & Scarpinatto, K. C. (2000). Modeling student knowledge: Cognitive tutors in high school and college. *User Modeling and User-Adapted Interaction, 10(2-3),* 81-108.

Costa, P. T., & McCrae, R. R. (1985). The *NEO personality inventory manual.* Odessa, FL: Psychological Assessment Resources.

Costa, P. T., & McCrae, R. R. (1992). *NEO PI-R professional manual.* Odessa, FL: Psychological Assessment Resources.

Craiger, J. P. (1997). Technology, organizations, and work in the 20th century. *The Industrial-Organizational Psychologist, 34(3),* 89-96.

Dourish, P. (2001). *Where the action is: The foundations of embodied interaction.* MIT Press.

Ducatel K., Bogdanowicz M., Scapolo F., Leijten J., & Burgelman J.-C. (2001) *ISTAG Scenarios for Ambient Intelligence in 2010.* European Commission, Community Research.

Fischer, G. (2001). User modeling in human-computer interaction. *User Modeling and User-Adapted Interaction, 11,* 65-86.

Goldberg, L. R. (1981). Language and individual differences: The search for universals in personality lexicons. In Wheeler (Ed.), *Review of Personality and social psychology, 1,* 141-165. Beverly Hills, CA: Sage.

Goldberg, L. R. (1990). An alternative "description of personality": The big-five factor structure. *Journal of Personality and Social Psychology, 59,* 1216-1229.

Goldberg, L. R. (1992). The development of markers for the Big-Five factor structure. *Psychological Assessment, 4,* 26-42.

Grudin, J. (2002). Group Dynamics and Ubiquitous Computing. *Communications of the ACM, 45,* 74-78.

Hatala, M., & Wakkary, R. (2005). Ontology-based user modeling in an augmented audio reality system for museums. *User Modeling and User-Adapted Interaction, 15,* 339–380, Springer.

Hayes, P. J., & Rosner, M. A. (1976). ULLY: A program for handling conversations. *AISB Summer Conference,* Edinburgh, Scotland.

Hornecker, E., & Buur, J. (2006). Getting a grip on tangible interaction: A framework on physical space and social interaction. *CHI 2006,* Montreal, Canada, pp. 437-446.

Hirashima, T., Hachiya, K., Kashihara, A., & Toyoda, J. (1997). Information filtering using user's context on browsing in hypertext. *User Modeling and UserAdapted Interaction, 7,* 239-256.

Hogg, L., & Jennings, N. R. (2000). Variable sociability in agent-based decision making. In N. Jennings and Y. Lesperance (Eds.) *Intelligent Agents VI* (LNAI-1757), pp. 305-318.

Huizinga, J. (1964). *Homo Ludens: A study of the play-element in culture.* Beacon Press, Boston.

Jacobsen, T. (2004). Individual and group modelling of aesthetic judgment strategies. *British Journal of Psychology, 95,* 41-56.

Jameson, A., Baldes, S., & Kleinbauer, T. (2003). Generative models of group members as support for group collaboration. *User Modeling*, Johnstown, PA.. pp. 1–14.

Jiang, Y., Wakkary, R., Hatala, M., Droumeva, M., & Hosseini, M. (2008). *Exploring group user modelling in an embedded interaction environment utilizing Keirsey Temperaments as a classifier.* Submitted for publication.

John, O. P., Angleitner, A., & Ostendorf, F. (1988). The lexical approach to personality: A historical review of trait taxonomic research. *European Journal of Personality, 2,* 171-203.

John, O. P., Donahue, E. M., & Kentle, R. L. (1991). *The Big Five Inventory--Versions 4a and 54.* Berkeley: University of California, Berkeley, Institute of Personality and Social Research.

John, O. P., & Srivastava, S. (1999). The Big Five Trait taxonomy: History, measurement, and theoretical perspectives. In L. A. Pervin & Oliver P. John (Eds.), *Handbook of personality: Theory and research* (2nd ed., pp. 102–139). New York: Guilford Press.

Jung, C. G. (1921/1971). *Psychological Types,* Princeton University Press.

Kaplan, C., Fenwick, J., & Chen, J. (1993). Adaptive hypertext navigation based on user goals and context. *User Modeling and User-Adapted Interaction,* 3, 193-220.

Kaptelinin, V., Nardi, B. A., & Macaulay, C. (1999). The activity checklist: A tool for representing the "space" of context. *ACM interactions. 6,* 27-39.

Kay, J. (2001). Learner control. *User Modeling and User-Adapted Interaction, 11,* 111-127.

Keirsey, D. (1978). *Please Understand Me.* Del Mar, CA, Prometheus Nemesis Books.

Keirsey, D. (1998). *Please understand me II: Temperament, character, intelligence.* Del Mar, CA, Prometheus Nemesis.

Kelly, K. R., & Jugovic, H. (2001). Concurrent Validity of the Online Version of the Keirsey Temperament Sorter II. *Journal Of Career Assessment, 9,* 49-59.

Kobsa, A., & Wahlster, W. (1989). *User Models in Dialog Systems.* Springer.

Kudenko, D., Bauer, M., & Dengler, D. (2003). Group decision making through mediated discussions. *Lecture Notes in Computer Science, 2702,* 238-247.

Lincoln, Y., & Guba, E. (1985). Naturalistic inquiry. 37. New York: Sage.

Linton, F., & Schaefer, H. P. (2000). Recommender systems for learning: Building user and expert models through long-term observation of application use. *User Modeling and User-Adapted Interaction, 10,* 181-207.

Malone, T. W., & Lepper, M. R. (1987). Making learning fun: A taxonomy of intrinsic motivations for learning, in R. Snow & M. Farr (Eds.), *Aptitude, learning, and instruction: Cognitive and affective process analyses,* Lawrence Erlbaum, Hillsdale, N.J.

Maslow, A. H. (1943). A theory of human motivation. *Psychological Review, 50,* 370-396.

Masthoff, J. (2002). Modeling a group of television viewers. *Future TV: Adaptive Instruction in Your Living Room Workshop,* San Sebastian, Spain, pp. 34-42.

Masthoff, J. (2004). Group modeling: Selecting a sequence of television items to suit a group of viewers. *User Modeling and User-Adapted Interaction 14,* 37-85.

Mathé, N., & Chen, J. R. (1996). User-centered indexing for adaptive information access. *User Modeling and User-Adapted Interaction, 6,* 225-261.

McCarthy, J., & Anagnost, T. (1998). MusicFX: An arbiter of group preferences for computer supported collaborative workouts. ACM 1998 Conference on CSCW, Seattle, WA, pp. 363-372.

Molin, E. J. E. (1997). Modeling group preferences using a decompositional preference approach. *Group Decision and Negotiation, 6,* 339-350.

Mudrack, P. E. (1989). Group cohesiveness and productivity: A closer look. *Human Relations, 42*, 771-785.

Myers, I. B. (1962). *The Myers-Briggs Type Indicator manual*, Princeton, N.J.: Educational Testing Service.

Myers, I. B., & McCaulley, M. H. (1985). *Manual: A guide to the development and use of the Myers-Briggs Type Zndicaror.* Palo Alto, CA, Consulting Psychologists Press.

MyersBriggs.org. (2007). Retrieved June 16, 2007, from http://www.myersbriggs.org/.

Newell, S. C. (1997). User models and filtering agents for improved internet information retrieval. *User Modeling and UserAdapted Interaction, 7,* 223-237.

Norman, W. T. (1963). Toward an adequate taxonomy of personality attributes: Replicated factor structure in peer nomination personality ratings. *Journal of Abnormal and Social Psychology, 66,* 574-583.

Norman, W. T. (1967). *2800 personality trait descriptors: Normative operating characteristics for a university population.* Department of Psychology, University of Michigan.

O'Conner, M., Cosley, D., Konstan, J. A., & Riedl, J. (2001). PolyLens: A recommender system for groups of users. ECSCW 2001, Bonn, Germany, pp. 199-218.

Rich, E. (1979). User modeling via stereotypes. *Cognitive Science, 3,* 329-354.

Saucier, G. (1997). Effects of variable selection on the factor structure of person descriptors. *Journal of Personality and Social Psychology, 73,* 1296-1312.

Saucier, G., & Goldberg, L. R. (1996a). The language of personality: Lexical perspectives on the five-factor model. In J. S. Wiggins (Ed.), *The five-factor model of personality: Theoretical perspectives.* (pp. 21-50): Guilford Press, New York, NY.

Saucier, G., & Goldberg, L. R. (1996b). Evidence for the Big Five in analyses of familiar English personality adjectives. *European Journal of Personality, 10,* 61-77.

Seashore, S. (1954). *Group Cohesiveness in the Industrial Work Group.* University of Michigan Press, Ann Arbor.

Stock, O., Zancanaro, M., Busetta, P., Callaway, C., Krüger, A., Kruppa, M., Kuflik T., Not, E., & Rocchi, C. (2007). Adaptive, intelligent presentation of information for the museum visitor in PEACH. *User Modelling and User-Adapted Interaction, 17,* 257–304.

Suchman, L. (1987). *Plans and situated actions: The problem of human machine communication* (pp.8). Cambridge University Press, Cambridge, U.K.

Suebnukarn, S., & Haddawy, P. (2006). Modeling individual and collaborative problem-solving in medical problem-based learning. *User Modelling and User-Adapted Interaction, 16,* 211–248.

Tucker, I. F., & Gillespie, B. V. (1993). Correlations among three measures of personality type. *Perceptual and Motor Skills, 77,* 650.

Vildjiounaite, E., & Kallio, S. (2007). A layered approach to context-dependent user modelling. G. Amati, C. Carpineto, and G. Romano (Eds.): *ECIR 2007, LNCS 4425,* 749- 752.

Wakkary, R. (2006). Evaluating context in tangible and embedded interaction. Unpublished manuscript.

Wakkary, R., Hatala, M., Jiang, Y., Droumeva M., & Hosseini, M. (2008). Making sense of group interaction in an ambient intelligent environment for physical play. *Tangible Embedded Interaction 2008*, Bonn, Germany, pp. 179-186.

Wakkary, R., Hatala, M., Lovell, R. & Droumeva, M. (2005). An ambient intelligence platform for physical play. *ACM Multimedia 2005,* Singapore, pp. 764-773.

Wakkary, R., Hatala, M., Lovell, R., Droumeva, M., Antle, A., Evernden, D. & Bizzocchi, J. (2005). socio-ec(h)o: Ambient intelligence and gameplay. In S. Castell & J. Jenson (Eds.), *Digital Games Research Association's International Conference 2005,* Vancouver, Canada, pp. 217-226.

Yee, N. (2004). Unmasking the avatar: The demographics of MMO player motivations, in-game preferences, and attrition, Retrieved May 28, 2008, from http://www.gamasutra.com/resource_guide/20040920/yee_pfv.htm

Yee, N. (2005). Motivations of Play in MMORPGs. In Castell, S. & Jenson, J. (Eds.), *Digital Games Research Association's International Conference 2005*, Vancouver, Canada, pp. 157-164.

Zack, M. H., & McKenney,J. L. (1995). Social Context and Interaction in Ongoing Computer-Supported Management Groups, *Organization Science, 6,* 394-422.